A multiresolution random field model for estimating fossil-fuel CO_2 emissions

J. Ray[1], V. Yadav[2], J. Lee[1], A. M. Michalak[2], S. Lefantzi[1] and Bart van Bloemen Waanders[3]
[1]Sandia National Laboratories, Livermore, CA, [2]Carnegie Institution for Science, Stanford and [3]Sandia National Laboratories, Albuquerque, NM

OBJECTIVE

To develop an estimation method for fossil-fuel CO_2 (ffCO2) emissions

- Construct a random field model for non-stationary ffCO2 emission fields
- Design a reconstruction method to infer ffCO2 emissions from limited concentration measurements of ffCO2

Demonstration problem

- Estimate weekly-averaged ffCO2 emissions in $\mathcal{R}$, the lower 48 states of US, at 1° x 1° resolution, for a year
- Pseudo-data or synthetic observations generated using the Vulcan inventory of ffCO2 emissions (2002)
- Examine accuracy and spatial fidelity of the inferred ffCO2 emissions

BACKGROUND

CO_2 flux estimation on a grid

- Currently only done for biogenic CO_2 fluxes
- Obtained by optimizing an objective function J

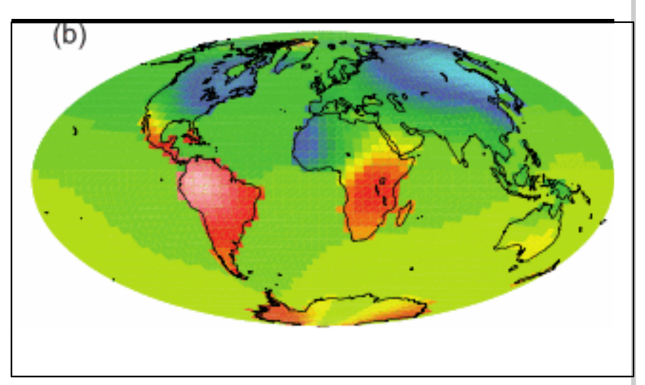
$$J = (y^{obs} - Hs)^T R^{-1} (y^{obs} - HS) + (s - s_{pr})^T Q^{-1} (s - s_{pr})$$

  - s : CO2 fluxes being estimated
  - y^{obs} : observations of CO_2 concentrations at a few sites
  - s_{pr}: prior belief re fluxes from a process-based model like CASA
  - H : atmospheric transport model
  - R: diagonal matrix of measurement error estimates (variances)
  - Q: covariance matrix for the multivariate Gaussian field model for $(s - s_{pr})$
- Inferred fluxes represent a balance between observations and prior beliefs

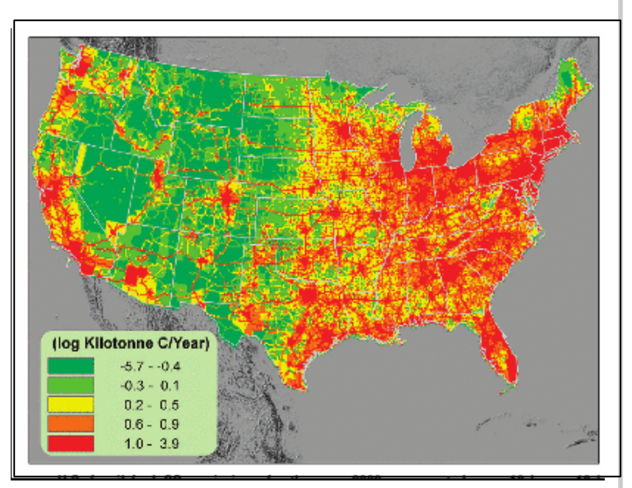
This method cannot be used for estimating ffCO2 emission

- s_{pr}, obtained from inventories, tends to be inaccurate at fine resolutions
  - Especially for developing countries
- A multivariate Gaussian approximation for $(s - s_{pr})$ is unlikely to be accurate (Q is hard to model)
  - Biospheric fluxes are smoothly distributed in space (right, above)
  - ffCO2 emissions are a lot more complex.

Need to construct a new random field model that can efficiently represent ffCO2 emission fields



Global biospheric fluxes

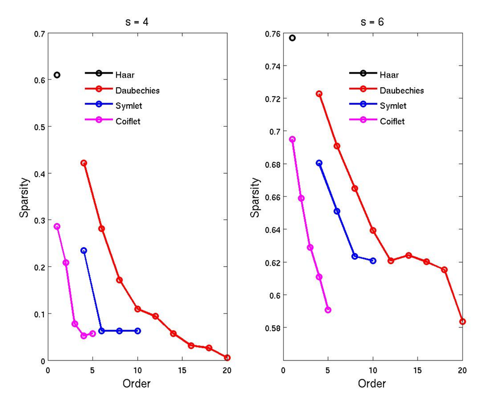


US ffCO2 emissions from the Vulcan inventory

RANDOM FIELD MODEL

Model ffCO2 emissions using wavelets in the box [24.5N, -63.5W] [87.5N, -126.5W]

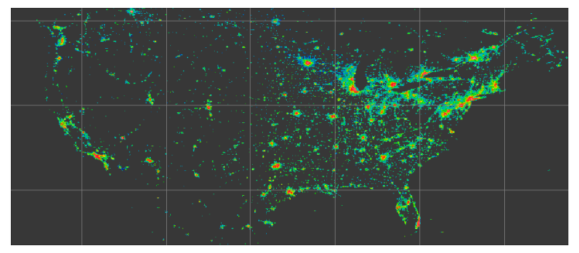
- Modeled Vulcan emissions using Daubechies, Symmlets, etc. on a 1° resolution mesh i.e. 64 x 64
- Haar wavelets proved to provide the most *compressible* representation



Sparsity of Vulcan emission wavelet coefficients at different scales

Reducing dimensionality from 64^2

- ffCO2 emissions occur near regions of human activity
- Used radiance-calibrated images of lights at night to identify such regions
- And removed Haars modeling dark/uninhabited areas
- Retained 1031/64^2 wavelets



Radiance-calibrated image of lights at night for the US

We call this the Multiscale Random Field (MsRF) model for ffCO2 emissions

Converted the MsRF into a prior ffCO2 emission f_{pr} model

$$f_{pr} = c \sum_{w_i^{(X)} \in W^s} w_i^{(X)} \phi_i$$

- W^s is the set of wavelets in the MsRF, $w^{(X)}_i$ are the Haar wavelet coefficients for the lights-at-night radiances
- C is a scaling constant that renders $\mathcal{R}$-integrated emissions equal to EDGAR emissions

SPARSE RECONSTRUCTION

Inverse problem

- Model: $y^{obs} = Hf + \varepsilon$, $f = \Phi_R w$, $w = \{w_i\}$, $w_i \in W^s$
- y^{obs} may not inform all w_i
- Use sparse reconstruction to remove w_i that cannot be estimated

$$\min \; \|w\|_1, \quad subject\ to \left\| y^{obs} - H\Phi_R w \right\|_2 < \varepsilon$$

Impose boundary conditions using compressive sampling

- Wavelets in W^s model emissions in a rectangle
- Non-zero ffCO2 emissions are restricted to $\mathcal{R}$
- Permute the wavelets of Φ to separate out emissions in $\mathcal{R}$ and outside – $f_{\mathcal{R}} = \Phi_{\mathcal{R}} w$ and $f'_{\mathcal{R}} = \Phi'_{\mathcal{R}} w$

$$\begin{bmatrix} y^{obs} \\ 0 \end{bmatrix} = \begin{bmatrix} Hf_R \\ Uf'_R \end{bmatrix} = \begin{bmatrix} H\Phi_R \\ U\Phi'_R \end{bmatrix} w$$

- U is a random matrix – uniform random ensemble
- Randomly project ffCO2 emissions outside $\mathcal{R}$ and set them to zero
  - No. of rows in U << number of grid-cells outside $\mathcal{R}$ (about 1/10); a much more efficient way of enforcing zero ffCO2 emissions outside $\mathcal{R}$

PRIOR MODELS AND NON-NEGATIVITY

Solution for non-negative ffCO2 emissions

- Incorporate priors

$$y = Hf_R = H\Phi_R diag(cw_i^{(X)})w', w_i^{'} = w_i / cw_i^{(X)}$$

- Solve, using StOMP [1]

$$\min \|w\|_1 \quad subject\ to \left\| \begin{bmatrix} y^{obs} \\ 0 \end{bmatrix} - \begin{bmatrix} H\tilde{\Phi}_R \\ U\tilde{\Phi}'_R \end{bmatrix} w' \right\|_2 < \varepsilon$$

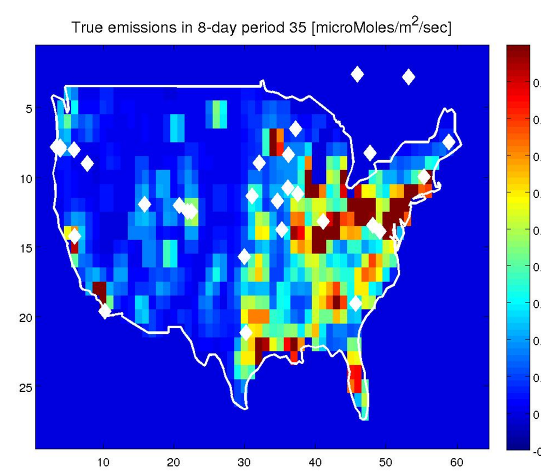
- StOMP solution does not give non-negative ffCO2 emissions; a few grid cells end up with small, negative ffCO2 emissions
- Designed an iterative post-processing method to enforce non-negativity on the StOMP solution [2]

RESULTS

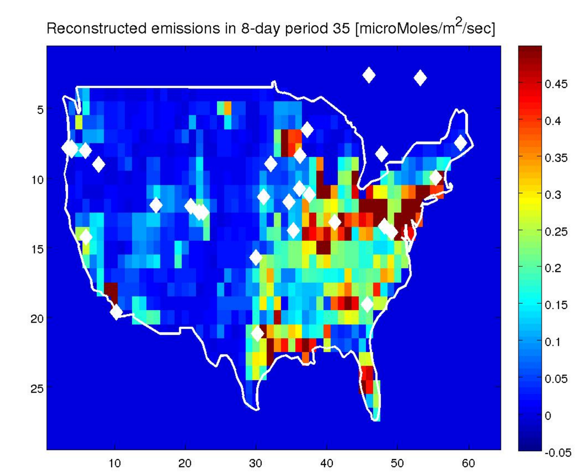
Generation of synthetic observations y^{obs}

- Generated using emissions from the Vulcan inventory, coarsened to 1° x 1° resolution
- Added an uncorrelated noise $\varepsilon \sim N(0, \sigma^2)$, $\sigma = 0.1$ ppmv

Estimated ffCO2 emissions for a week in August 2002



True emissions

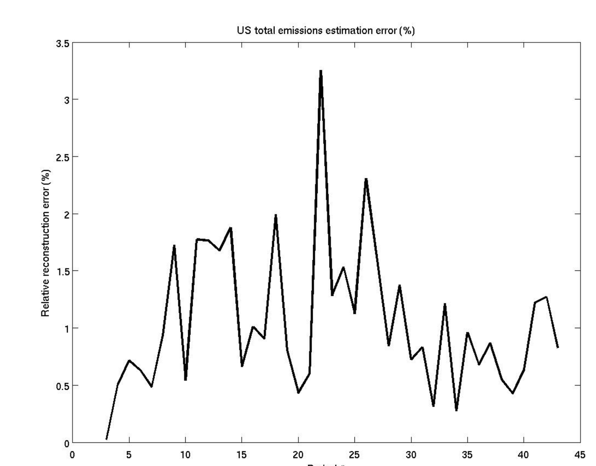


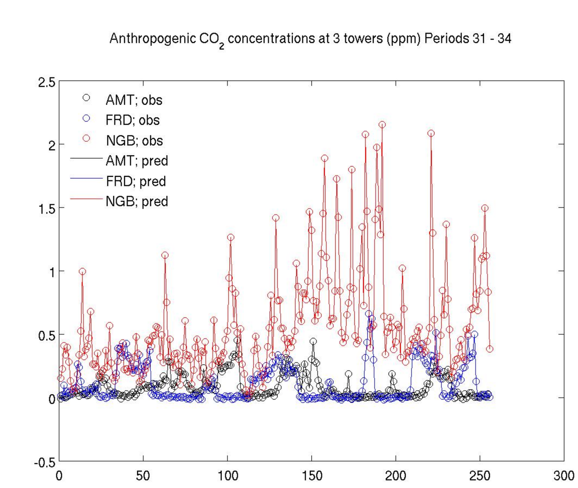
Reconstructed emissions

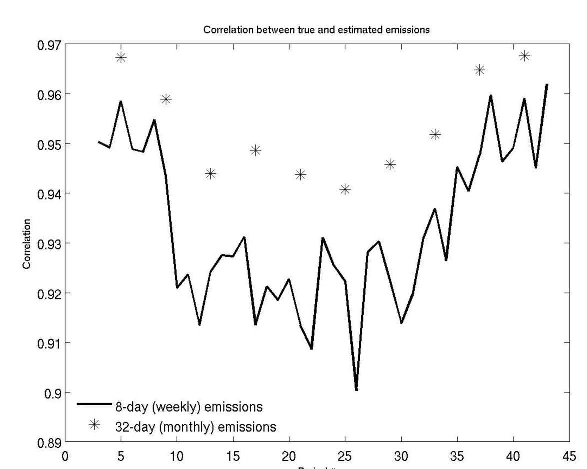
- White diamonds indicate measurement sites
- Estimates in the Northeast (with many measurement sites) are accurate
- Emission estimates in the West are not very accurate (not many measurement sites)

Accuracy of estimates, aggregated over $\mathcal{R}$

- Right: Relative error w.r.t. Vulcan emissions around 4%; becomes worst in spring
- Bottom right: Very high spatial correlations between estimated and Vulcan emissions
- Bottom: Estimated emissions can reproduce observations at various sites



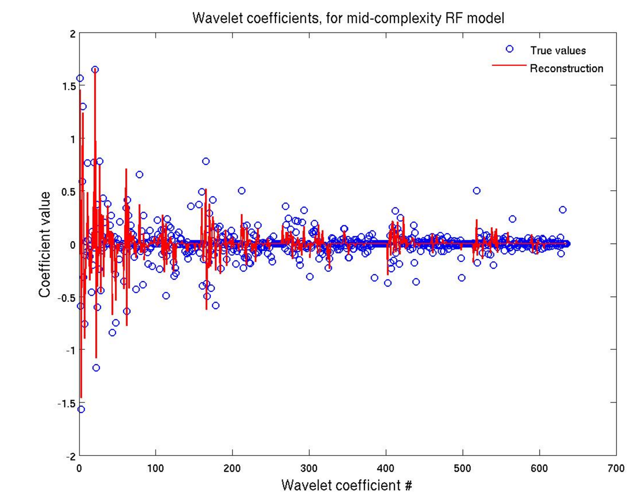




NUMERICAL PERFORMANCE

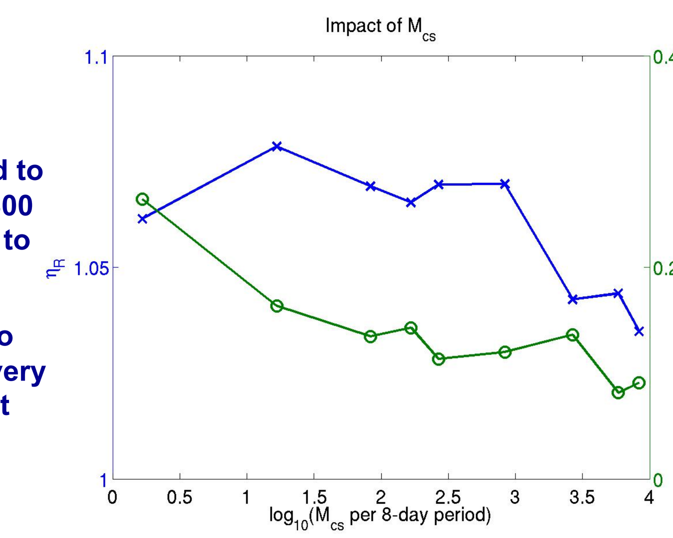
Enforcement of sparsity

- StOMP finds that only about 50% of the wavelet coefficients can be estimated from y^{obs}
- These are the large support wavelets and those near the measurement sites



Limiting emissions in $\mathcal{R}$ via compressive sampling

- About 300 compressive samples per week are need to limit emissions within $\mathcal{R}$ (300 extra constraint equations to be solved by StOMP)
- A naïve way to enforce zero emissions outside $\mathcal{R}$ (in every grid-cell) would need about 3000 constraints



CONCLUSIONS

- Estimation of ffCO2 emissions required a multiresolution random field model to capture its non-stationary behavior in space.
  - Used Haar wavelets in the random field model
  - Reduced its dimensionality using images of lights at night
- Designed a sparse reconstruction method that fits the random field model and preserves the non-negative nature of ffCO2 emssions
  - Could estimate only about 50% of the wavelets from limited observations
  - The method is an extension of StOMP
- Devised a method based on compressive sampling to limit ffCO2 emissions inside an irregular region $\mathcal{R}$ while using a model for rectangular random fields

Acknowledgements

The project was funded by Sandia National Laboratories LDRD (Laboratory Directed Research and Development program, administered by the Geosciences IA.

References

1. D. L. Donoho et al, "Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit", IEEE Transactions on Information Theory, 58(2):1094-1121, 2012.
2. J. Ray et al, "A multiresolution spatial parametrization for the estimation of fossil-fuel carbon dioxide emissions via atmospheric inversions", Sandia Technical Report, SAND2013-2919. Printed April 2013. Sandia National Laboratories, Livermore

For additional information, please contact:

J. Ray, Sandia National Laboratories, jairay@sandia.gov