Wavelet-based reconstruction of fossil-fuel CO_2 emissions from sparse measurements

J. Ray[1], V. Yadav[3], B. van Bloemen Waanders[2], A. Michalak[3] and S. A. McKenna[2]

jairay [at] sandia [dot] gov

[1]Sandia National Laboratories, Livermore, CA,
[2]Sandia National Laboratories, Albuquerque, NM
[3]Carnegie Institution for Science, Stanford, CA

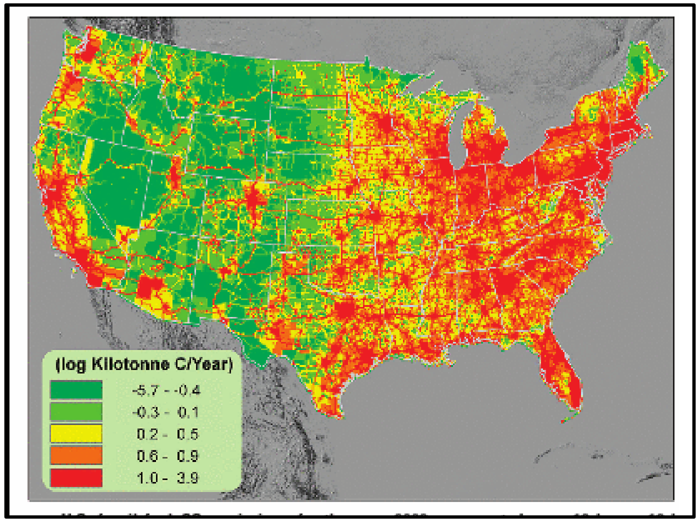
Funded by the LDRD program in Sandia National Labs, Livermore, CA
SAND2012-7811P

The ffCO2 estimation problem

- **Aim:** Develop a technique to estimate anthropogenic (fossil-fuel) CO_2 emissions from sparse observations
- **Motivations:**
  - An alternative to estimating ffCO_2 emission using bottom-up (economic model) techniques
    - Databases: Vulcan (2002, US-only); EDGAR, CDIAC (ORNL) etc
  - Can provide independent verification in case of ffCO_2 abatement treaties
- **How is it done?**
  - Measure CO_2 concentrations in flasks at measurement sites; also column-averaged satellite measurements
  - Use an atmospheric transport model to invert for source locations

Background on $ffCO_2$ inversion

- Unlike biogenic CO2 emissions, anthropogenic emissions are very nonstationary and multiscale
  - The challenge is in devising a low-dimensional random field model, for use in the inverse problem
- NOAA runs a set of towers which measure CO_2 concentrations every 3 hours – main data source
  - Meant for biospheric fluxes (far from cities); about 100 today
  - We will assume that they can measure $^{14}CO_2$ and $^{12}CO_2$ separately



- Picture of 2002 $ffCO_2$ emissions from Vulcan database

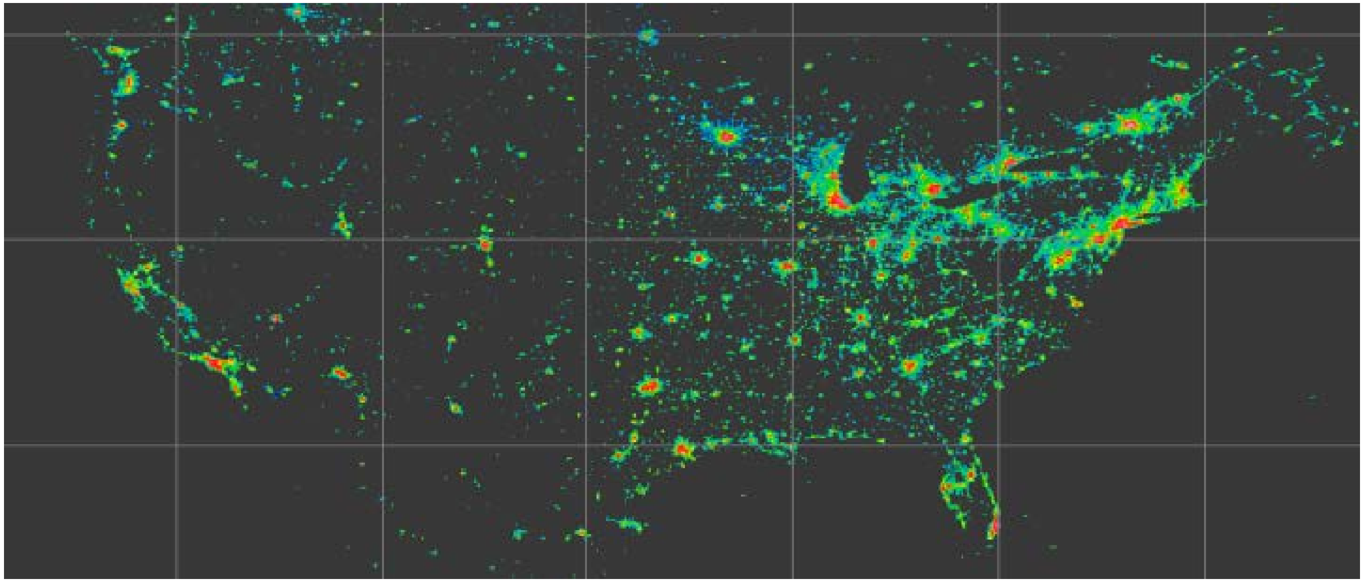
Random field model for ffCO2

- An emission field on a dyadic grid, modeled with wavelets

$$e(x) = \sum_{s=1}^{N} \sum_{i=1}^{2^s} \sum_{j=1}^{2^s} w_{s,i,j} \phi_{s,i,j}(x) = \mathbf{\Phi w}$$

ϕ are orthogonal wavelets (bases)

- Conjecture: $w_{s,i,j}$ are mostly zero (i.e., is sparse)
  - Can be hugely sparsified by pictures of lights at night
  - The remaining could be estimated using sparse data (perhaps)



Sandia National Laboratories

Sparsity enforced reconstruction

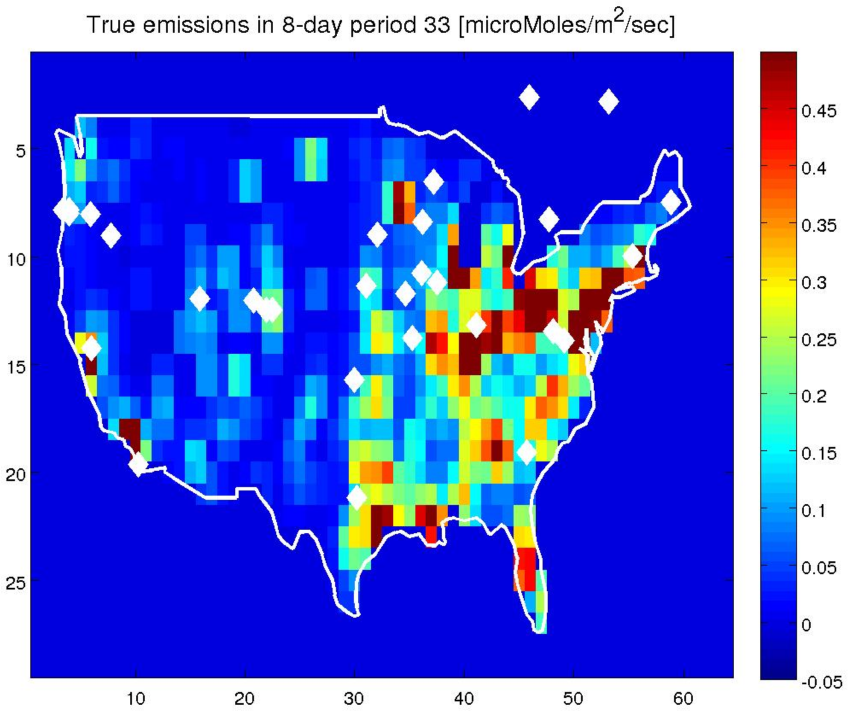
- Time-dependent CO_2 concentration measurements, y^{obs} ,at a sampling location, due to ffCO_2 emissions $e(\mathbf{x})$ modeled as

$$y^{obs} = \mathbf{H}e = \mathbf{H}\boldsymbol{\Phi}w, \quad e(\mathbf{x}) = \boldsymbol{\Phi}w$$

- **H** = transport matrix, obtained using WRF
- $\boldsymbol{\Phi}$, matrix; columns are wavelets; w are the wavelet weights
- Sparsity-enforced reconstruction:
  - Cannot estimate all elements of w, even after sparsifying with nightlights
  - minimize $| y^{(obs)} - [\mathbf{H}][\boldsymbol{\Phi}]w |_2 + |w|_1$
- Many algorithms to solve this – usually formulated as
  - Minimize $|\mathbf{w}|_1$ under the constraint $| \mathbf{y}^{(obs)} - [\mathbf{A}][\boldsymbol{\Phi}]\mathbf{w} |_2 < \varepsilon_s$
  - We use StOMP – Stagewise Orth. Matching Pursuit

Setting up the synthetic data inversion

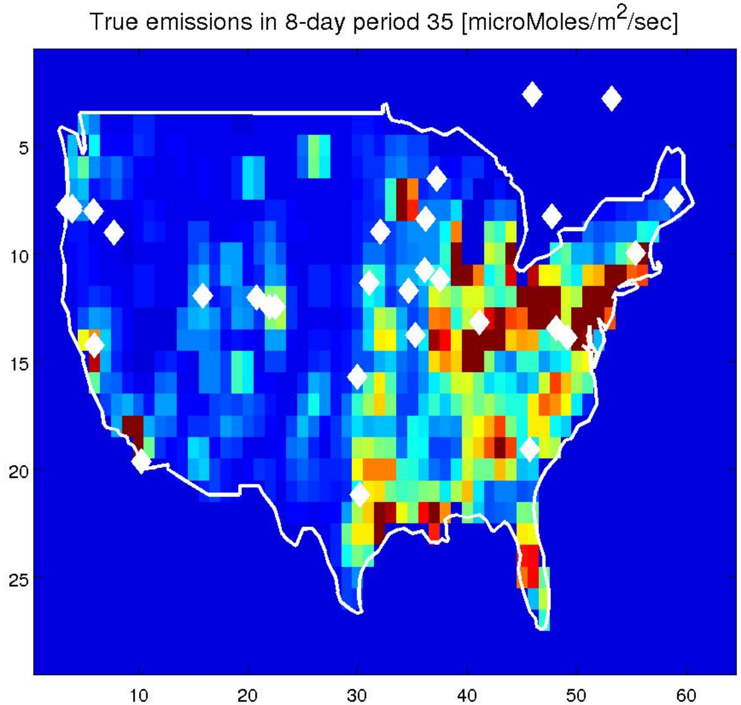
- True emissions – Vulcan database for US, 2002
  - Used to generate CO_2 concentrations at towers
  - 3 hr temporal resolution
- Nightlight images (for 1997) – used to remove wavelets from "dark" areas
- Emissions discretized on a grid
  - 1 degree spatial resolution Fluxes assumed to be constant over 8-day periods ("a week")



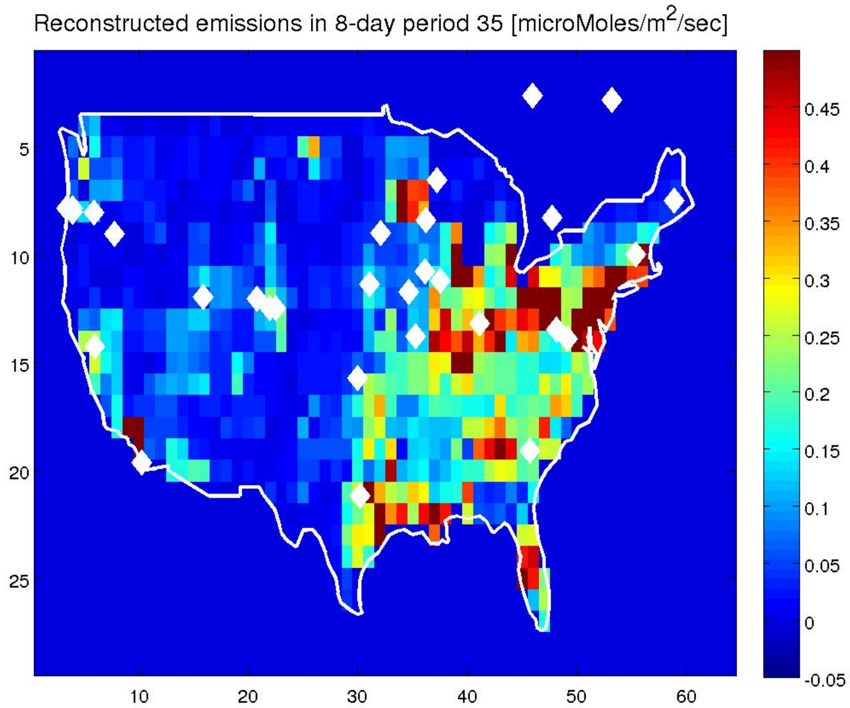
Emissions for a week in August 2002 (Vulcan database, 1 deg resolution)

Sandia National Laboratories

How good is the reconstruction?



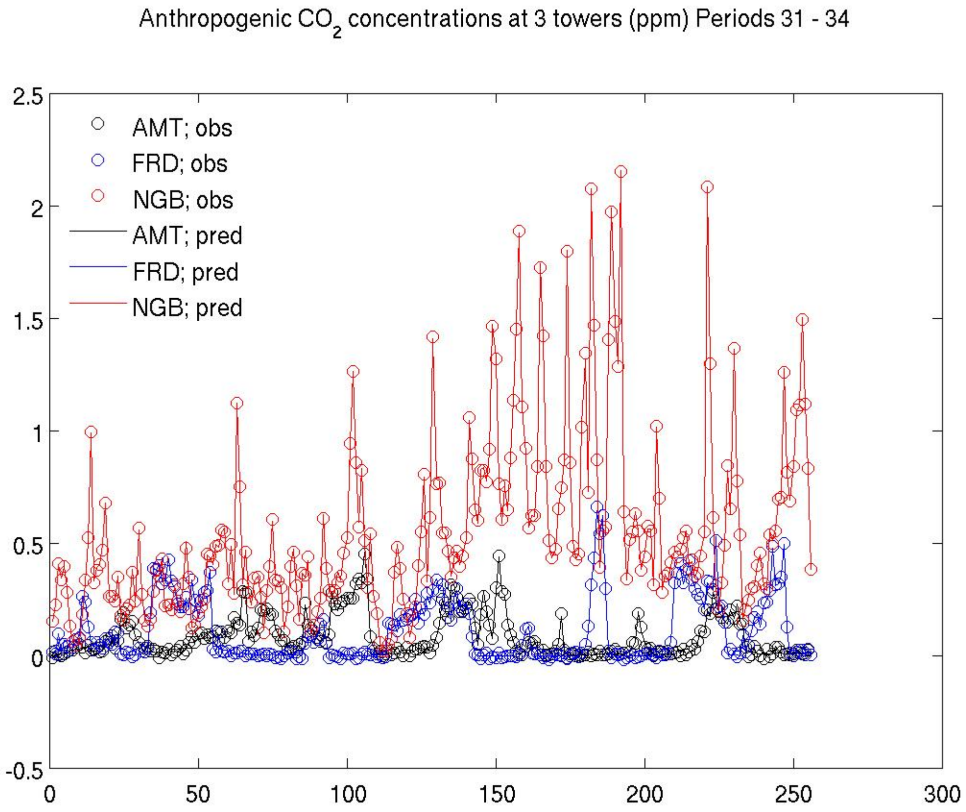
True emissions



Reconstructed emissions

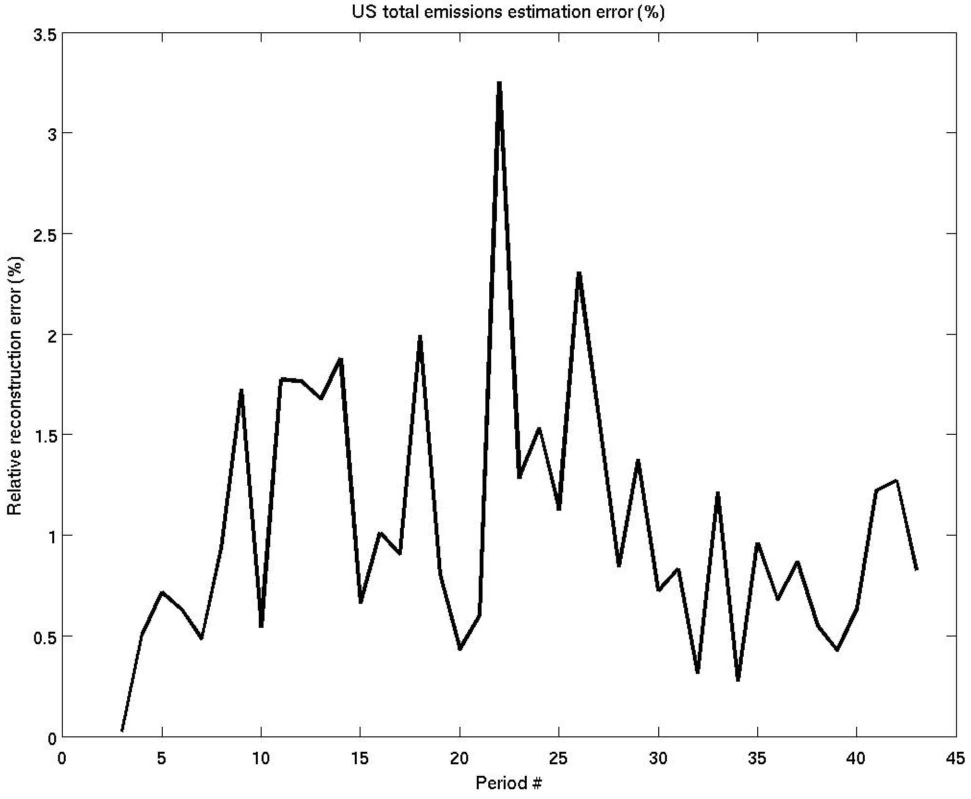
- A week in September 2002

Can we reproduce tower observations?



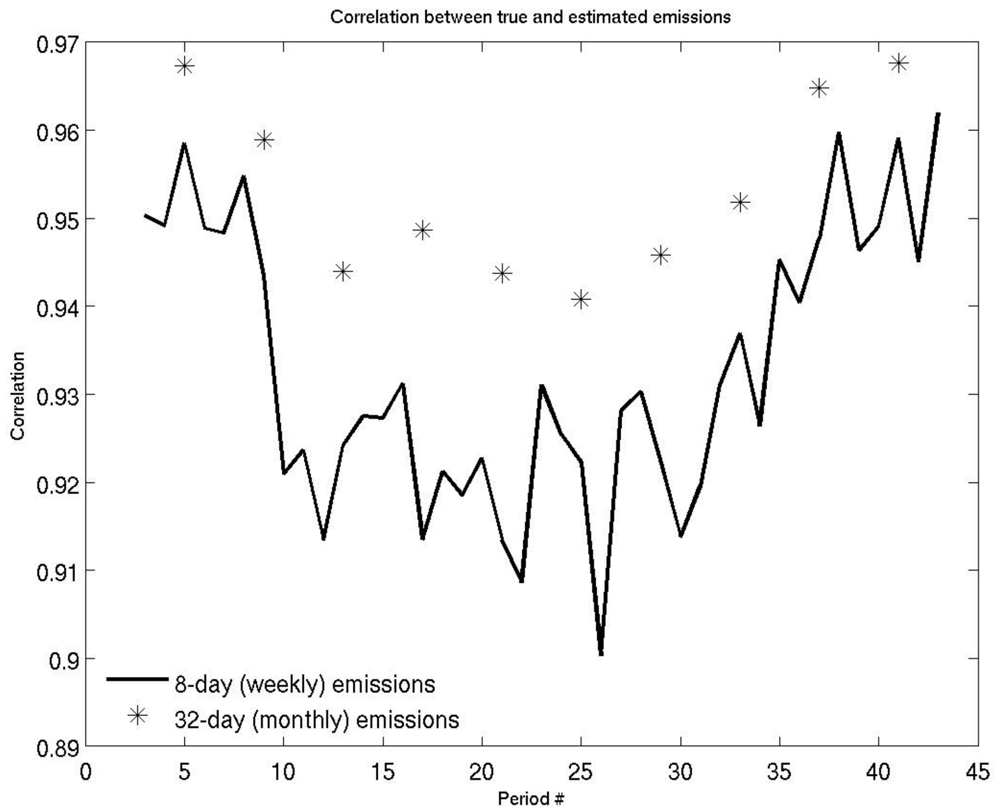
- Tower concentration predictions with reconstructed fluxes (only 3 weeks)
  - Symbols : observations used in the inverse problem.

Reconstruction error in total US emission



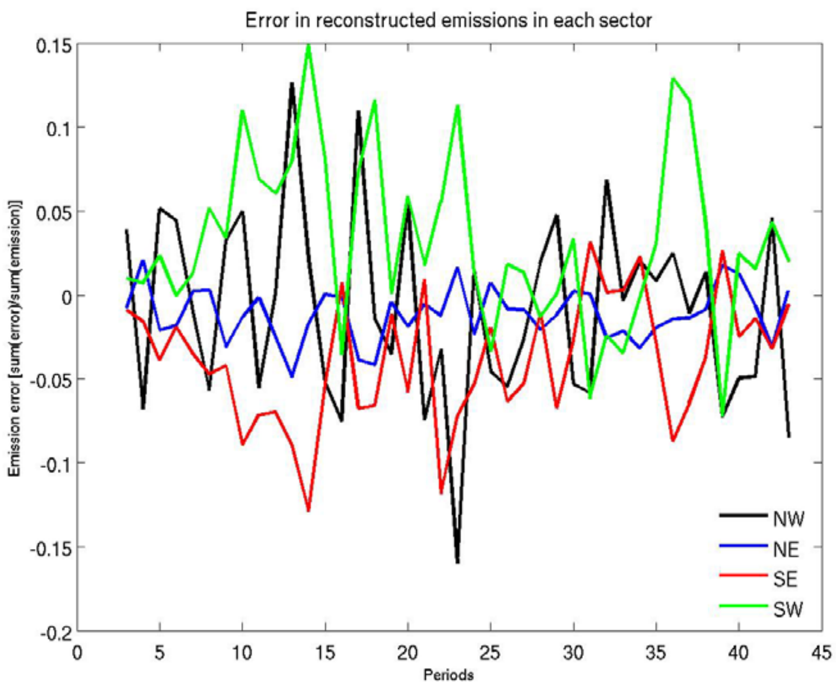
- We get about 3.5% error, worst case

Is the spatial distribution correct?

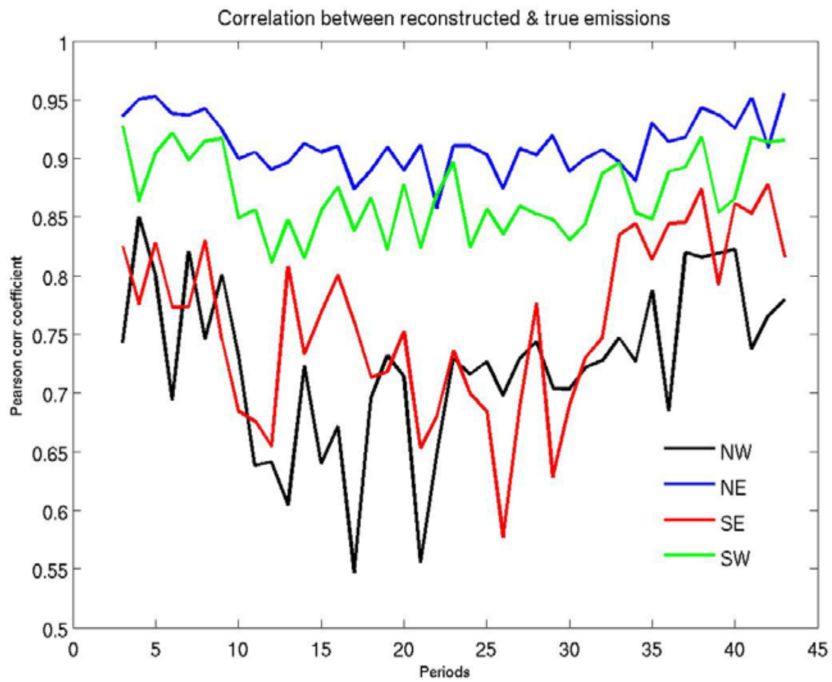


- The spatial distribution of emissions is very close to truth
- Especially, if considering monthly fluxes

Which parts of US are well estimated?



Errors



Spatial Correlation

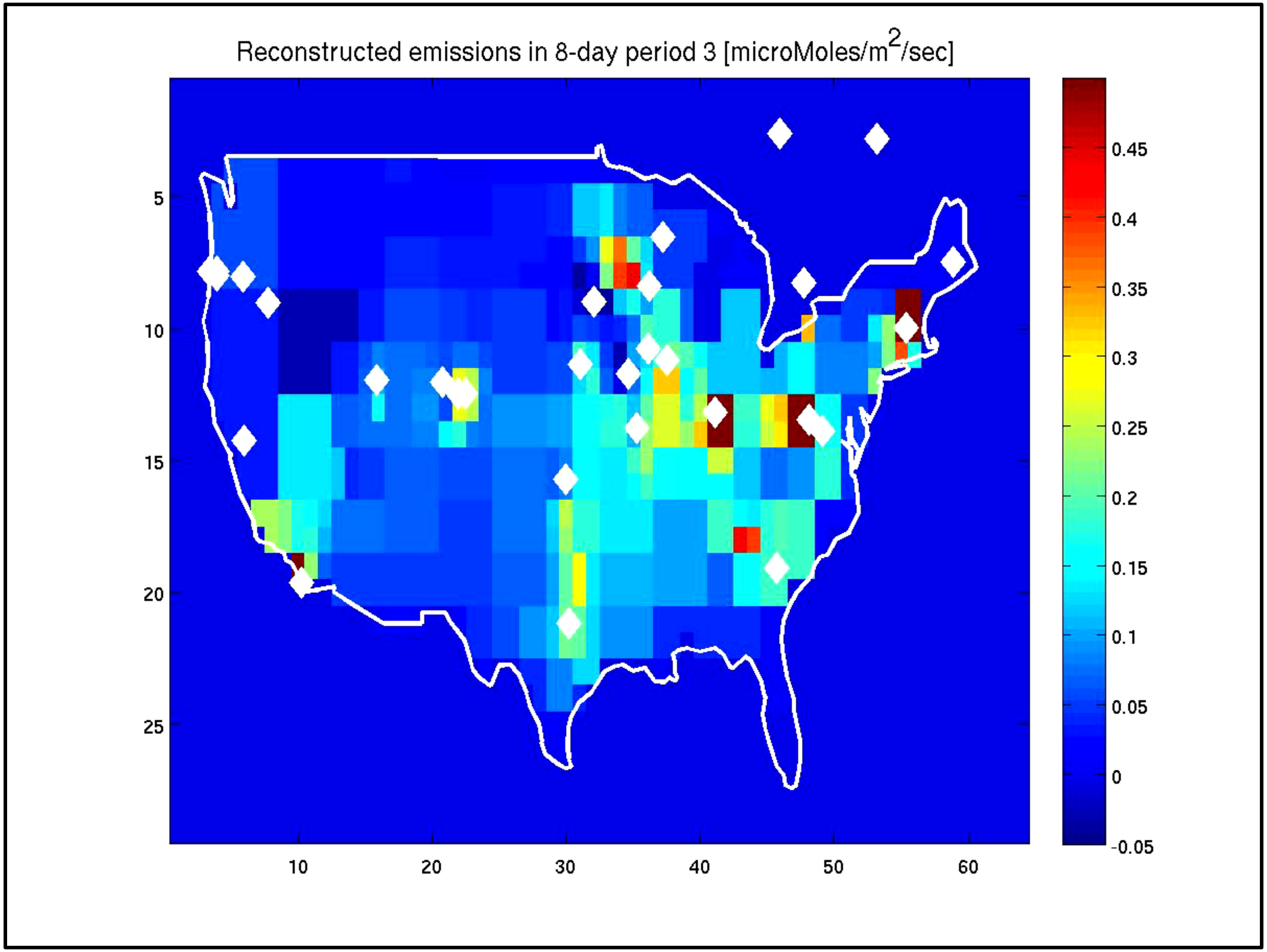
- The NE has the lowest errors and best correlations
- The NW is generally the worst estimated

Sandia National Laboratories

Conclusions

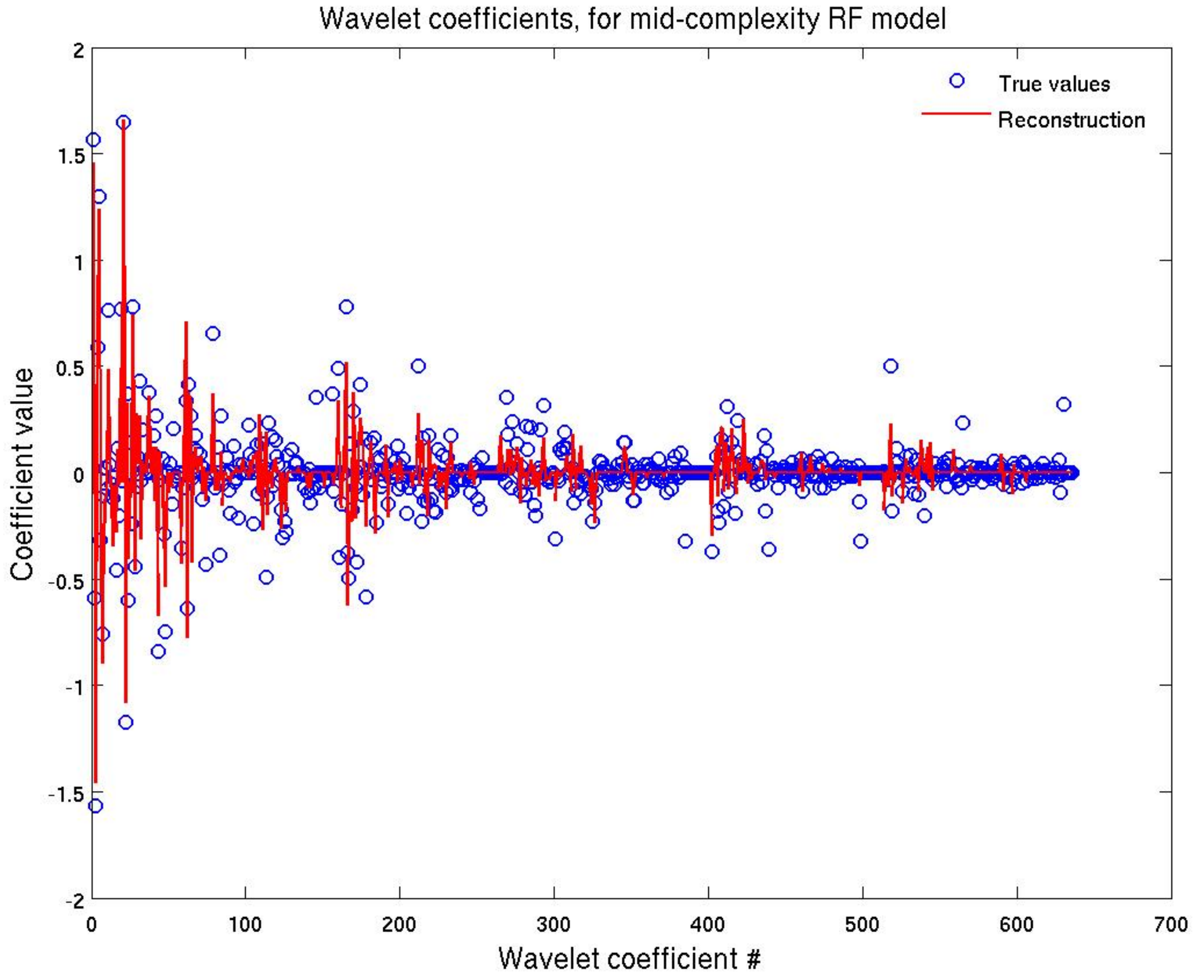
- A wavelet-based random field model can be used to represent $ffCO_2$ emissions
- Sparsifying using nightlights does not reduce dimensionality sufficiently
  - need sparsity-enforced estimation in light of sparse measurements
- Not discussed here – non-negativity enforcement
  - The emissions estimated by sparsity enforcement can sometimes be negative
  - A post-processing step (non-sparsity enforcing) corrects it
  - Simple and works only because we start with a very good guess
- Under the simplifying assumption of being able to measure $ffCO_2$, high accuracies can be obtained

Questions?



Reconstructed emissions in 8-day period 3 [microMoles/m2/sec]
5
10
15
20
25
10
20
30
40
50
60
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05

Did sparsification work?



- Only about half the wavelets could be estimated
- We are probably not over-fitting the problem
  - Data-driven sparsification works

Sandia National Laboratories