Spatial Statistical Data Fusion for Remote Sensing Applications

Amy Braverman[1] Hai Nguyen[1] Noel Cressie[2] Matthias Katzfuss[2]
Edward Olsen[1] Rui Wang[2] Anna Michalak[3] Charles Miller[1]

[1]Jet Propulsion Laboratory,
California Institute of Technology

[2]Department of Statistics,
The Ohio State University

[3]Department of Civil and Environmental Engineering,
University of Michigan

June 22, 2010

- The work presented here is based on three related bodies of research:
  1. a theoretical foundation developed by Noel Cressie,
  2. a method for spatial interpolation of very large data sets called Fixed-Rank Kriging, also developed by Noel Cressie and students,
  3. Spatial Statistical Data Fusion, the Ph.D. dissertation of Hai Nguyen, which extends Fixed-Rank Kriging to address the data fusion problem.

- This research is supported by the NASA Earth Science Technology Office through its Advanced Information Systems Technology Program.

- Some small notational errors in the published version of the paper. If you'd like a corrected copy, please email me at Amy.Braverman@jpl.nasa.gov.

Outline

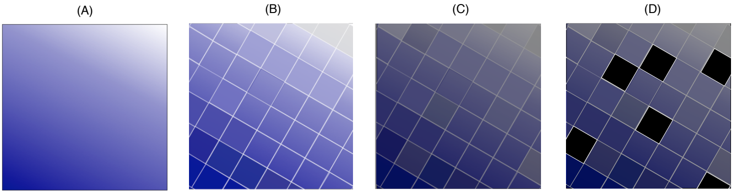
Introduction

- Data fusion means many things to many different people.
- This is true even within the remote sensing community (e.g. "image fusion").
- Our definition focusses on Earth science: *infer the true value of some quantity of interest from multiple data sources with different statistical characteristics (e.g. resolutions, systematic and random errors, etc.).*
- The fused data are these estimates (also called the "predictions"). Uncertainties of the estimates (mean squared prediction errors, *MSPE*'s) must accompany the predictions.
- Inference from spatial data relies on the theory of spatial statistics, so that is the formalism we use.
- First, consider the problem of inferring the true value from a single data source.

Introduction

Remote sensing data:

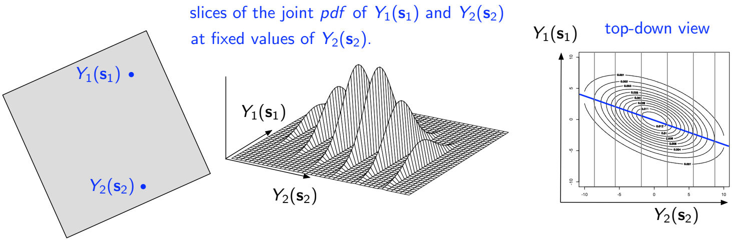
(A) is the true field. (B) is discretized into pixels. (C) is noisy (measurement bias and variance added). (D) has missing data.



Given only (D), can we infer (A)?

Let $\mathbf{s}_1$ and $\mathbf{s}_2$ be the (lat,lon) pairs of two point locations.

Let $Y_1(\cdot)$ and $Y_2(\cdot)$ be random variables representing the values of two quantities (e.g. air temperature and humidity) at the locations of their arguments.

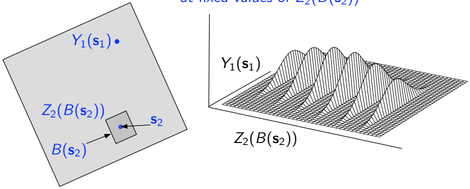


$E[Y_1(\mathbf{s}_1)|Y_2(\mathbf{s}_2)]$ = projection of the slice means onto the floor is a line (linear regression).

$E[\cdot]$ = expected value. $[Y_1(\mathbf{s}_1)|Y_2(\mathbf{s}_2)]$ = conditional distribution of $Y_1(\mathbf{s}_1)$ given $Y_2(\mathbf{s}_2)$.

- Remote sensing data typically have areal extent (support), not point support.

slices of the joint *pdf* of $Y_1(\mathbf{s}_1)$ and $Z_2B(\mathbf{s}_2))$
at fixed values of $Z_2(B(\mathbf{s}_2))$



Slices have larger variances, and linear regression may be biased relative to that of $Y_1(\mathbf{s}_1)$ on $Y_2(\mathbf{s}_2)$.

$$Z_2(B(\mathbf{s}_2)) = \left[\frac{1}{|B(\mathbf{s}_2)|} \int_{\mathbf{u} \in B(\mathbf{s}_2)} Y_2(\mathbf{u}) d\mathbf{u} \right] + \epsilon_2(B(\mathbf{s}_2)),$$

where $|\cdot|$ is size, and $\epsilon_2(\cdot)$ is measurement error.

- Remote sensing data have measurement error: $E(\epsilon_2(\cdot)) = b\mu$ (bias), $\mu = E(Y(\cdot))$; $Var(\epsilon_2(\cdot)) = \sigma^2$ (variance).

Kriging

- Optimal statistical spatial interpolation (for point support) using covariances to determine weights.
- Georges Matheron (1963).
- Let $Y(\cdot)$ be a statistical "process" (random variable) to be estimated at location $\mathbf{s}_0$ from observations at locations $\mathbf{s}_1, \ldots, \mathbf{s}_N \in D$.
- Assume $Z(\mathbf{s}_i) = Y(\mathbf{s}_i) + \epsilon(\mathbf{s}_i)$, where $\epsilon(\cdot)$ is zero-mean white noise with finite variance, $\sigma^2 v(\mathbf{s})$, $\sigma^2 > 0$, and $v(\cdot)$ is assumed known.
- Assume Y has linear mean structure:

$$Y(\mathbf{s}) = \mathbf{t}(\mathbf{s})'\boldsymbol{\alpha} + \nu(\mathbf{s}),$$

  where $\mathbf{t}(\cdot)$ is a vector of known covariates (e.g. latitude and longitude), $\boldsymbol{\alpha}$ is estimated from the data, and $\nu(\cdot)$ is small scale variation.
- $\nu(\mathbf{s})$ is assumed to be zero mean with finite, non-zero variance, and (spatial) covariance function,

$$Cov(\nu(\mathbf{u}), \nu(\mathbf{v})) = C(\mathbf{u}, \mathbf{v}).$$

Kriging

- Combine all this and get:

$$\mathbf{Z} = \mathbf{T}\boldsymbol{\alpha} + \boldsymbol{\delta}, \quad \boldsymbol{\delta} = \boldsymbol{\nu} + \boldsymbol{\epsilon},$$

  where $\boldsymbol{\delta}$, $\boldsymbol{\nu}$, and $\boldsymbol{\epsilon}$ are vectors of length N, $\mathbf{T}$ is an $N \times p$ matrix of covariates ($p = 2$ for lat/lon).

- Note: $\boldsymbol{\delta}$ is a combination of small scale variation and measurement error. Write:

$$Cov(\boldsymbol{\delta}) = \boldsymbol{\Sigma} = \mathbf{C} + \sigma^2\mathbf{V},$$

  where $\mathbf{C}$ is an $N \times N$ matrix of spatial covariances, $[\mathbf{C}]_{ij} = C(\mathbf{s}_i, \mathbf{s}_j)$, and $\mathbf{V} = diag(v(\mathbf{s}1), \ldots, v(\mathbf{s}_N))$. Note: $\mathbf{V}$ allows for non-constant measurement error variance.

- The kriging estimator of $Y(\mathbf{s}_0)$ is $\hat{Y}(\mathbf{s}_0) = \mathbf{a}'\mathbf{Z}$ where $\mathbf{a}$ is chosen to minimize $E\|Y(\mathbf{s}_0) - \mathbf{a}'\mathbf{Z}\|^2$ subject to the unbiasedness condition, $E(\mathbf{a}'\mathbf{Z}) = E(Y(\mathbf{s}_0)) = \mu$.

Answer:

$$\hat{Y}(\mathbf{s}_0) = \mathbf{t}(\mathbf{s}_0)\hat{\alpha} + \mathbf{a}'(\mathbf{Z} - \mathbf{T}\hat{\alpha}),$$

$$RMSPE(\hat{Y}(\mathbf{s}_0)) = \{C(\mathbf{s}_0, \mathbf{s}_0) - \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a} + (\mathbf{t}(\mathbf{s}_0) - \mathbf{T}'\mathbf{a})'(\mathbf{T}'\boldsymbol{\Sigma}^{-1}\mathbf{T})^{-1}(\mathbf{t}(\mathbf{s}_0) - \mathbf{T}'\mathbf{a})\}^{\frac{1}{2}},$$

where

$$\hat{\alpha} = (\mathbf{T}'\boldsymbol{\Sigma}^{-1}\mathbf{T})^{-1}\mathbf{T}'\boldsymbol{\Sigma}^{-1}\mathbf{Z}, \quad \mathbf{a} = \mathbf{c}(\mathbf{s}_0)\boldsymbol{\Sigma}^{-1}, \quad \mathbf{c}(\mathbf{s}_0) \equiv (C(\mathbf{s}_0, \mathbf{s}_1), \ldots, C(\mathbf{s}_0, \mathbf{s}_N))'.$$

If there is bias in the measurement, $E(\epsilon(\mathbf{s})) = b\mu$, then

$$\mathbf{a} = \left(\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Sigma}^{-1}\mathbf{1}(1+b)\left[-\mathbf{1}'(1+b)\boldsymbol{\Sigma}^{-1}\mathbf{1}(1+b)\right]^{-1}\mathbf{1}'(1+b)\boldsymbol{\Sigma}^{-1}\right)\mathbf{c}(\mathbf{s}_0) + \boldsymbol{\Sigma}^{-1}\mathbf{1}(1+b)\left[-\mathbf{1}'(1+b)\boldsymbol{\Sigma}^{-1}\mathbf{1}(1+b)\right]^{-1}\mathbf{1}'(1+b)\boldsymbol{\Sigma}^{-1}.$$

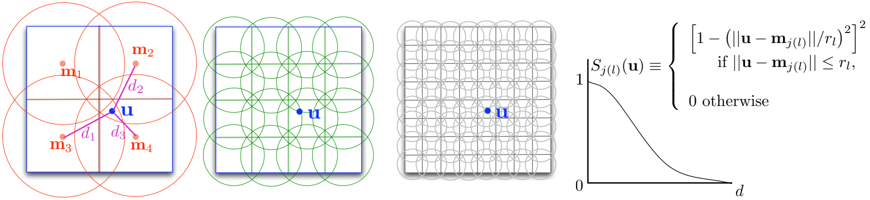
Fixed-Rank Kriging

- Unless C is isotropic and stationary, it is hard to invert the $N \times N$ covariance matrix $\mathbf{\Sigma}$ because N is very large. Isotropy and stationarity are unrealistic for most geophysical processes, particularly at large scales.
- We don't observe $Z(\mathbf{s})$ (point support), we observe $Z(B(\mathbf{s}))$ (footprint or "block" support).
- Cressie and Johannesson (2008) introduced Fixed-Rank Kriging (FRK) as a way to deal with these problems. Model the covariance function as

$$C(\mathbf{u}, \mathbf{v}) = \mathbf{S}(\mathbf{u})'\mathbf{K}\mathbf{S}(\mathbf{v}), \quad \mathbf{u}, \mathbf{v} \in D,$$

  for some $r \times r$ positive-definite matrix $\mathbf{K}$, $r << N$. $\mathbf{S}(\cdot)$ is the basis expansion of a point location into a fixed set of (not necessarily orthogonal) basis functions, $S_j(\cdot)$: $\mathbf{S}(\mathbf{u}) = (S_1(\mathbf{u}), \ldots, S_r(\mathbf{u}))'$.

Fixed-Rank Kriging



Resolution 1 | Resolution 2 | Resolution 3 | Local Bisquare Functions

Multi-resolution spatial basis functions. The spatial domain is subdivided into three levels of resolution, each a factor of two finer than its parent. At resolution l, each cell center ($m_1, \ldots, m_4$ in the left panel, but not shown in the others) is the center of a circle of diameter r_l . Point location $\mathbf{u}$ belongs to three circles, and is distances d_1, d_2, and d_3 from the cell centers, respectively. These distances determine the basis function values at resolution l, as shown in the right panel.

Fixed-Rank Kriging

Then $\mathbf{\Sigma} = \sigma^2\mathbf{V} + \mathbf{S}'\mathbf{K}\mathbf{S}$, and

$$\mathbf{\Sigma}^{-1} = (\sigma^2\mathbf{V})^{-1} - (\sigma^2\mathbf{V})^{-1}\mathbf{S}'\left(\mathbf{K}^{-1} + \mathbf{S}(\sigma^2\mathbf{V})^{-1}\mathbf{S}'\right)^{-1}\mathbf{S}'(\sigma^2\mathbf{V})^{-1},$$

by the Sherman-Morrison-Woodbury formula (Henderson and Searle, 1981). This only requires the inversion of $\mathbf{K}$ and $(\mathbf{K}^{-1} + \mathbf{S}'(\sigma^2\mathbf{V})^{-1}\mathbf{S})$, both of which are $r \times r$ matrices. Order of computation is $O(Nr^2)$, not $O(N^3)$.

The FRK kriging predictors and their uncertainties are

$$\hat{Y}(\mathbf{s}_0) = \mathbf{t}(\mathbf{s}_0)\hat{\alpha} + \mathbf{S}(\mathbf{s}_0)'\hat{\mathbf{K}}\mathbf{S}\check{\mathbf{\Sigma}}^{-1}(\mathbf{Z} - \mathbf{T}\hat{\alpha}),$$

$$\begin{aligned} RMSPE(\hat{Y}(\mathbf{s}_0)) = \{&\mathbf{S}(\mathbf{s}_0)'\hat{\mathbf{K}}\mathbf{S} - \mathbf{S}(\mathbf{s}_0)'\hat{\mathbf{K}}\mathbf{S}\check{\mathbf{\Sigma}}^{-1}\mathbf{S}'\hat{\mathbf{K}}\mathbf{S}(\mathbf{s}_0) + \\ &(\mathbf{t}(\mathbf{s}_0) - \mathbf{T}'\check{\mathbf{\Sigma}}^{-1}\mathbf{S}'\hat{\mathbf{K}}\mathbf{S}(\mathbf{s}_0))'(\mathbf{T}'\check{\mathbf{\Sigma}}^{-1}\mathbf{T})^{-1}(\mathbf{t}(\mathbf{s}_0) - \mathbf{T}'\check{\mathbf{\Sigma}}^{-1}\mathbf{S}'\hat{\mathbf{K}}\mathbf{S}(\mathbf{s}_0))\}^{\frac{1}{2}}. \end{aligned}$$

$\check{\mathbf{\Sigma}} = \sigma^2\mathbf{V} + \mathbf{S}'\hat{\mathbf{K}}\mathbf{S}$ and $\hat{\mathbf{K}}$ is estimated from the data. (Assume σ^2 is given and no measurement bias.)

Fixed-Rank Kriging

What about estimating $\mathbf{K}$? We can use the footprint-level data:

$$Cov(Z(B(\mathbf{s}_k)), Z(B(\mathbf{s}_l))) =$$

$$Cov\left[\frac{1}{|B(\mathbf{s}_k)|}\int_{\mathbf{u}\in B(\mathbf{s}_k)} Y(\mathbf{u})d\mathbf{u} + \epsilon(B(\mathbf{s}_k)), \frac{1}{|B(\mathbf{s}_l)|}\int_{\mathbf{v}\in B(\mathbf{s}_l)} Y(\mathbf{v})d\mathbf{v} + \epsilon(B(\mathbf{s}_l)\right],$$

$$= \frac{1}{|B(\mathbf{s}_k)|}\frac{1}{|B(\mathbf{s}_l)|}\int_{\mathbf{u}\in B(\mathbf{s}_k)}\int_{\mathbf{v}\in B(\mathbf{s}_l)} Cov(Y(\mathbf{u}), Y(\mathbf{v}))d\mathbf{u}d\mathbf{v}$$

$$= \frac{1}{|B(\mathbf{s}_k)|}\int_{\mathbf{u}\in B(\mathbf{s}_k)} \mathbf{S}(\mathbf{u})'d\mathbf{u}\ \mathbf{K}\ \frac{1}{|B(\mathbf{s}_l)|}\int_{\mathbf{v}\in B(\mathbf{s}_l)} \mathbf{S}(\mathbf{v})'d\mathbf{v},$$

$$= \tilde{\mathbf{S}}(B(\mathbf{s}_k))'\,\mathbf{K}\,\tilde{\mathbf{S}}(B(\mathbf{s}_l)),$$

where $\tilde{\mathbf{S}}(B(\mathbf{s})) = \left(\tilde{S}_1(B(\mathbf{s})), \ldots, \tilde{S}_r(B(\mathbf{s}))\right)$, $\tilde{S}_j(B(\mathbf{s})) = \frac{1}{|B(\mathbf{s})|}\int_{\mathbf{u}\in B(\mathbf{s})} S_j(\mathbf{u})d\mathbf{u}$.

Fixed-Rank Kriging

Estimate $\mathbf{K}$ by:

1. Subdividing the domain into coarse bins (e.g. resolution of a coarse level of $\mathbf{S}$) and calculating $\hat{\mathbf{\Sigma}}$, an empirical estimate of the spatial covariance matrix (details omitted in the interest of time).
2. Find $\mathbf{K}$ that minimize the distance between $\check{\mathbf{\Sigma}}$ and $\hat{\mathbf{\Sigma}}$:

$$\|\hat{\mathbf{\Sigma}} - \check{\mathbf{\Sigma}}(\mathbf{K})\|_F = tr\left((\hat{\mathbf{\Sigma}} - \check{\mathbf{\Sigma}}(\mathbf{K}))'(\hat{\mathbf{\Sigma}} - \check{\mathbf{\Sigma}}(\mathbf{K}))\right).$$

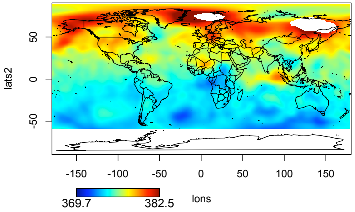
This yields a method-of-moments estimate, $\hat{\mathbf{K}}$.

National Aeronautics and Space Administration

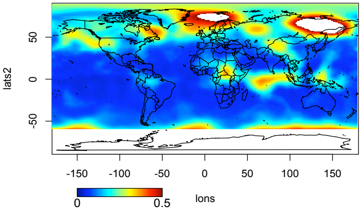
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Example: FRK AIRS CO2, May 1-3, 2003

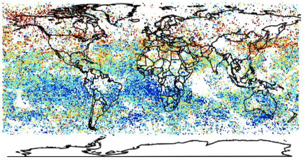
Prediction



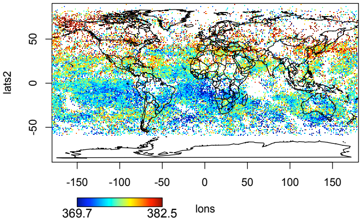
RMSPE



Raw Data



Simple average, one degree grid

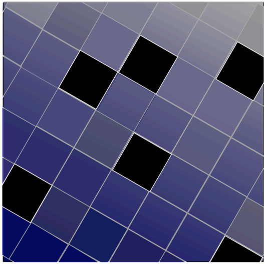


Example: FRK AIRS CO2, May 1-3, 2003

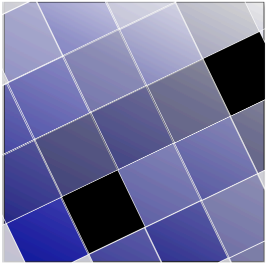
- Locations south of 60°S screened out.
- Results with $RMSPE > .5$ screened out.
- Prediction grid is rectangular, $1^\circ \times 1^\circ$.
- 396 basis functions at three coarsest levels of resolution on hexagonal discrete global grid (DGG; Sahr and White, 1998). Intercell distances at level 1 $\approx$ 4,400 km; at level 2 $\approx$ 2,500 km; at level 3 $\approx$ 1,400 km.
- Binning for method-of-moments estimation of $\mathbf{K}$ uses level 5 DGG hexagons (intercell distance $\approx$ 400 km).
- Computation time: 15 seconds of a 3 GHz MacBook Pro.

Data Fusion

Surely having a second data set must help, especially if it's measuring the same thing. If not, it will still help if the second measurement is correlated with the first.



Instrument 1



Instrument 2

Single process multiple source (SPMS), $Y_1(\cdot) = Y_2(\cdot) = Y(\cdot)$ or multiple process multiple source (MPMS), $Y_1(\cdot) \neq Y_2(\cdot)$.

Data Fusion

$$E(Y_i(\mathbf{s})) = \mu_i, \quad E(\epsilon_j(B_{jk})) = b_j\mu_j, \quad Var(\epsilon_j(B_{jk})) = \sigma_j^2, \quad B_{jk} = B_j(\mathbf{s}_k),$$

where i indexes process (variable), j indexes instrument, and k indexes footprint.

$$\mathbf{Z}_j = (Z_j(B_{j1}), \ldots, Z_j(B_{jN_j}))',$$
$$Z_j(B_{jk}) = \frac{1}{|B_{jk}|}\int_{\mathbf{u}\in B_{jk}} Y_j(\mathbf{u})\, d\mathbf{u} + \epsilon_j(B_{jk}),$$

$$\hat{\mathbf{Y}}(\mathbf{s}) = \begin{bmatrix} \hat{Y}_1(\mathbf{s}) \\ \hat{Y}_2(\mathbf{s}) \end{bmatrix} = \begin{bmatrix} \mathbf{a}'_{11\mathbf{s}}\mathbf{Z}_1 + \mathbf{a}'_{12\mathbf{s}}\mathbf{Z}_2 \\ \mathbf{a}'_{21\mathbf{s}}\mathbf{Z}_1 + \mathbf{a}'_{22\mathbf{s}}\mathbf{Z}_2 \end{bmatrix}.$$

Minimize $E\left\{\left(\hat{Y}_1(\mathbf{s}) - Y_1(\mathbf{s})\right)^2 + \left(\hat{Y}_2(\mathbf{s}) - Y_2(\mathbf{s})\right)^2\right\}$ subject to

$$E(\hat{Y}_1(\mathbf{s})) = b_1\mathbf{a}'_{11\mathbf{s}}\mathbf{1}_{N_1}\mu_1 + b_2\mathbf{a}'_{12\mathbf{s}}\mathbf{1}_{N_2}\mu_2 = \mu_1 \text{ and}$$
$$E(\hat{Y}_2(\mathbf{s})) = b_1\mathbf{a}'_{21\mathbf{s}}\mathbf{1}_{N_1}\mu_1 + b_2\mathbf{a}'_{22\mathbf{s}}\mathbf{1}_{N_2}\mu_2 = \mu_2.$$

Data Fusion

$$\mathbf{a}_{i1\mathbf{s}} = \mathbf{A}_1^{-1}(\mathbf{B}_{i1} + \mathbf{C}_1 m_i), \quad \text{and} \quad \mathbf{a}_{i2\mathbf{s}} = \mathbf{A}_2^{-1}(\mathbf{B}_{i2} + \mathbf{C}_2 m_i).$$

where

$$m_i = \frac{(1 - \mathbf{1}'_{N_1}\mathbf{A}_1^{-1}\mathbf{B}_{i1}(1+b_1) + \mathbf{1}'_{N_2}\mathbf{A}_2^{-1}\mathbf{B}_{i2}(1+b_2))}{(\mathbf{1}'_{N_1}\mathbf{A}_1^{-1}\mathbf{C}_1(1+b_1) + \mathbf{1}'_{N_2}\mathbf{A}_2^{-1}\mathbf{C}_2(1+b_2))}, \quad \text{and}$$

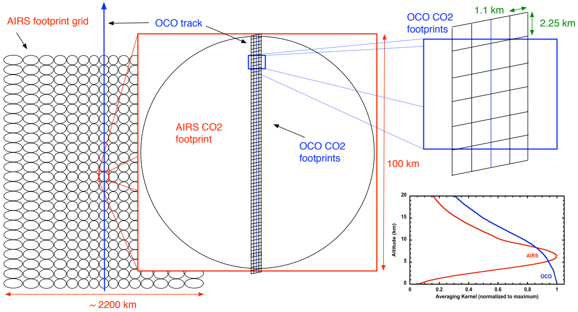
$$\begin{aligned}
\mathbf{A}_1 &\equiv (\mathbf{I}_{N_1} - \check{\boldsymbol{\Sigma}}_{11}^{-1}\check{\boldsymbol{\Sigma}}_{12}\check{\boldsymbol{\Sigma}}_{22}^{-1}\check{\boldsymbol{\Sigma}}_{21}),\\
\mathbf{A}_2 &\equiv (\mathbf{I}_{N_2} - \check{\boldsymbol{\Sigma}}_{22}^{-1}\check{\boldsymbol{\Sigma}}_{21}\check{\boldsymbol{\Sigma}}_{11}^{-1}\check{\boldsymbol{\Sigma}}_{12}),\\
\mathbf{B}_{i1} &\equiv \check{\boldsymbol{\Sigma}}_{11}^{-1}(\mathbf{c}_{i1s} - \check{\boldsymbol{\Sigma}}_{12}\check{\boldsymbol{\Sigma}}_{22}^{-1}\mathbf{c}_{i2s}),\\
\mathbf{B}_{i2} &\equiv \check{\boldsymbol{\Sigma}}_{22}^{-1}(\mathbf{c}_{i2s} - \check{\boldsymbol{\Sigma}}_{21}\check{\boldsymbol{\Sigma}}_{11}^{-1}\mathbf{c}_{i1s}),\\
\mathbf{C}_1 &= -\check{\boldsymbol{\Sigma}}_{11}^{-1}\left[\mathbf{1}_{N_1}(1+b_1) - \check{\boldsymbol{\Sigma}}_{12}\check{\boldsymbol{\Sigma}}_{22}^{-1}\mathbf{1}_{N_2}(1+b_2)\right],\\
\mathbf{C}_2 &= -\check{\boldsymbol{\Sigma}}_{22}^{-1}\left[\mathbf{1}_{N_2}(1+b_2) - \check{\boldsymbol{\Sigma}}_{21}\check{\boldsymbol{\Sigma}}_{11}^{-1}\mathbf{1}_{N_1}(1+b_1)\right].
\end{aligned}$$

$$\begin{aligned}
\check{\boldsymbol{\Sigma}}_{11} &= \tilde{\mathbf{S}}_1'\hat{\mathbf{K}}_{11}\tilde{\mathbf{S}}_1 + \sigma_1^2\mathbf{V}_1,\\
\check{\boldsymbol{\Sigma}}_{12} &= \tilde{\mathbf{S}}_1'\hat{\mathbf{K}}_{12}\tilde{\mathbf{S}}_2,\\
\check{\boldsymbol{\Sigma}}_{21} &= \tilde{\mathbf{S}}_2'\hat{\mathbf{K}}_{21}\tilde{\mathbf{S}}_1\\
\check{\boldsymbol{\Sigma}}_{22} &= \tilde{\mathbf{S}}_2'\hat{\mathbf{K}}_{22}\tilde{\mathbf{S}}_2 + \sigma_2^2\mathbf{V}_2,
\end{aligned}$$

($\mathbf{K}_{ij}$ estimated using method of moments as described earlier.)

Fusing AIRS and OCO-2 CO2

- OCO-2 will measure total column CO2 on 1.1×2.25 km footprints.
- AIRS measures mid-tropospheric (and above) CO2 at 90 km resolution.
- $Y_1(\cdot)$ = total column CO2, $Y_2(\cdot)$ = CO2 in the mid-troposphere and above.



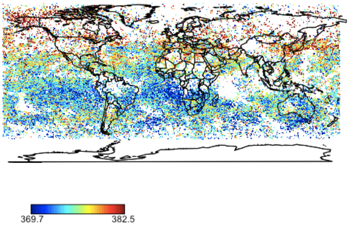
- $Y_1(\mathbf{s}) - Y_2(\mathbf{s})$ = planetary boundary layer CO2 (PBL CO2).

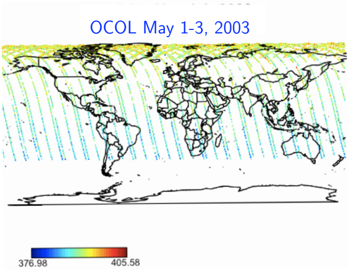
Fusing AIRS and OCO-2 CO_2

- PBL is that portion of the lower atmosphere that is dragged along by the rotation of the Earth.
- CO_2 fluxes from the surface enter the atmosphere at the PBL.
- Changes in PBL CO_2 at any time should generally be correlated with flux.
- Monitoring PBL CO_2 may allow monitoring of sources and sinks.

Fusing AIRS and OCO-2 CO2

- OCO-like (OCOL) synthetic data downscaled from Parallel Climate Transport Model (PCTM) at $1^\circ \times 1.25^\circ$ resolution to 1×2 km resolution.

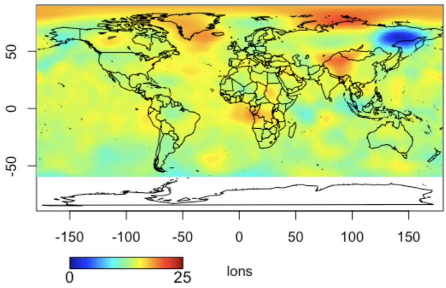




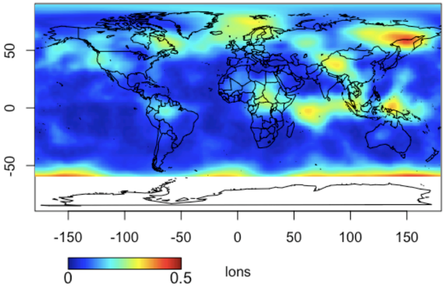
- Estimate $Y_{PBL}(\mathbf{s}) = Y_1(\mathbf{s}) - Y_2(\mathbf{s}) = (w_1, w_2) \cdot (\hat{Y}_1(\mathbf{s}), (\hat{Y}_2(\mathbf{s}))'$, with $(w_1, w_2) = (1, -1)$.
- $MSPE = (w_1, w_2) \cdot Cov \left[\begin{pmatrix} \hat{Y}_1(\mathbf{s}) \\ \hat{Y}_2(\mathbf{s}) \end{pmatrix} - \begin{pmatrix} Y_1(\mathbf{s}) \\ Y_2(\mathbf{s}) \end{pmatrix} \right] \cdot (w_1, w_2)'$.

Fusing AIRS and OCO-2 CO2

SSDF Predictions, May 1-3, 2003



SSDF Standard Errors, May 1-3, 2003



- Prediction grid: $1^\circ \times 1^\circ$ rectangular; 396 basis functions at three levels of DGG etc.. SE's shown are truncated at .5 ppm, some higher.
- Computation time on 3 GHz MacBook Pro $\approx$ 5 minutes, half for computation of *MSPE*'s.
- Performed for 89 overlapping three day blocks May 1 through July 31, 2003.

Conclusions

- Unlike more ad-hoc approaches, this methodology is *inferential*. It yields formal statistical estimates and their uncertainties relative to truth.
- We've demonstrated computational feasibility, but need to study trade-off's related to number and kind of basis functions, binning and estimation of **K**'s etc.
- Difficult to judge results because we have combined AIRS observations with synthetic OCO-2. Need to create synthetic "truth", derive synthetic AIRS and OCO-2, fuse and judge results against "truth".
- Results depend crucially on biases and variances of measurement error terms. Instrument team validation results are the only sources of this information.
- Methodology is potentially applicable to many other situations.
- Next: Spatio-temporal Data Fusion (STDF) based on Fixed-Rank Filtering (Kang, Cressie, and Shi, 2009).