Spatial Prediction of Fine Particulate Matter in Taiwan

Hsin-Cheng Huang
Institute of Statistical Science, Academia Sinica

July 25, 2017

Joint work with

Guowen Huang (National Tsing Hua University, Taiwan)
Wen-Han Hwang (National Chung Hsing University, Taiwan)

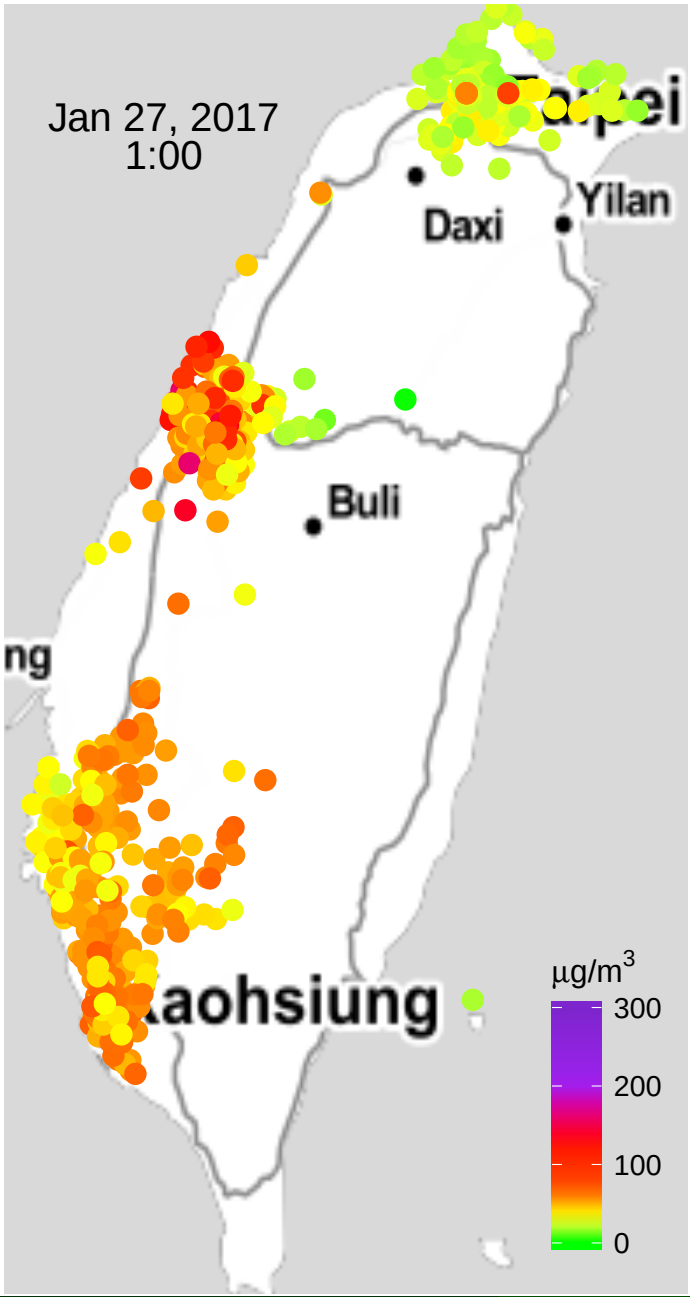
Outline

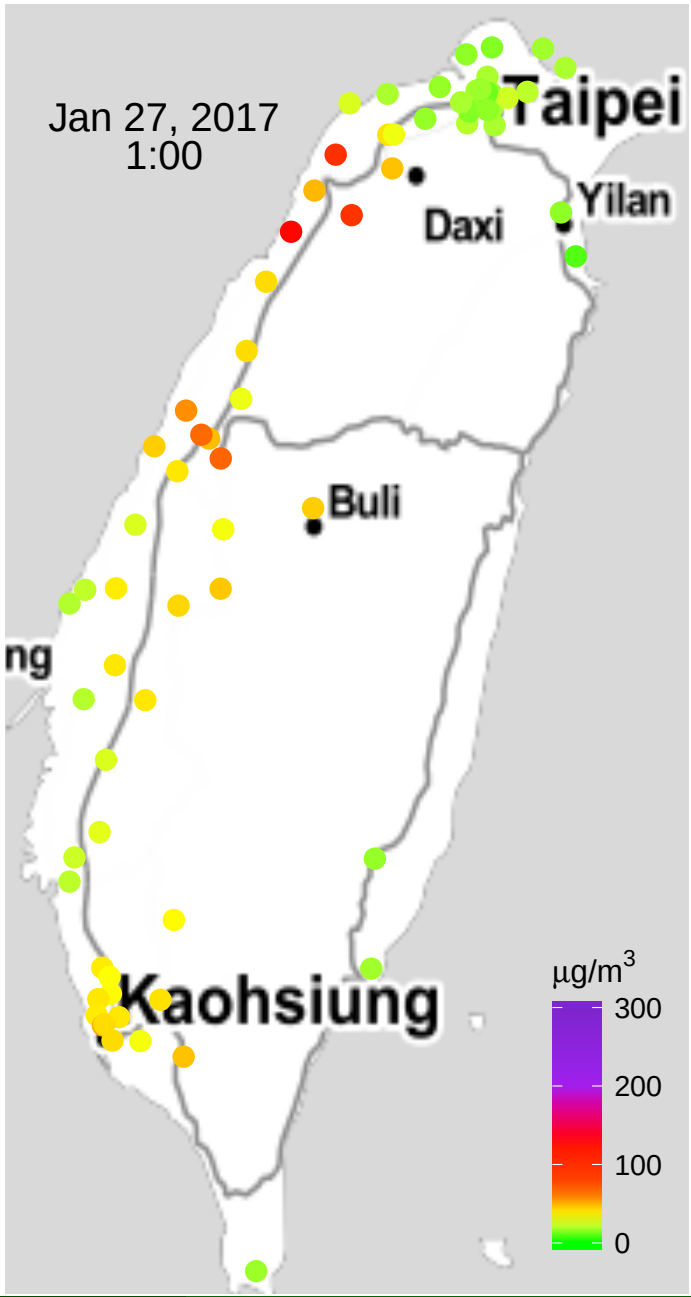
$PM_{2.5}$ Data

$PM_{2.5}$ Data Analysis

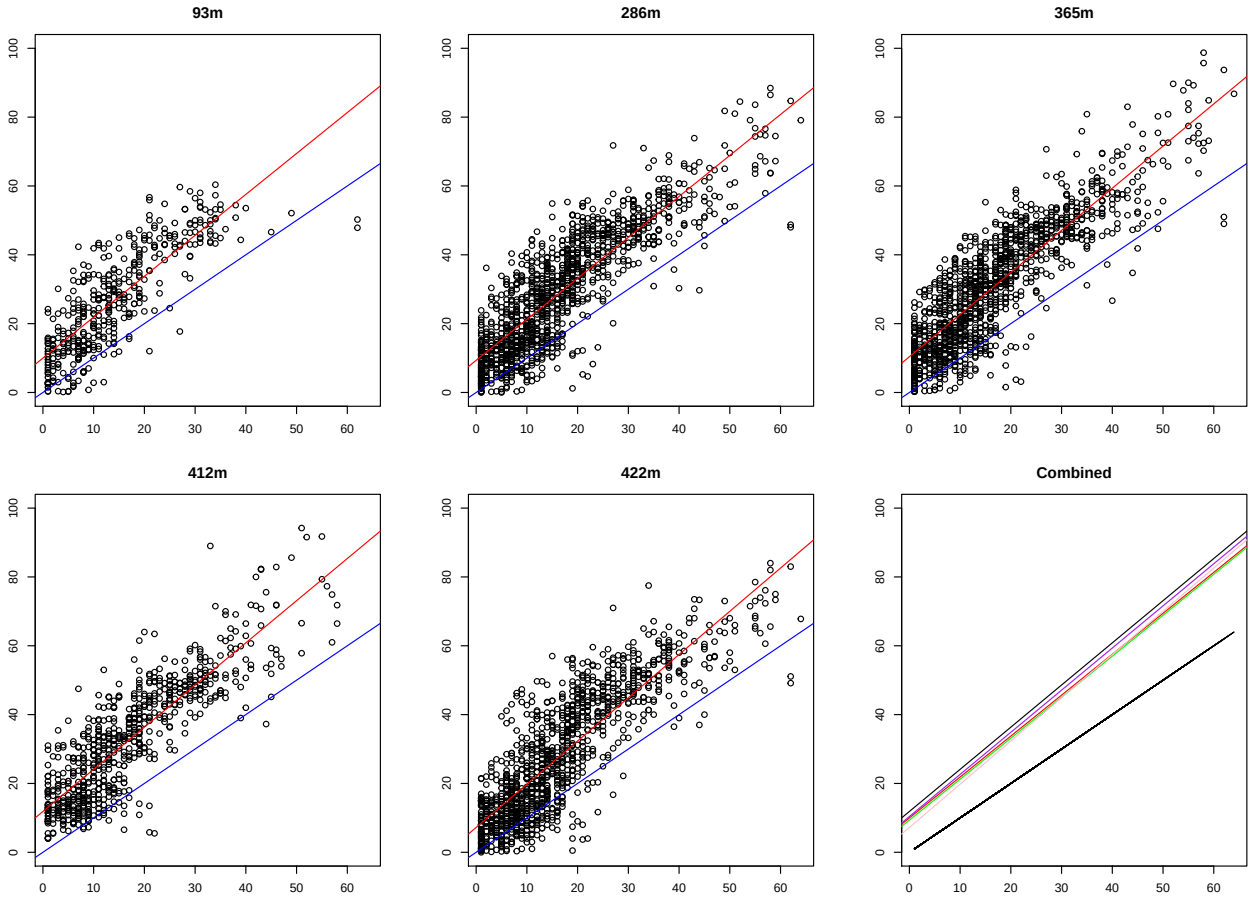
- Particulate matter less than 2.5 micrometers ($PM_{2.5}$) can cause cardiovascular and lung diseases
- Two data sources
  - 1,147 AirBoxes (available at `https://pm25.lass-net.org/`)
  - 76 Taiwan EPA monitoring stations
- Data: Hourly measurements from Jan 1 to Feb 28, 2017
- Goals
  - Spatial prediction of ground-level $PM_{2.5}$ concentration at any location in Taiwan
  - Detection of abnormal AirBox measurements and emission sources

1,147 AirBoxes (Left) & 76 Monitoring Stations (Right)



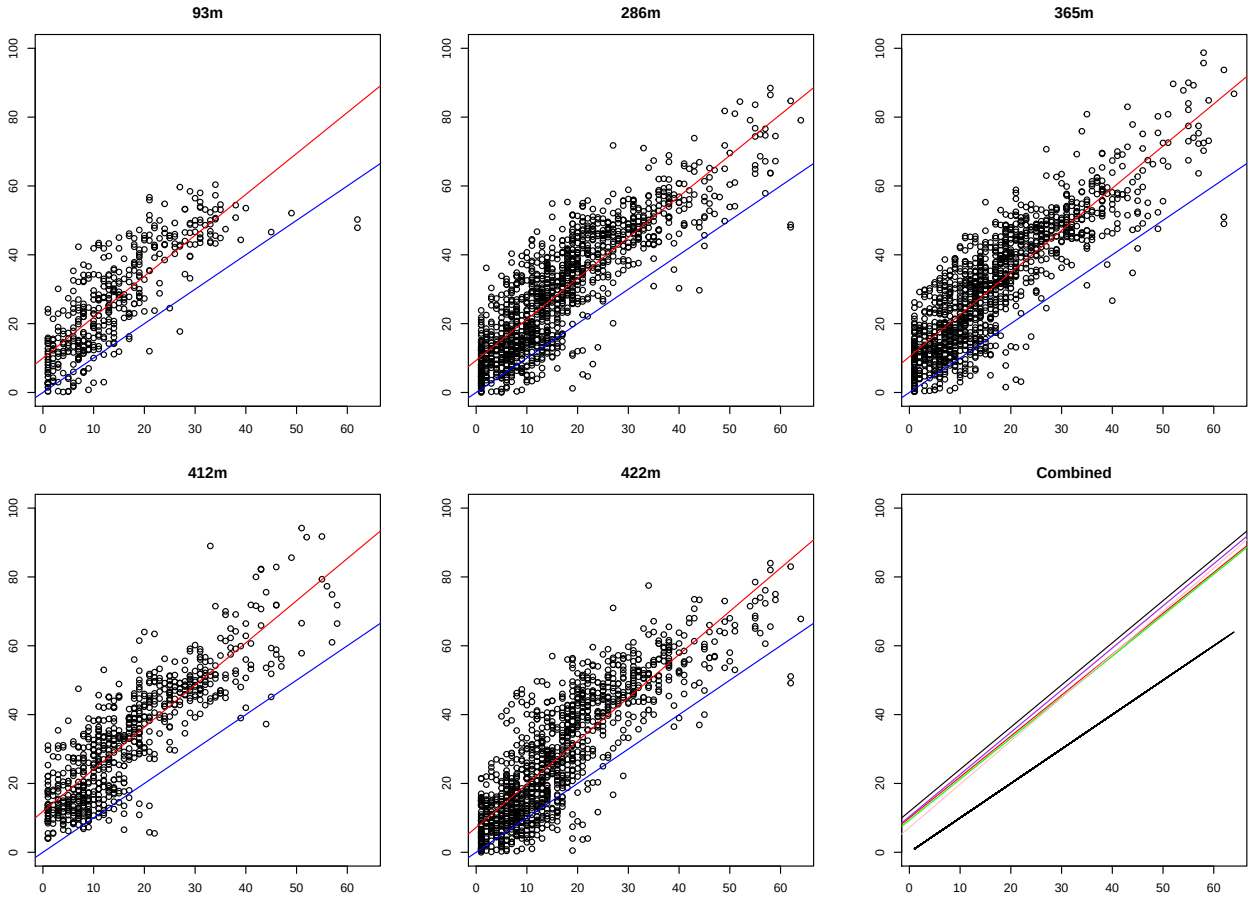


Calibration Between Two Types of Measurements



- x-axis: $PM_{2.5}$ measurements at an EPA monitoring station
- y-axis: $PM_{2.5}$ measurements at nearby AirBoxes
- Monitoring stations produce higher quality measurements but lower $PM_{2.5}$ values, because they are located at higher elevations

Calibration Between Two Types of Measurements



- x-axis: $PM_{2.5}$ measurements at an EPA monitoring station
- y-axis: $PM_{2.5}$ measurements at nearby AirBoxes
- Monitoring stations produce higher quality measurements but lower $PM_{2.5}$ values, because they are located at higher elevations

Resolution Adaptive Fixed Rank Kriging

Spatial Random Effects Models

- Spatial process of interest at time t; $t = 1, \ldots, T$

$$y_t(\boldsymbol{s}) = \sum_{k=1}^{K} w_{t,k}\varphi_k(\boldsymbol{s}); \quad \boldsymbol{s} \in D$$

  - $\boldsymbol{w}_t = (w_{t,1}, \ldots, w_{t,K})' \sim N(\boldsymbol{0}, \boldsymbol{M})$
  - $\varphi_1(\boldsymbol{s}), \ldots, \varphi_K(\boldsymbol{s})$ are basis functions
- Spatial covariance function

$$C_y(\boldsymbol{s}, \boldsymbol{s}^*) = \varphi(\boldsymbol{s})'\boldsymbol{M}\varphi(\boldsymbol{s}^*)$$

  - $\varphi(\boldsymbol{s}) = (\varphi_1(\boldsymbol{s}), \ldots, \varphi_K(\boldsymbol{s}))'$
- Data: $\boldsymbol{Z}_t = (z_t(\boldsymbol{s}_1), \ldots, z_t(\boldsymbol{s}_n))'$; $t = 1, \ldots, T$

$$z_t(\boldsymbol{s}_i) = y_t(\boldsymbol{s}_i) + \varepsilon_t(\boldsymbol{s}_i)$$

  - $\varepsilon_t(\boldsymbol{s}_i) \sim N(0, \sigma^2)$

Spatial Random Effects Models

- Spatial process of interest at time t; $t = 1, \ldots, T$

$$y_t(\boldsymbol{s}) = \sum_{k=1}^{K} w_{t,k}\varphi_k(\boldsymbol{s}); \quad \boldsymbol{s} \in D$$

  - $\boldsymbol{w}_t = (w_{t,1}, \ldots, w_{t,K})' \sim N(\boldsymbol{0}, \boldsymbol{M})$
  - $\varphi_1(\boldsymbol{s}), \ldots, \varphi_K(\boldsymbol{s})$ are basis functions
- Spatial covariance function

$$C_y(\boldsymbol{s}, \boldsymbol{s}^*) = \boldsymbol{\varphi}(\boldsymbol{s})'\boldsymbol{M}\boldsymbol{\varphi}(\boldsymbol{s}^*)$$

  - $\boldsymbol{\varphi}(\boldsymbol{s}) = (\varphi_1(\boldsymbol{s}), \ldots, \varphi_K(\boldsymbol{s}))'$
- Data: $\boldsymbol{Z}_t = (z_t(\boldsymbol{s}_1), \ldots, z_t(\boldsymbol{s}_n))'$; $t = 1, \ldots, T$

$$z_t(\boldsymbol{s}_i) = y_t(\boldsymbol{s}_i) + \varepsilon_t(\boldsymbol{s}_i)$$

  - $\varepsilon_t(\boldsymbol{s}_i) \sim N(0, \sigma^2)$

Spatial Random Effects Models

- Spatial process of interest at time t; $t = 1, \ldots, T$

$$y_t(\boldsymbol{s}) = \sum_{k=1}^{K} w_{t,k}\varphi_k(\boldsymbol{s}); \quad \boldsymbol{s} \in D$$

  - $\boldsymbol{w}_t = (w_{t,1}, \ldots, w_{t,K})' \sim N(\boldsymbol{0}, \boldsymbol{M})$
  - $\varphi_1(\boldsymbol{s}), \ldots, \varphi_K(\boldsymbol{s})$ are basis functions
- Spatial covariance function

$$C_y(\boldsymbol{s}, \boldsymbol{s}^*) = \boldsymbol{\varphi}(\boldsymbol{s})'\boldsymbol{M}\boldsymbol{\varphi}(\boldsymbol{s}^*)$$

  - $\boldsymbol{\varphi}(\boldsymbol{s}) = (\varphi_1(\boldsymbol{s}), \ldots, \varphi_K(\boldsymbol{s}))'$
- Data: $\boldsymbol{Z}_t = (z_t(\boldsymbol{s}_1), \ldots, z_t(\boldsymbol{s}_n))'$; $t = 1, \ldots, T$

$$z_t(\boldsymbol{s}_i) = y_t(\boldsymbol{s}_i) + \varepsilon_t(\boldsymbol{s}_i)$$

  - $\varepsilon_t(\boldsymbol{s}_i) \sim N(0, \sigma^2)$

Spatial Random Effects Models

- Minimum mean-squared error predictor of $y_t(\boldsymbol{s})$

$$\hat{y}_t(\boldsymbol{s}) = \varphi(\boldsymbol{s})'\boldsymbol{M}\boldsymbol{\Phi}'(\boldsymbol{\Phi}\boldsymbol{M}\boldsymbol{\Phi}' + \sigma^2\boldsymbol{I}_n)^{-1}\boldsymbol{Z}_t$$

  - $\boldsymbol{\Phi} = \left(\varphi_j(\boldsymbol{s}_i)\right)_{n\times K}$
  - $\varphi(\boldsymbol{s}) = (\varphi_1(\boldsymbol{s}), \ldots, \varphi_K(\boldsymbol{s}))'$
- Mean squared prediction error (prediction variance)

$$\mathrm{E}(\hat{y}_t(\boldsymbol{s}) - y_t(\boldsymbol{s}))^2 = \varphi(\boldsymbol{s})'\left\{\boldsymbol{M} - \boldsymbol{M}\boldsymbol{\Phi}'(\boldsymbol{\Phi}\boldsymbol{M}\boldsymbol{\Phi}' + \sigma^2\boldsymbol{I})^{-1}\boldsymbol{\Phi}\boldsymbol{M}\right\}\varphi(\boldsymbol{s})$$

- Sherman-Morrison-Woodbury formula

$$(\boldsymbol{\Phi}\boldsymbol{M}\boldsymbol{\Phi}' + \sigma^2\boldsymbol{I}_n)^{-1} = \frac{1}{\sigma^2}\boldsymbol{I}_n - \frac{1}{\sigma^4}\boldsymbol{\Phi}\left\{\boldsymbol{M}^{-} + \frac{1}{\sigma^2}\boldsymbol{\Phi}'\boldsymbol{\Phi}\right\}^{-1}\boldsymbol{\Phi}'$$

- Models for $\boldsymbol{M}$
  - `LatticeKrig` (Nychka *et al.*, 2015): $\boldsymbol{M}(\theta)$ depends on a small number of parameters θ
  - Fixed rank kriging (Cressie and Johannesson, 2008): No structure is imposed on $\boldsymbol{M}$

Spatial Random Effects Models

- Minimum mean-squared error predictor of $y_t(\boldsymbol{s})$

$$\hat{y}_t(\boldsymbol{s}) = \varphi(\boldsymbol{s})'\boldsymbol{M}\boldsymbol{\Phi}'(\boldsymbol{\Phi}\boldsymbol{M}\boldsymbol{\Phi}' + \sigma^2\boldsymbol{I}_n)^{-1}\boldsymbol{Z}_t$$

  - $\boldsymbol{\Phi} = \left(\varphi_j(\boldsymbol{s}_i)\right)_{n\times K}$
  - $\varphi(\boldsymbol{s}) = (\varphi_1(\boldsymbol{s}), \ldots, \varphi_K(\boldsymbol{s}))'$
- Mean squared prediction error (prediction variance)

$$\mathrm{E}(\hat{y}_t(\boldsymbol{s}) - y_t(\boldsymbol{s}))^2 = \varphi(\boldsymbol{s})'\left\{\boldsymbol{M} - \boldsymbol{M}\boldsymbol{\Phi}'(\boldsymbol{\Phi}\boldsymbol{M}\boldsymbol{\Phi}' + \sigma^2\boldsymbol{I})^{-1}\boldsymbol{\Phi}\boldsymbol{M}\right\}\varphi(\boldsymbol{s})$$

- Sherman-Morrison-Woodbury formula

$$(\boldsymbol{\Phi}\boldsymbol{M}\boldsymbol{\Phi}' + \sigma^2\boldsymbol{I}_n)^{-1} = \frac{1}{\sigma^2}\boldsymbol{I}_n - \frac{1}{\sigma^4}\boldsymbol{\Phi}\left\{\boldsymbol{M}^{-} + \frac{1}{\sigma^2}\boldsymbol{\Phi}'\boldsymbol{\Phi}\right\}^{-1}\boldsymbol{\Phi}'$$

- Models for $\boldsymbol{M}$
  - `LatticeKrig` (Nychka *et al.*, 2015): $\boldsymbol{M}(\theta)$ depends on a small number of parameters θ
  - Fixed rank kriging (Cressie and Johannesson, 2008): No structure is imposed on $\boldsymbol{M}$

Fixed Rank Kriging (FRK)

$$y_t(\boldsymbol{s}) = \sum_{k=1}^{K} w_{t,k}\varphi_k(\boldsymbol{s}), \quad \boldsymbol{w}_t = (w_{t,1}, \ldots, w_{t,K})' \sim N(\boldsymbol{0}, \boldsymbol{M})$$

- No structure is assumed on $\boldsymbol{M}$ except that it is non-negative definite
- Parameters: $\boldsymbol{M}$ and σ^2 (with $K(K+1)/2+1$ parameters)
- Advantages
  - Flexible nonstationary spatial covariance models
  - Handle massive amounts of data
- Issues
  1. Maximum likelihood estimation: EM algorithm (Katzfuss and Cressie, 2009)
  2. How to allocate basis functions?
  3. How many basis functions?

Improvement 1: Maximum Likelihood Estimation

- Parameters: $\boldsymbol{M}$ and σ^2 (with $K(K+1)/2+1$ parameters)

Theorem (Tzeng and Huang. 2017+, Technometrics)

$$\hat{\boldsymbol{M}} = (\Phi'\Phi)^{-1/2}\boldsymbol{P}\,\mathrm{diag}(\hat{d}_1,\ldots,\hat{d}_K)\boldsymbol{P}'(\Phi'\Phi)^{-1/2},$$

$$\hat{\sigma}^2 = \max\left\{\frac{1}{n-K^*}\left(\mathrm{tr}(\boldsymbol{S}) - \sum_{k=0}^{K^*} d_k\right), 0\right\}$$

- $\boldsymbol{P}\,\mathrm{diag}(d_1,\ldots,d_K)\boldsymbol{P}'$ *is the eigen-decomposition of*
$$(\Phi'\Phi)^{-1/2}\Phi'\boldsymbol{S}\,\Phi(\Phi'\Phi)^{-1/2}$$
- $\boldsymbol{S} = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{Z}_t\boldsymbol{Z}_t'$
- $d_0 \equiv 0$ *and* $\hat{d}_k = \max(d_k - \hat{\sigma}^2, 0)$*;* $k = 1,\ldots,K$
- $K^* = \max\left(\{0\} \cup \left\{L : d_L > \frac{1}{n-L}\left(\mathrm{tr}(\boldsymbol{S}) - \sum_{k=0}^{L} d_k\right)\right\}\right)$

Improvement 1: Maximum Likelihood Estimation

- Parameters: $\boldsymbol{M}$ and σ^2 (with $K(K+1)/2+1$ parameters)

Theorem (Tzeng and Huang. 2017+, Technometrics)

$$\hat{\boldsymbol{M}} = (\boldsymbol{\Phi}'\boldsymbol{\Phi})^{-1/2}\boldsymbol{P}\,\mathrm{diag}(\hat{d}_1,\ldots,\hat{d}_K)\boldsymbol{P}'(\boldsymbol{\Phi}'\boldsymbol{\Phi})^{-1/2},$$

$$\hat{\sigma}^2 = \max\left\{\frac{1}{n-K^*}\left(\mathrm{tr}(\boldsymbol{S}) - \sum_{k=0}^{K^*} d_k\right), 0\right\}$$

- $\boldsymbol{P}\,\mathrm{diag}(d_1,\ldots,d_K)\boldsymbol{P}'$ *is the eigen-decomposition of*

$$(\boldsymbol{\Phi}'\boldsymbol{\Phi})^{-1/2}\boldsymbol{\Phi}'\boldsymbol{S}\,\boldsymbol{\Phi}(\boldsymbol{\Phi}'\boldsymbol{\Phi})^{-1/2}$$

- $\boldsymbol{S} = \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{Z}_t\boldsymbol{Z}_t'$
- $d_0 \equiv 0$ *and* $\hat{d}_k = \max(d_k - \hat{\sigma}^2, 0)$*;* $k = 1,\ldots,K$
- $K^* = \max\left(\{0\} \cup \left\{L : d_L > \frac{1}{n-L}\left(\mathrm{tr}(\boldsymbol{S}) - \sum_{k=0}^{L} d_k\right)\right\}\right)$

Improvement 2: Multi-Resolution Basis Functions

- $J(f) = \int_{\mathbb{R}^d} \sum_{\nu_1+\cdots+\nu_d=2} \frac{2}{\nu_1!\cdots\nu_d!} \left(\frac{\partial^2 f(\boldsymbol{s})}{\partial x_1^{\nu_1} \cdots \partial x_d^{\nu_d}} \right)^2 d\boldsymbol{s}; \quad \boldsymbol{s} = (x_1, \ldots, x_d)'$

Definition (Multiresolution spline basis functions)

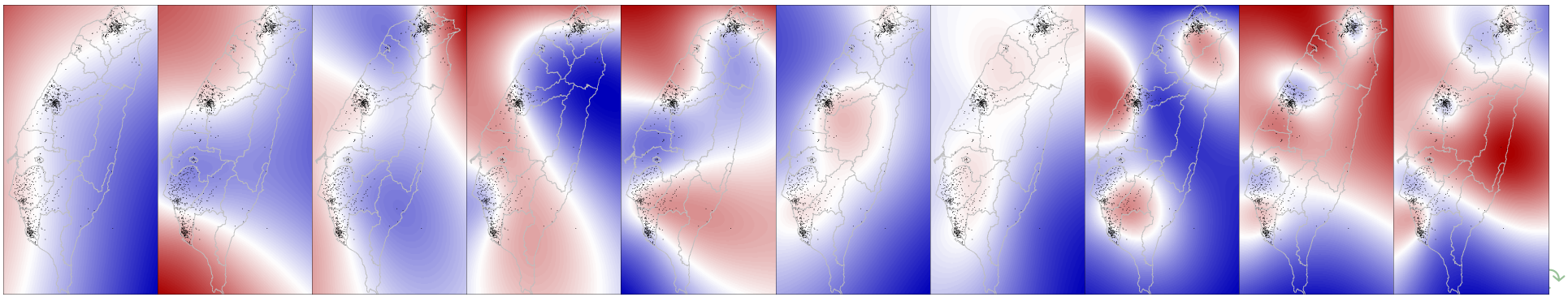
$$\varphi_1(\boldsymbol{s}) = 1$$

$$\varphi_k(\boldsymbol{s}) = x_k; \quad k = 2, \ldots, d+1$$

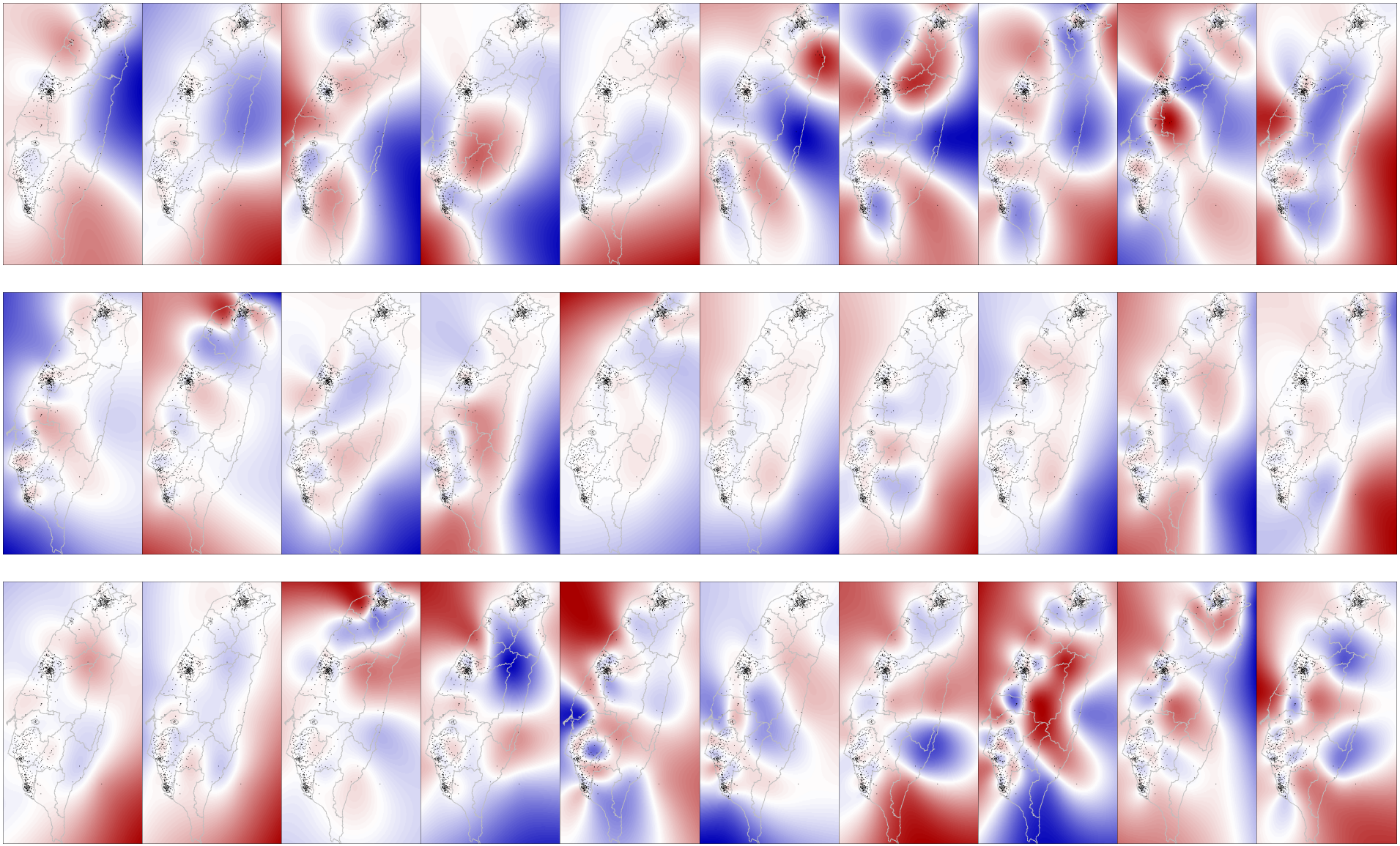
$$\varphi_k(\boldsymbol{s}) = \arg\min_{g(\cdot)} \left\{ J(g) : \boldsymbol{g}'\phi_1 = \cdots = \boldsymbol{g}'\phi_{k-1} = 0, \|\boldsymbol{g}\| = 1 \right\};$$

$$k = d+2, \ldots, n$$

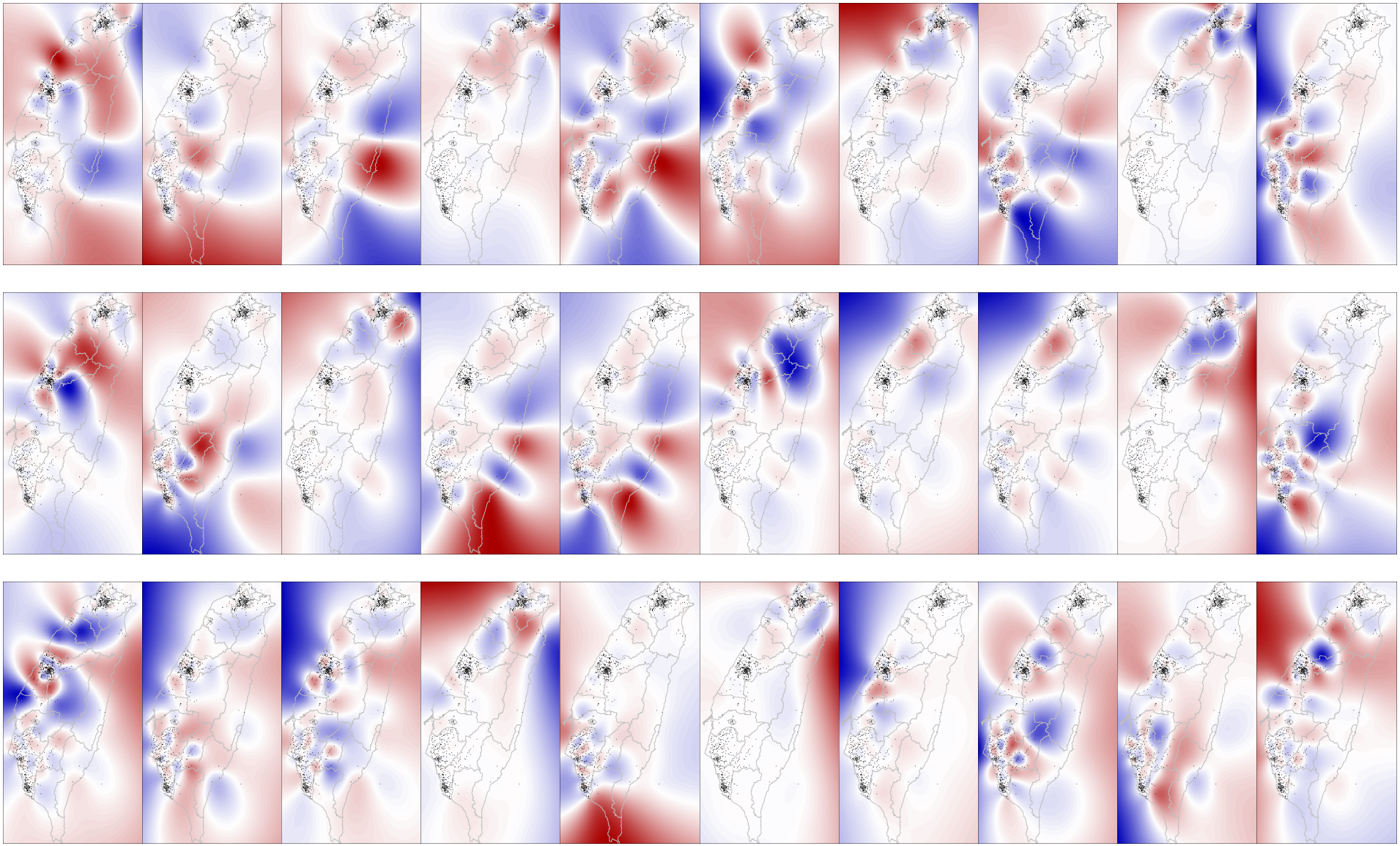
- $\phi_j = (\varphi_j(\boldsymbol{s}_1), \ldots, \varphi_j(\boldsymbol{s}_n))'$ and $\boldsymbol{g} = (g(\boldsymbol{s}_1), \ldots, g(\boldsymbol{s}_n))'$
- $0 = J(\varphi_1) = \cdots = J(\varphi_{d+1}) < J(\varphi_{d+2}) \leq \cdots \leq J(\varphi_n)$



Basis Functions: $\varphi_{14}(\cdot), \ldots, \varphi_{43}(\cdot)$



Basis Functions: $\varphi_{44}(\cdot), \ldots, \varphi_{73}(\cdot)$



Improvement 2: Multi-Resolution Basis Functions

Proposition

$$\varphi_k(\boldsymbol{s}) = \begin{cases} 1; & k=1 \\ x_k; & k=2,\ldots,d+1 \\ \dfrac{1}{\lambda_{k-d-1}}\left\{\psi(\boldsymbol{s}) - \boldsymbol{\Psi}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{x}\right\}'\boldsymbol{v}_{k-d-1}; & k=d+2,\ldots,n \end{cases}$$

- $\boldsymbol{x} = (1, x_1, \ldots, x_d)'$
- $\boldsymbol{X}$ *is an* $n \times (d+1)$ *matrix with the* i*-th row* $(1, x_{i1}, \ldots, x_{id})$ *corresponding to* $\boldsymbol{s}_i$
- λ_k *and* $\boldsymbol{v}_k$ *are the* k*-th eigenvalue and eigenvector of* $\boldsymbol{Q}\boldsymbol{\Psi}\boldsymbol{Q}$ *with* $\lambda_1 \geq \cdots \geq \lambda_n$
- $\boldsymbol{Q} = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$ *and* $\boldsymbol{\Psi} = \left(\psi_j(\boldsymbol{s}_i)\right)_{n\times n}$
- $\psi(\boldsymbol{s}) = (\psi_1(\boldsymbol{s}), \ldots, \psi_n(\boldsymbol{s}))'$
- $\psi_j(\boldsymbol{s}) = \dfrac{1}{8\pi}\|\boldsymbol{s} - \boldsymbol{s}_j\|^2 \log\left(\|\boldsymbol{s} - \boldsymbol{s}_j\|\right), \quad$ *if* $d = 2$

Improvement 3: Selection of K

- Akaike's information criterion (Akaike, 1973)

$$\mathrm{AIC}(K) = \underbrace{-2 \quad \ell(\hat{\boldsymbol{M}}, \hat{\sigma}^2)}_{\text{log-likelihood}} + \underbrace{2 \quad \mathrm{df}(K, T)}_{\text{\# free parameters}}$$

$$= \frac{T\,\mathrm{tr}(\boldsymbol{S})}{\hat{\sigma}^2} + T\sum_{k=1}^{K}\left\{\log\left(\hat{d}_k + \hat{\sigma}^2\right) - \frac{d_k \hat{d}_k}{\hat{\sigma}^2\left(\hat{d}_k + \hat{\sigma}^2\right)}\right\} + 2\,\mathrm{df}(K, T)$$

  - $\mathrm{df}(K, T) = \begin{cases} K(K+1)/2 + 1; & \text{if } K \le T \\ KT - T(T-1)/2 + 1; & \text{if } K > T \end{cases}$

- $\hat{K} = \underset{d+1 \le K \le K^*}{\arg\min}\ \mathrm{AIC}(K)$

Resolution Adaptive FRK

- Automatic FRK (Tzeng and Huang, 2017+)

$$\hat{y}_t(\boldsymbol{s}) = \boldsymbol{\varphi}(\boldsymbol{s})'\hat{\boldsymbol{M}}\boldsymbol{\Phi}'(\boldsymbol{\Phi}\hat{\boldsymbol{M}}\boldsymbol{\Phi}' + \hat{\sigma}^2\boldsymbol{I})^{-1}\boldsymbol{Z}_t,$$

$$(\boldsymbol{\Phi}\hat{\boldsymbol{M}}\boldsymbol{\Phi}' + \hat{\sigma}^2\boldsymbol{I})^{-1} = \frac{1}{\hat{\sigma}^2}\boldsymbol{I}_n - \frac{1}{\hat{\sigma}^4}\boldsymbol{\Phi}\left\{\hat{\boldsymbol{M}}^{-1} + \frac{1}{\hat{\sigma}^2}\boldsymbol{\Phi}'\boldsymbol{\Phi}\right\}^{-1}\boldsymbol{\Phi}'$$

  - $\boldsymbol{\Phi}$: simple closed-form expression (multiresolution spline basis functions)
  - $\hat{\boldsymbol{M}}$ and $\hat{\sigma}^2$: simple closed-form expressions $\Rightarrow$ allow large K
  - $\mathrm{AIC}(K)$: simple closed-form expression
- Numerically stable with no undesirable artifacts in predicted surfaces
- Directly applicable to sparse or irregularly spaced data

Resolution Adaptive FRK

- Automatic FRK (Tzeng and Huang, 2017+)

$$\hat{y}_t(\boldsymbol{s}) = \boldsymbol{\varphi}(\boldsymbol{s})'\hat{\boldsymbol{M}}\boldsymbol{\Phi}'(\boldsymbol{\Phi}\hat{\boldsymbol{M}}\boldsymbol{\Phi}' + \hat{\sigma}^2\boldsymbol{I})^{-1}\boldsymbol{Z}_t,$$

$$(\boldsymbol{\Phi}\hat{\boldsymbol{M}}\boldsymbol{\Phi}' + \hat{\sigma}^2\boldsymbol{I})^{-1} = \frac{1}{\hat{\sigma}^2}\boldsymbol{I}_n - \frac{1}{\hat{\sigma}^4}\boldsymbol{\Phi}\left\{\hat{\boldsymbol{M}}^{-1} + \frac{1}{\hat{\sigma}^2}\boldsymbol{\Phi}'\boldsymbol{\Phi}\right\}^{-1}\boldsymbol{\Phi}'$$

  - $\boldsymbol{\Phi}$: simple closed-form expression (multiresolution spline basis functions)
  - $\hat{\boldsymbol{M}}$ and $\hat{\sigma}^2$: simple closed-form expressions $\Rightarrow$ allow large K
  - $\mathrm{AIC}(K)$: simple closed-form expression
- Numerically stable with no undesirable artifacts in predicted surfaces
- Directly applicable to sparse or irregularly spaced data

Extensions

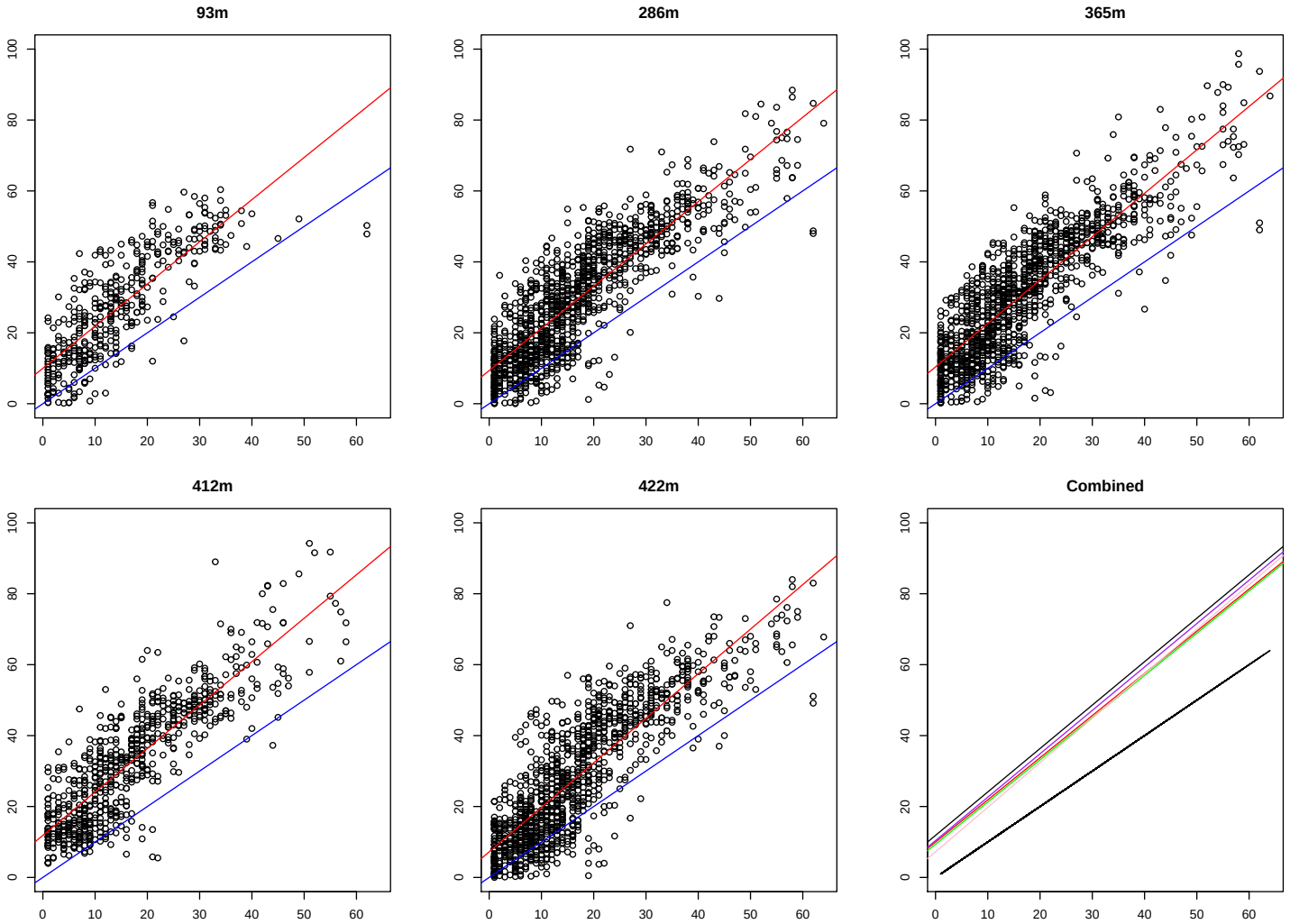
Spatial Random Effects Model with Nonzero Means

- Spatial $PM_{2.5}$ process at time $t = 1, \ldots, T$

$$y_t(\boldsymbol{s}) = \boldsymbol{w}_t' \boldsymbol{\varphi}(\boldsymbol{s}) = \sum_{k=1}^{K} w_{t,k} \varphi_k(\boldsymbol{s}); \quad \boldsymbol{s} \in D$$

  - $\boldsymbol{w}_t = (w_{t,1}, \ldots, w_{t,K})' \sim N(\boldsymbol{\mu}, \boldsymbol{M})$
  - $\varphi_1(\boldsymbol{s}), \ldots, \varphi_K(\boldsymbol{s})$ is an order set of multi-resolution basis functions
- Parameters: $\boldsymbol{\mu}$, $\boldsymbol{M}$ and σ^2
  - Given $\boldsymbol{\mu}$, the ML estimators of $\boldsymbol{M}$ and σ^2 have simple closed-form expressions
  - Given $\boldsymbol{M}$ and σ^2, the ML estimator of $\boldsymbol{\mu}$ has a simple closed-form expression

Spatial Prediction Based on Two Types of Data



- x-axis: $PM_{2.5}$ measurements at an EPA monitoring station
- y-axis: $PM_{2.5}$ measurements at nearby AirBoxes

Spatial Prediction Based on Two Types of Data

$$y_t(\boldsymbol{s}) = \boldsymbol{w}_t' \varphi(\boldsymbol{s}), \quad \boldsymbol{w}_t \sim N(\mu, \boldsymbol{M})$$

- Data
  - Airboxes: $\boldsymbol{Z}_t = (z_t(\boldsymbol{s}_1), \ldots, z_t(\boldsymbol{s}_m))'$
  - Monitoring stations: $\boldsymbol{Z}_t^* = (z_t^*(\boldsymbol{s}_{m+1}), \ldots, z_t^*(\boldsymbol{s}_n^*))'$
- Measurement equations

$$z_t(\boldsymbol{s}_i) = y_t(\boldsymbol{s}_i) + \varepsilon_t(\boldsymbol{s}_i); \quad i = 1, \ldots, m,$$
$$z_t^*(\boldsymbol{s}_j) = \alpha_0 + \alpha_1 y_t(\boldsymbol{s}_i) + \varepsilon_t^*(\boldsymbol{s}_i); \quad i = m+1, \ldots, n$$

  - $\varepsilon_t(\boldsymbol{s}_1), \ldots, \varepsilon_t(\boldsymbol{s}_m) \sim N(0, \sigma^2)$
  - $\varepsilon_t^*(\boldsymbol{s}_{m+1}), \ldots, \varepsilon_t^*(\boldsymbol{s}_n) \sim N(0, r\sigma^2)$ with $r < 1$
- Parameters: $\alpha_0, \alpha_1, r, \boldsymbol{M}$ and σ^2
- Given α_1 and r, the ML estimators of α_0, $\boldsymbol{M}$ and σ^2 have closed-form expressions

Spatial Prediction Based on Two Types of Data

$$y_t(\boldsymbol{s}) = \boldsymbol{w}_t' \varphi(\boldsymbol{s}), \quad \boldsymbol{w}_t \sim N(\mu, \boldsymbol{M})$$

- Data
  - Airboxes: $\boldsymbol{Z}_t = (z_t(\boldsymbol{s}_1), \ldots, z_t(\boldsymbol{s}_m))'$
  - Monitoring stations: $\boldsymbol{Z}_t^* = (z_t^*(\boldsymbol{s}_{m+1}), \ldots, z_t^*(\boldsymbol{s}_n^*))'$
- Measurement equations

$$z_t(\boldsymbol{s}_i) = y_t(\boldsymbol{s}_i) + \varepsilon_t(\boldsymbol{s}_i); \quad i = 1, \ldots, m,$$
$$z_t^*(\boldsymbol{s}_j) = \alpha_0 + \alpha_1 \, y_t(\boldsymbol{s}_i) + \varepsilon_t^*(\boldsymbol{s}_i); \quad i = m+1, \ldots, n$$

  - $\varepsilon_t(\boldsymbol{s}_1), \ldots, \varepsilon_t(\boldsymbol{s}_m) \sim N(0, \sigma^2)$
  - $\varepsilon_t^*(\boldsymbol{s}_{m+1}), \ldots, \varepsilon_t^*(\boldsymbol{s}_n) \sim N(0, r\sigma^2)$ with $r < 1$
- Parameters: $\alpha_0, \alpha_1, r, \boldsymbol{M}$ and σ^2
- Given α_1 and r, the ML estimators of α_0, $\boldsymbol{M}$ and σ^2 have closed-form expressions

Spatial Prediction Based on Two Types of Data

$$y_t(\boldsymbol{s}) = \boldsymbol{w}_t' \varphi(\boldsymbol{s}), \quad \boldsymbol{w}_t \sim N(\mu, \boldsymbol{M})$$

- Data
  - Airboxes: $\boldsymbol{Z}_t = (z_t(\boldsymbol{s}_1), \ldots, z_t(\boldsymbol{s}_m))'$
  - Monitoring stations: $\boldsymbol{Z}_t^* = (z_t^*(\boldsymbol{s}_{m+1}), \ldots, z_t^*(\boldsymbol{s}_n^*))'$
- Measurement equations

$$z_t(\boldsymbol{s}_i) = y_t(\boldsymbol{s}_i) + \varepsilon_t(\boldsymbol{s}_i); \quad i = 1, \ldots, m,$$
$$z_t^*(\boldsymbol{s}_j) = \alpha_0 + \alpha_1 y_t(\boldsymbol{s}_i) + \varepsilon_t^*(\boldsymbol{s}_i); \quad i = m+1, \ldots, n$$

  - $\varepsilon_t(\boldsymbol{s}_1), \ldots, \varepsilon_t(\boldsymbol{s}_m) \sim N(0, \sigma^2)$
  - $\varepsilon_t^*(\boldsymbol{s}_{m+1}), \ldots, \varepsilon_t^*(\boldsymbol{s}_n) \sim N(0, r\sigma^2)$ with $r < 1$
- Parameters: $\alpha_0, \alpha_1, r, \boldsymbol{M}$ and σ^2
- Given α_1 and r, the ML estimators of α_0, $\boldsymbol{M}$ and σ^2 have closed-form expressions

Spatial Prediction Based on Two Types of Data

$$y_t(\boldsymbol{s}) = \boldsymbol{w}_t' \varphi(\boldsymbol{s}), \quad \boldsymbol{w}_t \sim N(\mu, \boldsymbol{M})$$

- Data
  - Airboxes: $\boldsymbol{Z}_t = (z_t(\boldsymbol{s}_1), \ldots, z_t(\boldsymbol{s}_m))'$
  - Monitoring stations: $\boldsymbol{Z}_t^* = (z_t^*(\boldsymbol{s}_{m+1}), \ldots, z_t^*(\boldsymbol{s}_n^*))'$
- Measurement equations

$$z_t(\boldsymbol{s}_i) = y_t(\boldsymbol{s}_i) + \varepsilon_t(\boldsymbol{s}_i); \quad i = 1, \ldots, m,$$
$$z_t^*(\boldsymbol{s}_j) = \alpha_0 + \alpha_1 y_t(\boldsymbol{s}_i) + \varepsilon_t^*(\boldsymbol{s}_i); \quad i = m+1, \ldots, n$$

  - $\varepsilon_t(\boldsymbol{s}_1), \ldots, \varepsilon_t(\boldsymbol{s}_m) \sim N(0, \sigma^2)$
  - $\varepsilon_t^*(\boldsymbol{s}_{m+1}), \ldots, \varepsilon_t^*(\boldsymbol{s}_n) \sim N(0, r\sigma^2)$ with $r < 1$
- Parameters: $\alpha_0, \alpha_1, r, \boldsymbol{M}$ and σ^2
- Given α_1 and r, the ML estimators of α_0, $\boldsymbol{M}$ and σ^2 have closed-form expressions

Spatial Prediction Based on Two Types of Data

- Minimum mean squared error predictor of $y_t(\boldsymbol{s})$

$$\hat{y}_t(\boldsymbol{s}) = \boldsymbol{\varphi}(\boldsymbol{s})'\boldsymbol{\mu} + \boldsymbol{\varphi}(\boldsymbol{s})'\boldsymbol{M}\boldsymbol{\Phi}'\boldsymbol{A}'(\boldsymbol{A}\boldsymbol{\Phi}\boldsymbol{M}\boldsymbol{\Phi}'\boldsymbol{A}' + \boldsymbol{V})^{-1} \begin{pmatrix} \boldsymbol{Z}_t - \boldsymbol{f}_t \\ \boldsymbol{Z}_t^* - \alpha_0\boldsymbol{1} - \alpha_1\boldsymbol{f}_t^* \end{pmatrix}$$

  - $\boldsymbol{f}_t = \big(\boldsymbol{\varphi}(\boldsymbol{s}_1), \ldots, \boldsymbol{\varphi}(\boldsymbol{s}_m)\big)\boldsymbol{\mu}$
  - $\boldsymbol{f}_t^* = \big(\boldsymbol{\varphi}(\boldsymbol{s}_{m+1}), \ldots, \boldsymbol{\varphi}(\boldsymbol{s}_n)\big)\boldsymbol{\mu}$
  - $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{I}_m & \boldsymbol{0} \\ \boldsymbol{0} & \alpha_1\boldsymbol{I}_{n-m} \end{pmatrix}$
  - $\boldsymbol{V} = \begin{pmatrix} \sigma^2\boldsymbol{I}_m & \boldsymbol{0} \\ \boldsymbol{0} & r\sigma^2\boldsymbol{I}_{n-m} \end{pmatrix}$

- Sherman-Morrison-Woodbury formula

$$(\boldsymbol{A}\boldsymbol{\Phi}\boldsymbol{M}\boldsymbol{\Phi}'\boldsymbol{A}' + \boldsymbol{V})^{-1} = \boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{\Phi}\big\{\boldsymbol{M}^{-} + \boldsymbol{A}\boldsymbol{\Phi}'\boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{\Phi}\big\}^{-1}\boldsymbol{\Phi}'\boldsymbol{A}'\boldsymbol{V}^{-1}$$

Other Extensions

- Some data are missing: $\boldsymbol{Z}_t = \begin{pmatrix} \boldsymbol{Z}_t^{(obs)} \\ \boldsymbol{Z}_t^{(mis)} \end{pmatrix}$
  - Compute the ML estimators using the EM algorithm
  - Treat $\boldsymbol{Z}_t^{(mis)}$ as missing data directly instead of $\boldsymbol{w}_t$
- Models with covariates

$$\boldsymbol{Z}_t \sim N(\boldsymbol{X}\beta,\ \Phi \boldsymbol{M} \Phi' + \sigma^2 \boldsymbol{I}_n)$$

  - $\boldsymbol{X}$: $n \times (p+1)$ design matrix of rank $(p+1)$
  - β: Regression parameter vector
  - Restricted ML estimators of $\boldsymbol{M}$ and σ^2 have closed-form expressions
- Space-time models

$$y_t(\boldsymbol{s}) = \boldsymbol{w}_t' \varphi(\boldsymbol{s}), \quad \boldsymbol{w}_t \sim N(\mu, \boldsymbol{M}),$$
$$\boldsymbol{w}_t = \boldsymbol{A}\boldsymbol{w}_{t-1} + \eta_t \quad \text{(multivariate stationary AR(1))},$$
$$\boldsymbol{Z}_t = \Phi \boldsymbol{w}_t + \varepsilon_t, \quad \varepsilon_t \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$$

  - Given $\boldsymbol{A}$, the ML estimators of $\boldsymbol{M}$ and σ^2 have closed-form expressions

Other Extensions

- Some data are missing: $\boldsymbol{Z}_t = \begin{pmatrix} \boldsymbol{Z}_t^{(obs)} \\ \boldsymbol{Z}_t^{(mis)} \end{pmatrix}$
  - Compute the ML estimators using the EM algorithm
  - Treat $\boldsymbol{Z}_t^{(mis)}$ as missing data directly instead of $\boldsymbol{w}_t$
- Models with covariates

$$\boldsymbol{Z}_t \sim N(\boldsymbol{X}\beta,\ \boldsymbol{\Phi}\boldsymbol{M}\boldsymbol{\Phi}' + \sigma^2 \boldsymbol{I}_n)$$

  - $\boldsymbol{X}$: $n \times (p+1)$ design matrix of rank $(p+1)$
  - β: Regression parameter vector
  - Restricted ML estimators of $\boldsymbol{M}$ and σ^2 have closed-form expressions
- Space-time models

$$y_t(\boldsymbol{s}) = \boldsymbol{w}_t'\varphi(\boldsymbol{s}), \quad \boldsymbol{w}_t \sim N(\mu, \boldsymbol{M}),$$
$$\boldsymbol{w}_t = \boldsymbol{A}\boldsymbol{w}_{t-1} + \eta_t \quad \text{(multivariate stationary AR(1))},$$
$$\boldsymbol{Z}_t = \boldsymbol{\Phi}\boldsymbol{w}_t + \varepsilon_t, \quad \varepsilon_t \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$$

  - Given $\boldsymbol{A}$, the ML estimators of $\boldsymbol{M}$ and σ^2 have closed-form expressions

Other Extensions

- Some data are missing: $\boldsymbol{Z}_t = \begin{pmatrix} \boldsymbol{Z}_t^{(obs)} \\ \boldsymbol{Z}_t^{(mis)} \end{pmatrix}$
  - Compute the ML estimators using the EM algorithm
  - Treat $\boldsymbol{Z}_t^{(mis)}$ as missing data directly instead of $\boldsymbol{w}_t$
- Models with covariates

$$\boldsymbol{Z}_t \sim N\left(\boldsymbol{X}\beta,\ \boldsymbol{\Phi}\boldsymbol{M}\boldsymbol{\Phi}' + \sigma^2 \boldsymbol{I}_n\right)$$

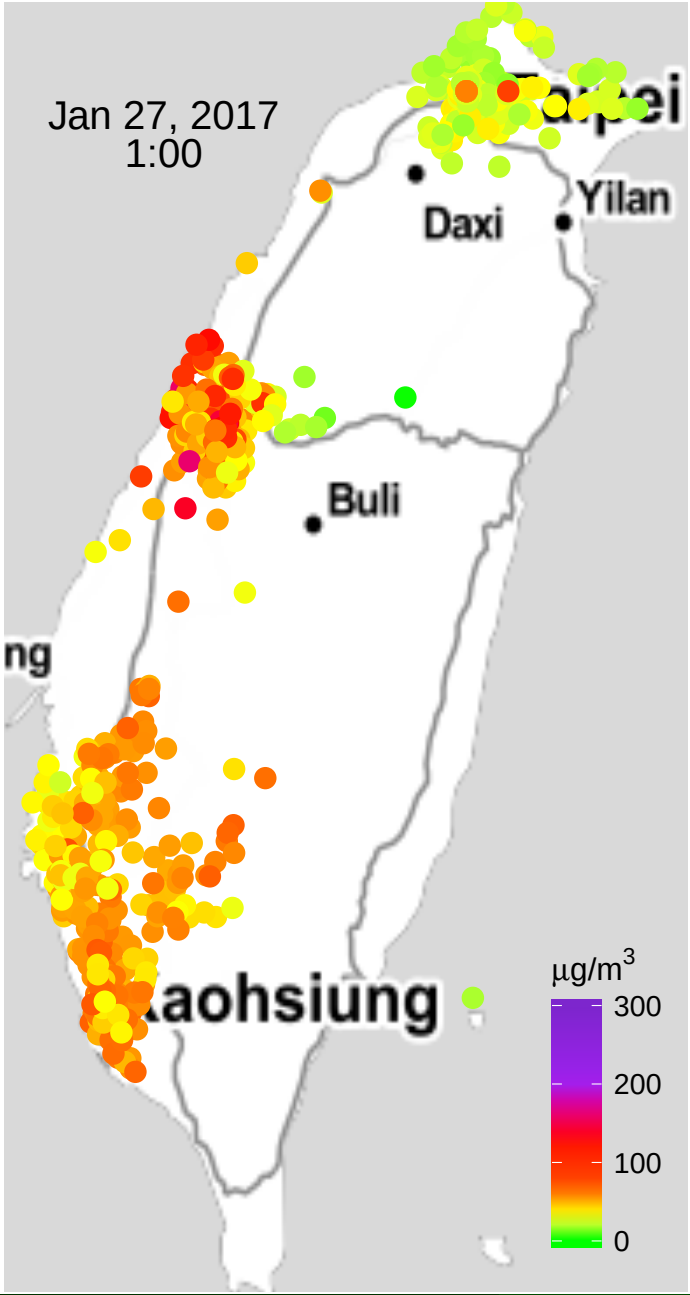
  - $\boldsymbol{X}$: $n \times (p+1)$ design matrix of rank $(p+1)$
  - β: Regression parameter vector
  - Restricted ML estimators of $\boldsymbol{M}$ and σ^2 have closed-form expressions
- Space-time models

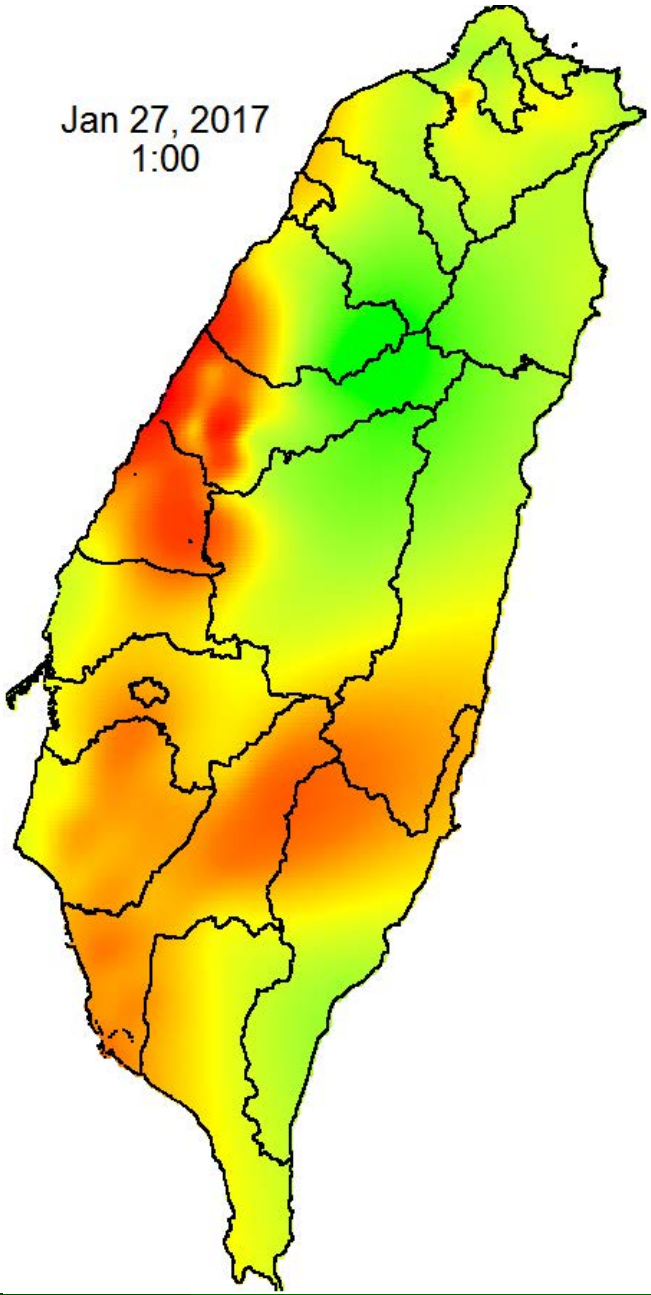
$$y_t(\boldsymbol{s}) = \boldsymbol{w}_t' \varphi(\boldsymbol{s}), \quad \boldsymbol{w}_t \sim N(\mu, \boldsymbol{M}),$$
$$\boldsymbol{w}_t = \boldsymbol{A}\boldsymbol{w}_{t-1} + \eta_t \quad \text{(multivariate stationary AR(1))},$$
$$\boldsymbol{Z}_t = \boldsymbol{\Phi}\boldsymbol{w}_t + \varepsilon_t, \quad \varepsilon_t \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$$

  - Given $\boldsymbol{A}$, the ML estimators of $\boldsymbol{M}$ and σ^2 have closed-form expressions

$PM_{2.5}$ Applications

Spatial Prediction of $PM_{2.5}$





Detect Abnormal AirBox Measurements

- AirBoxes closer together in space are likely to produce similar measurements
- $\hat{y}_t(\boldsymbol{s})$: Predicted surface obtained from the proposed fixed rank kriging
- Detect outliers by monitoring standardized residuals

$$r_t(\boldsymbol{s}_i) = \frac{z_t(\boldsymbol{s}_i) - \hat{y}_t(\boldsymbol{s}_i)}{v_t}; \quad i = 1, \ldots, n,\ t = 1, \ldots, T$$

- Monitor excessively large values of $r_t^2(\boldsymbol{s}_i)$
  - Small (negative) $r_t(\boldsymbol{s}_i)$ corresponds to the AirBoxes that may be put indoors at $\boldsymbol{s}_l$
  - Large (positive) $r_t(\boldsymbol{s}_i)$ corresponds to potential emission sources of $PM_{2.5}$ around location $\boldsymbol{s}_i$ and time t

Detect Abnormal AirBox Measurements

- AirBoxes closer together in space are likely to produce similar measurements
- $\hat{y}_t(\boldsymbol{s})$: Predicted surface obtained from the proposed fixed rank kriging
- Detect outliers by monitoring standardized residuals

$$r_t(\boldsymbol{s}_i) = \frac{z_t(\boldsymbol{s}_i) - \hat{y}_t(\boldsymbol{s}_i)}{v_t}; \quad i = 1, \ldots, n,\ t = 1, \ldots, T$$

- Monitor excessively large values of $r_t^2(\boldsymbol{s}_i)$
  - Small (negative) $r_t(\boldsymbol{s}_i)$ corresponds to the AirBoxes that may be put indoors at $\boldsymbol{s}_l$
  - Large (positive) $r_t(\boldsymbol{s}_i)$ corresponds to potential emission sources of $PM_{2.5}$ around location $\boldsymbol{s}_i$ and time t

Detect Abnormal AirBox Measurements

- AirBoxes closer together in space are likely to produce similar measurements
- $\hat{y}_t(\boldsymbol{s})$: Predicted surface obtained from the proposed fixed rank kriging
- Detect outliers by monitoring standardized residuals

$$r_t(\boldsymbol{s}_i) = \frac{z_t(\boldsymbol{s}_i) - \hat{y}_t(\boldsymbol{s}_i)}{v_t}; \quad i = 1, \ldots, n,\ t = 1, \ldots, T$$

- Monitor excessively large values of $r_t^2(\boldsymbol{s}_i)$
  - Small (negative) $r_t(\boldsymbol{s}_i)$ corresponds to the AirBoxes that may be put indoors at $\boldsymbol{s}_l$
  - Large (positive) $r_t(\boldsymbol{s}_i)$ corresponds to potential emission sources of $PM_{2.5}$ around location $\boldsymbol{s}_i$ and time t

R Package

- `autoFRK`: To be submitted to CRAN