Tomoko Matsuo's collaborators

Thermosphere-Ionosphere GCM Modeling

A. Richmond (NCAR-HAO)
T. Fuller-Rowell and M. Codrescu (NOAA)

Polar Ionosphere Data Assimilation

A. Richmond, G. Lu, and B. Emery (NCAR-HAO)
D. Lummerzheim (U of Alaska)
M. Hairston (U of Texas, Dallas)

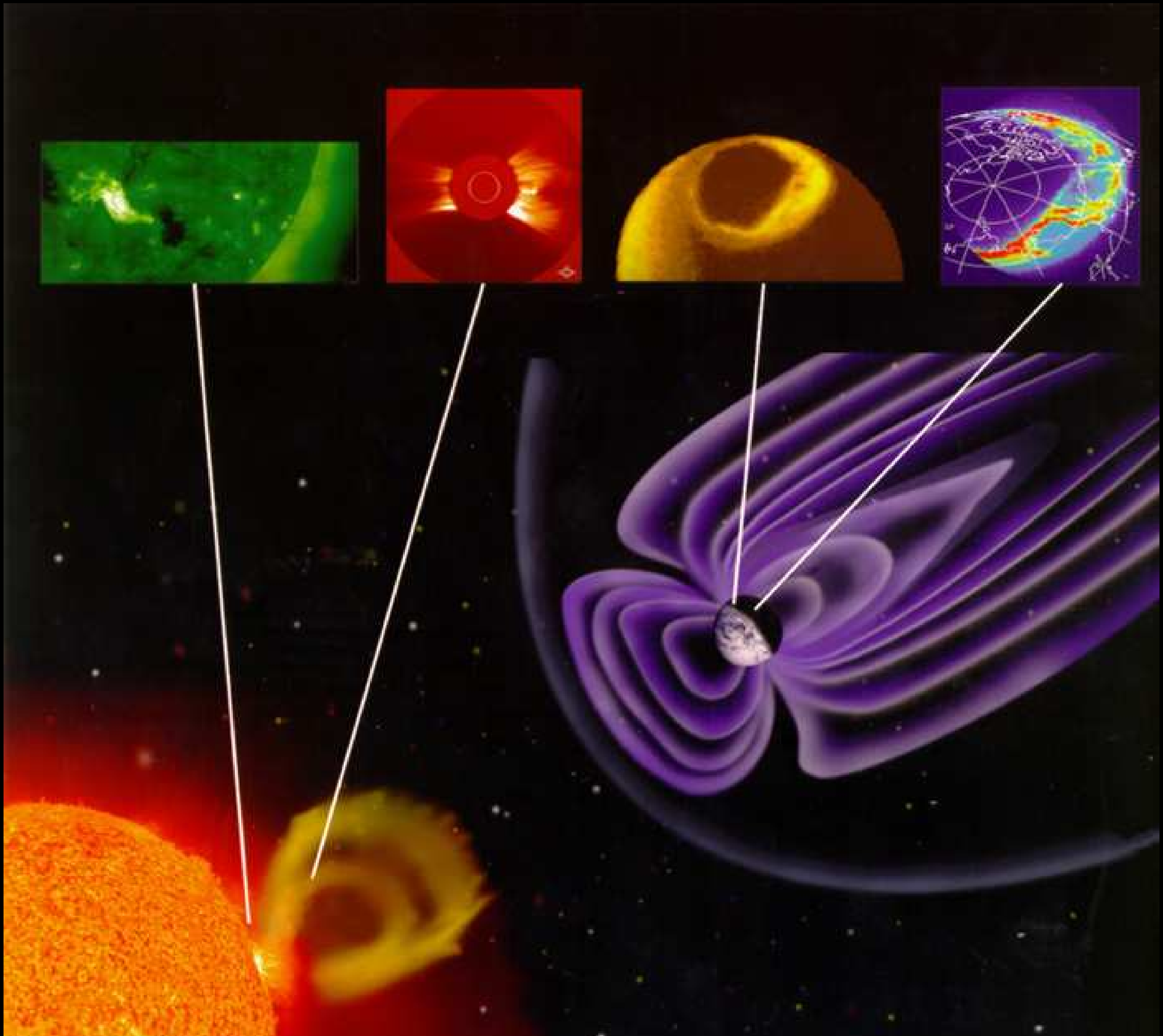
Spatial Statistics

D. Nychka (NCAR-IMAGe)
D. Paul (U of California, Davis)

Middle Atmosphere Data Assimilation

J. Anderson (NCAR-IMAGe)
D. Marsh and A. Smith (NCAR-ACD)

Sun-Earth Connection



GRL'98 Cover Page

Multi-resolution Based Nonstationary Covariance Modeling: Monte-Carlo EM approach

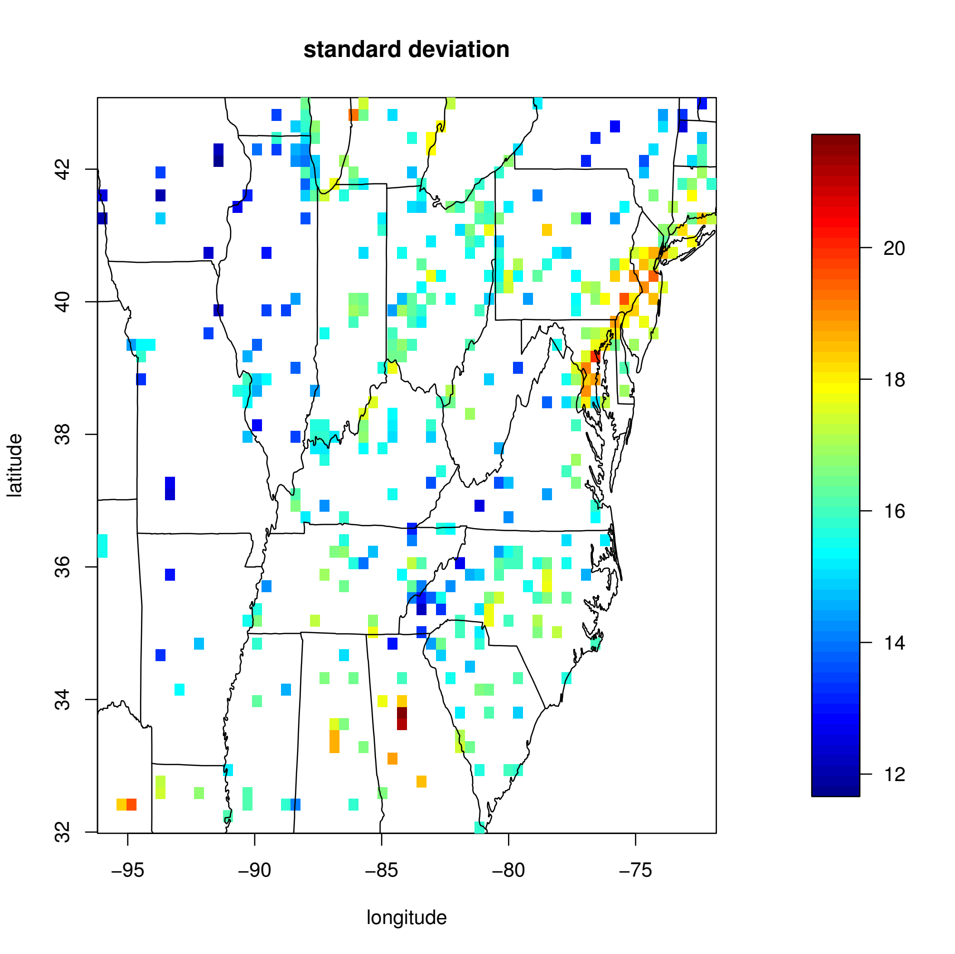
Tomoko Matsuo

in collaboration with

Doug Nychka & **Debashis Paul** (UC, Davis)

Surface Ozone

standard deviation



- O_3 is one of six common pollutants
- EPA's national air quality standards (80 ppb)
- 1997 Data Set
  - 364 locations on 48×48 grid
  - 184 days from May to Oct

Motivation and Goal

- Motivation:

  Flexible Nonstationary Covariance Model

  Gaussian Model in Spatial Statistics
  - Kriging (geostatistics)
    [e.g., Higdon et al., 1999; Fuentes, 2001; Fuentes and Smith, 2001; Nychka et al., 2003; Sampson and Guttorp, 1992; Anderes and Stein, 2005]
  - Variational and OI methods (data assimilation)
    [e.g., Purser et al., 2005; Gaspari et al., 2006]

- Goal and Challenges: Computational efficiency
  - Irregularly distributed observational data
  - Large data set

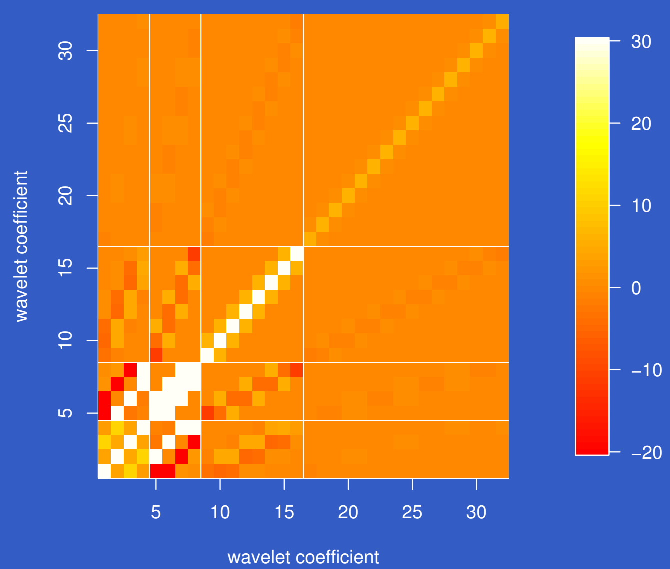
Nonparametric Model

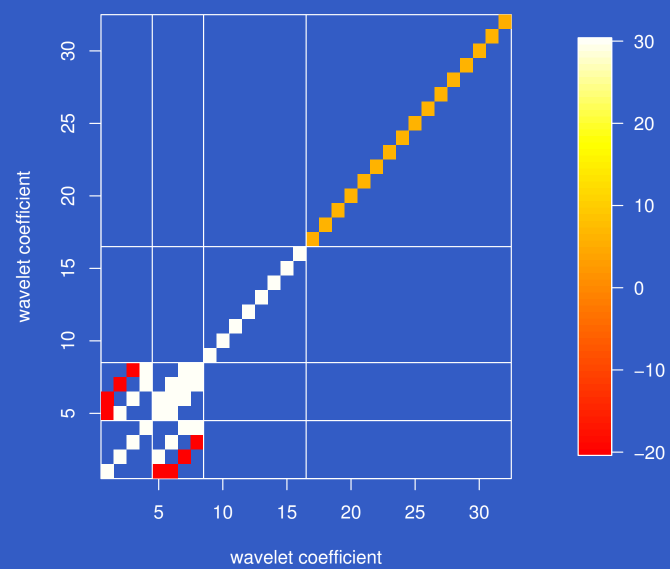
- $\boxed{\Sigma = \mathcal{W} H^2 \mathcal{W}^T}$ [Nychka, Wikle, and Royle, 2003]

$$H = \begin{bmatrix} H_{00} & H_{01} \\ H_{10} & H_{11} \end{bmatrix} \approx \begin{bmatrix} \tilde{H}_{00} & 0 \\ 0 & \tilde{H}_1 \end{bmatrix}$$

where $\tilde{H}_{00}$ is thresholded and $\tilde{H}_1 = \text{diag}(H_{11})$

- Enforced sparsity in H





Covariance Estimator

- Gaussian Model:

$$\mathbf{f} = \begin{pmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{pmatrix} \sim \mathcal{N}(0, \Sigma_\theta) \qquad \text{where} \;\; \Sigma_\theta = \mathcal{W}\tilde{H}^2(\theta)\mathcal{W}^T.$$

Covariance Estimator

- Gaussian Model:

$$\mathbf{f} = \begin{pmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{pmatrix} \sim \mathcal{N}(0, \Sigma_\theta) \qquad \text{where} \;\; \Sigma_\theta = \mathcal{W}\tilde{H}^2(\theta)\mathcal{W}^T.$$

- Monte-Carlo EM [e.g., Wei and Tanner, 1990]:

$$\begin{aligned} Q(\theta, \theta^*) &= E[\mathcal{L}(\mathbf{f}, \theta) | \mathbf{f}_1, \theta^*] \\ &\approx \frac{1}{N} \sum_{n=1}^{N} \mathcal{L}\left(\begin{pmatrix} \mathbf{f}_1 \\ \mathbf{f}_2^{(n)} \end{pmatrix}, \theta \right) \end{aligned}$$

  - MC sampling: $\mathbf{f}_2^{(n)} \sim [\; \mathbf{f}_2 | \mathbf{f}_1, \mathcal{W}\tilde{H}^2(\theta^*)\mathcal{W}^T \;]$

Covariance Estimator

- Gaussian Model:

$$\mathbf{f} = \begin{pmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{pmatrix} \sim \mathcal{N}(0, \Sigma_\theta) \qquad \text{where} \;\; \Sigma_\theta = \mathcal{W}\tilde{H}^2(\theta)\mathcal{W}^T.$$

- Monte-Carlo EM [e.g., Wei and Tanner, 1990]:

$$\begin{aligned} Q(\theta, \theta^*) &= E[\mathcal{L}(\mathbf{f}, \theta) | \mathbf{f}_1, \theta^*] \\ &\approx \frac{1}{N} \sum_{n=1}^{N} \mathcal{L}\left(\begin{pmatrix} \mathbf{f}_1 \\ \mathbf{f}_2^{(n)} \end{pmatrix}, \theta \right) \end{aligned}$$

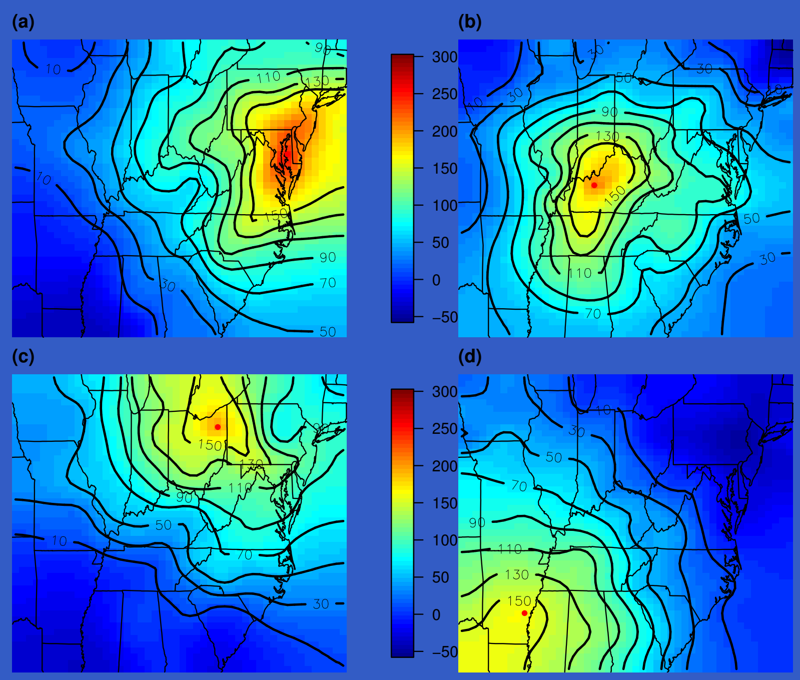
  - MC sampling: $\mathbf{f}_2^{(n)} \sim [\; \mathbf{f}_2 | \mathbf{f}_1, \mathcal{W}\tilde{H}^2(\theta^*)\mathcal{W}^T \;]$

- Smoothed MC EM [e.g., Silverman et al., 1990]:

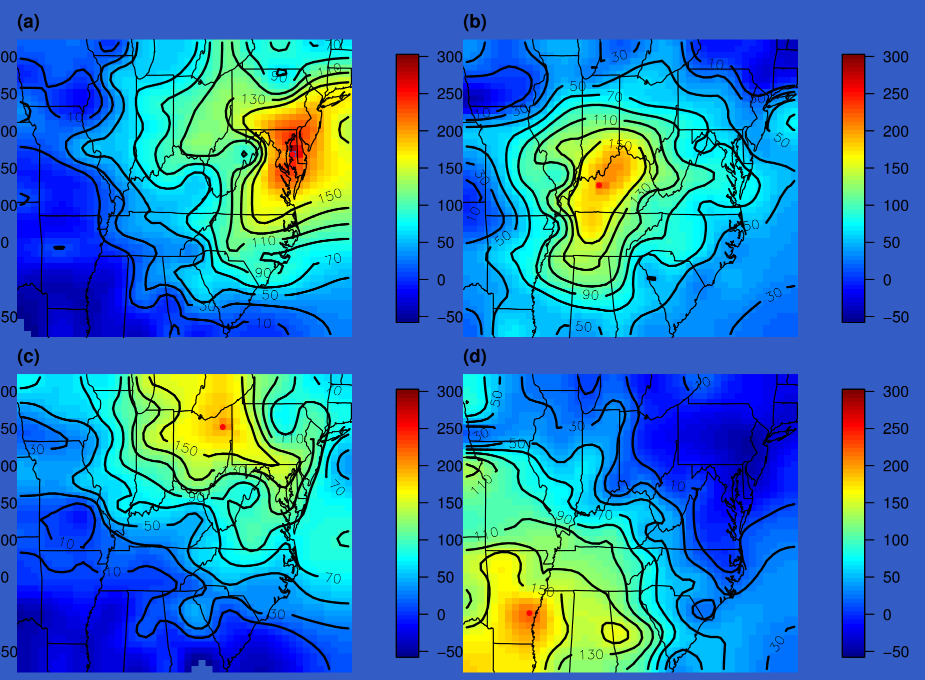
$$Q(\theta, \theta^*) + p(\theta)$$

Different thresholding in $\tilde{H}$

Lev 2, 86.08% (5045)

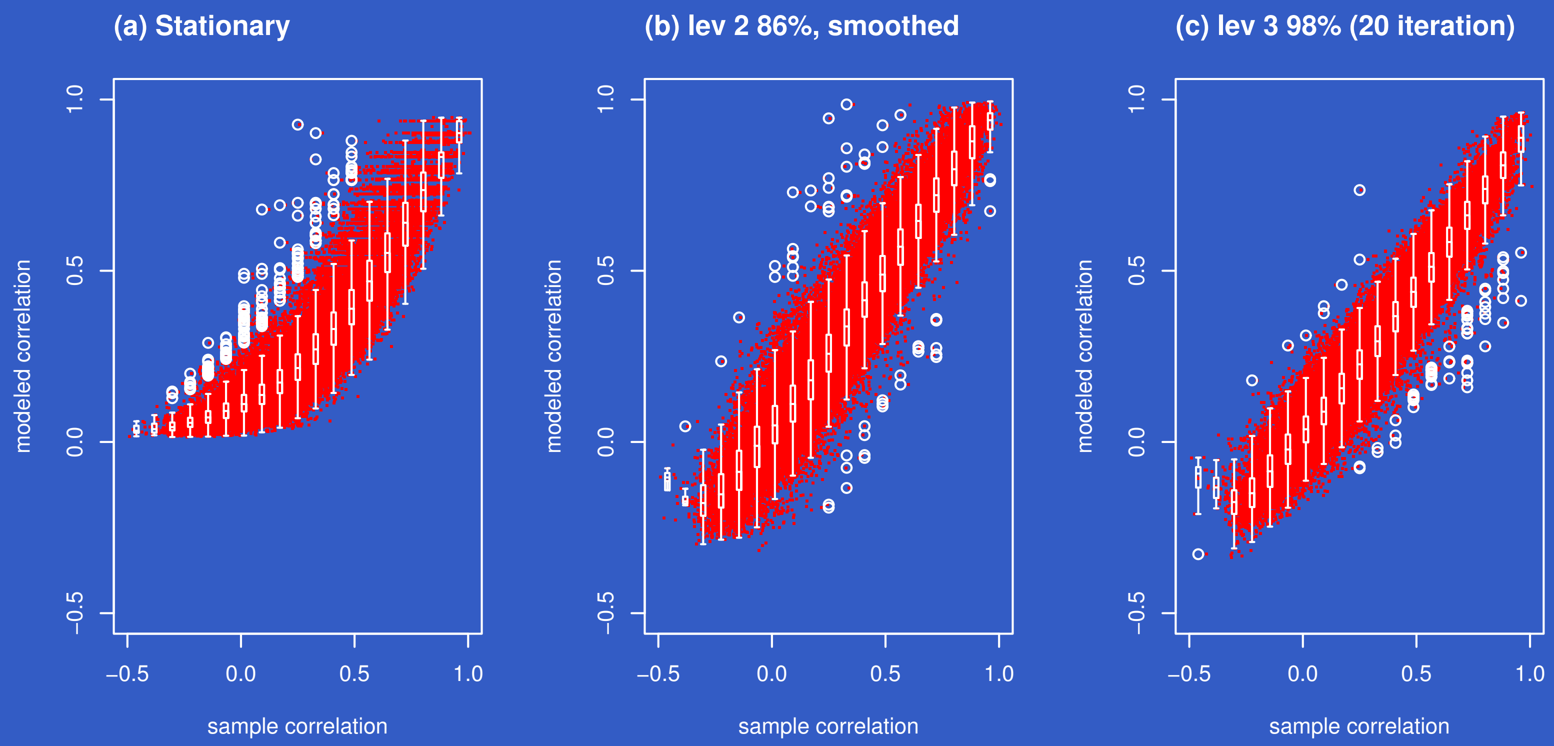


Lev 3, 98% (8363)



Validation

- Sample v.s. Modeled Correlation



Summary and Future Work

- To be submitted to JASA or JRSS-B
  - Flexible nonstationary covariance model $\mathcal{W}\tilde{H}^2\mathcal{W}^T$
  - Theory to support sparsity in $\tilde{H}$
  - Practical estimator (Monte-Carlo EM) to handle the incomplete data
  - Examples using surface ozone data

Summary and Future Work

- To be submitted to JASA or JRSS-B
  - Flexible nonstationary covariance model $\mathcal{W}\tilde{H}^2\mathcal{W}^T$
  - Theory to support sparsity in $\tilde{H}$
  - Practical estimator (Monte-Carlo EM) to handle the incomplete data
  - Examples using surface ozone data
- Another ozone application
  in collaboration with Eric Gilleland (NCAR-RAL)

Summary and Future Work

- To be submitted to JASA or JRSS-B
  - Flexible nonstationary covariance model $\mathcal{W}\tilde{H}^2\mathcal{W}^T$
  - Theory to support sparsity in $\tilde{H}$
  - Practical estimator (Monte-Carlo EM) to handle the incomplete data
  - Examples using surface ozone data
- Another ozone application
  in collaboration with Eric Gilleland (NCAR-RAL)
- Application to the Polar Ionosphere
  - Aurora image data ($\sim$ 100K)
  - Prior covariance for ionospheric data assimilation