A Bayesian Multidimensional Trend Filter

Stella Self, Christopher McMahan, and Andrew Brown

Trend Filtering

- Trends filters are time series analysis tools that remove noise or seasonal cycles.
- Trend filters penalize the 2nd (or other) derivative of the estimated trend function
- Hodrick-Prescott Trend Filtering (1997):

$$\underset{g_1,\ldots,g_T}{\text{argmin}}\{\sum_{t=1}^{T}(y_t-g_t)^2 + \lambda\sum_{t=3}^{T}((g_t-g_{t-1}-(g_{t-1}-g_{t-2}))^2\}$$

where y_t is the observation from time t, g_t is the trend estimate, and λ is a smoothness parameter.

- l_1 Trend Filtering (Seung-Jean et al 2009)

$$\underset{g_1,\ldots,g_T}{\text{argmin}}\{\sum_{t=1}^{T}(y_t-g_t)^2/2 + \lambda\sum_{t=2}^{T-1}|g_{t-1}-2g_t+g_{t+1}|\}$$

The Model

- Impose a regular grid over the support.
- Let $y_1,\ldots,y_n$ be observed data, and $\boldsymbol{\phi} = (\phi_1,\ldots,\phi_Q)'$ be the trend function at each grid cell. Assume the following model:

$$\begin{aligned} y_i|\phi_i,\boldsymbol{\theta} &\overset{iid}{\sim} \mathrm{D}(g(\phi_i,\boldsymbol{\theta})) \\ \pi(\boldsymbol{\phi}|\tau^2) &\sim \text{Multivariate Normal}(\mathbf{0},\tau^2\boldsymbol{M}^{-1}) \\ \pi(\tau^2) &\sim \text{Inverse Gamma}(\alpha_{\tau^2},\beta_{\tau^2}) \end{aligned}$$

- $\boldsymbol{M} = (\boldsymbol{D}-\boldsymbol{W})(\boldsymbol{D}-\boldsymbol{W})$, where $\boldsymbol{W}$ is the adjacency matrix of the grid (i.e. $\boldsymbol{W}(i,j)=1$ if cells i and j are adjacent and 0 otherwise), and $\boldsymbol{D}$ is diagonal with $\boldsymbol{D}(i,i)=\sum_{j=1}^{Q}\boldsymbol{W}(i,j) := d_i$.
- $\boldsymbol{\theta}$ includes other components of the mean structure (covariate effects, random effects, etc).
- The prior imposes a penalty in abrupt changes in $\boldsymbol{\phi}$ from cell to cell

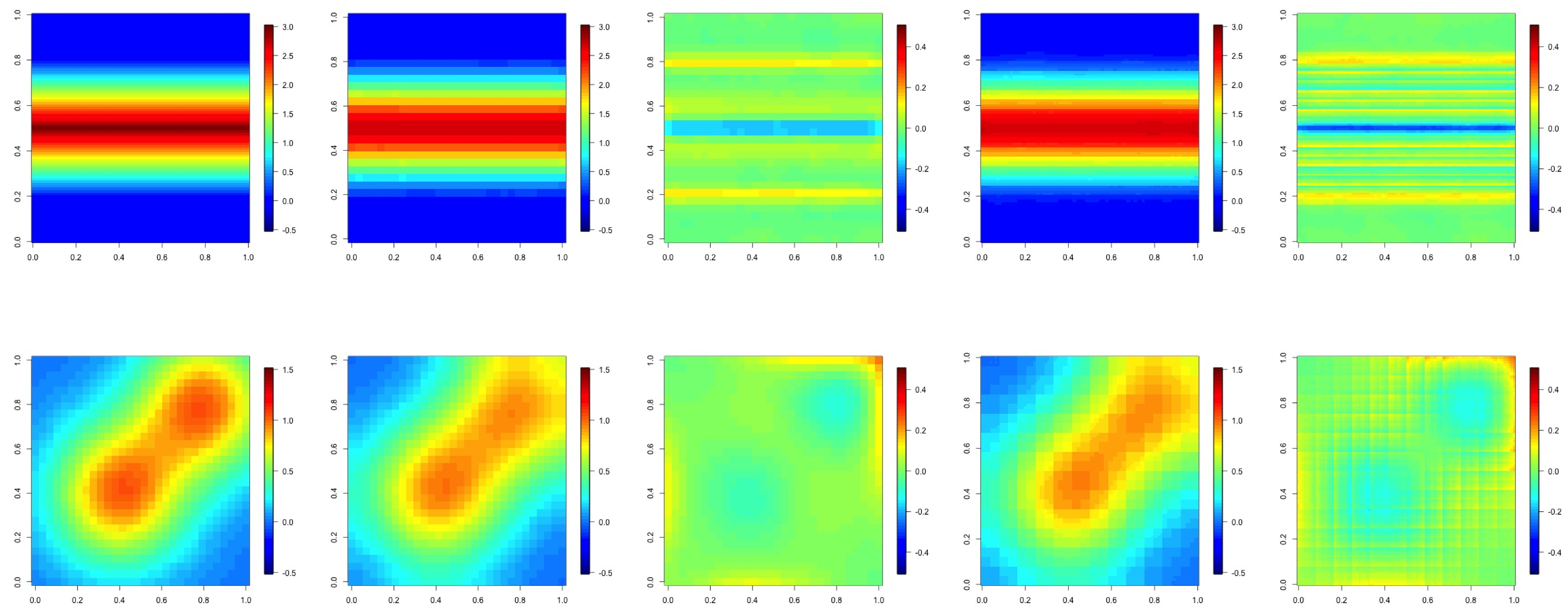
$$\pi(\boldsymbol{\phi}|\tau^2) \propto \exp\{-\frac{1}{2\tau^2}\sum_{s=1}^{Q}d_s^2\left(\phi_s-\sum_{j\in N(s)}\frac{\phi_j}{d_s}\right)^2\}$$

Adaptive Descretization

- Goal: Use the adaptive mesh refinement idea from PDEs to develop a way to let the data determine the resolution of the descretization of the support space.
- Basic Idea:
  1. Run the model using a coarse mesh.
  2. Tag certain cells for refinement using a pre-specified refinement criteria.
  3. Split the tagged cells into 4 equally sized cells.
  4. Refit the model using the refined grid.
  5. Return to step 2 until no cells are refined or a max number of iterations is reached.
- Refinement criteria: k cells with the largest mean estimated squared error: $e_i = \frac{1}{\boldsymbol{D}(i,i)}\sum_{Y_j:\boldsymbol{\ell}_j\in N(i)}(Y_j-\hat{\phi}_j)^2$.
- Grid selection: compute a measure of goodness of fit in each iteration (AIC, BIC, cross validation score, etc.) and select the grid having the best score.

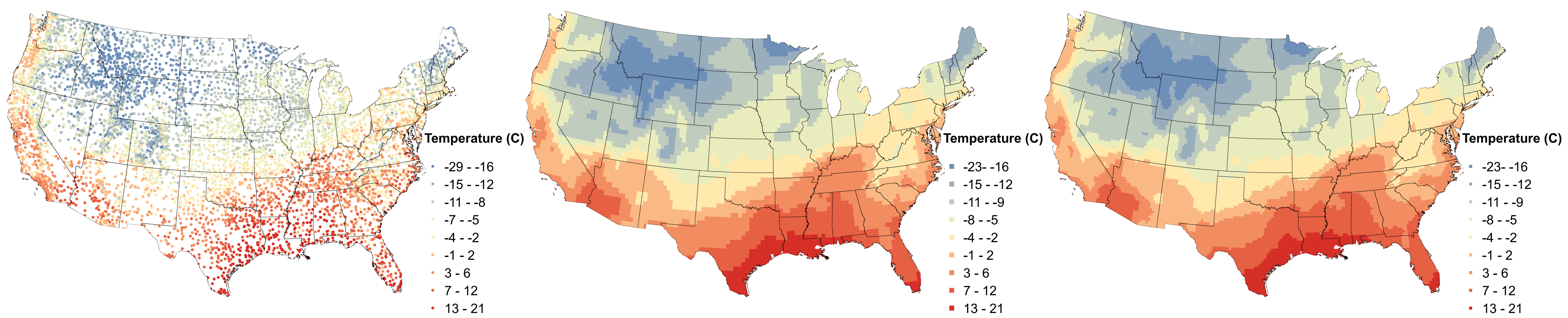
Simulation Results

Simulation results using data generated from an absolute value function (top row), and bimodel function (bottom row). From left to right the columns display the true function, the posterior mean estimate from the regular grid trend filter, the empirical bias from the regular grid trend filter, the posterior mean estimate from the adaptive mesh trend filter, and the empirical bias from the adaptive mesh trend filter.



Data Application

The data consists of the minimum daily temperature in degrees Celsius from January 1st, 2017 from 7276 weather stations across the United States. The figure shows the raw data (left), the estimated trend using the regular grid trend filter (center) and the estimated trend using the adaptive mesh trend filter (right).



Acknowledgements & Contact Information

- Partially funded by the Companion Animal Parasite Council
- Contact info: Stella Self, stellaw@clemson.edu