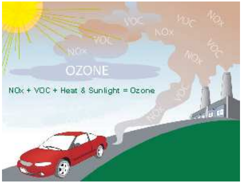
Bivariate Splines for Ozone Concentration Predictions

Bree Ettinger
University of Georgia

July 9, 2009

Ground Level Ozone



Ground-level or "bad" ozone is not emitted directly into the air, but is created by chemical reactions between oxides of nitrogen (NOx) and volatile organic compounds (VOC) in the presence of sunlight. Emissions from industrial facilities and electric utilities, motor vehicle exhaust, gasoline vapors, and chemical solvents are some of the major sources of NOx and VOC. (http://www.epa.gov/air/ozonepollution/basic.html)

Consider a real-valued time series observed p times a period over n periods sampled over a regional domain $\mathcal{D} \subset \mathbf{R}^2$. That is, we have

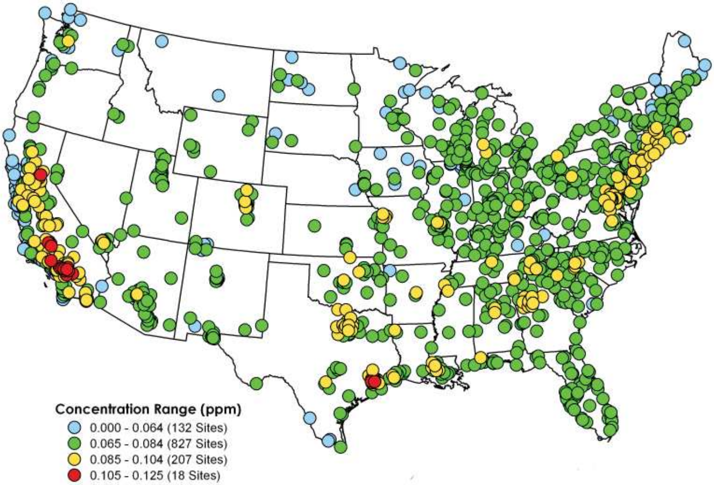
$$\{y_{ij}(x_k), k = 1, \cdots, N, j = 1, \cdots, p, i = 1, \cdots, n, x_k \in \mathcal{D}\}.$$

Example

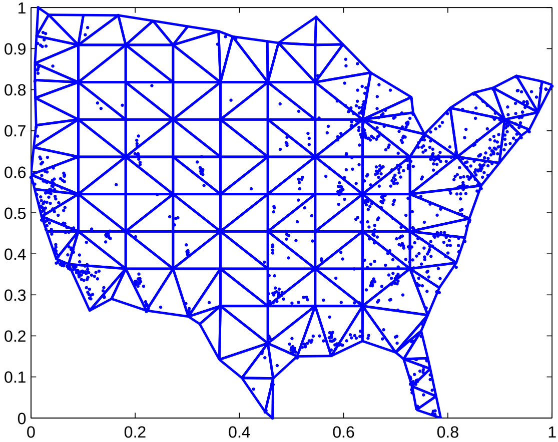
Consider the ozone concentrations over 969 locations around the U.S. whose measurements are taken every hour per day for 90 days. In this case, we have $n = 90$, $p = 24$ and $N = 969$ with $x_k \in \mathcal{D}$.

NOTE: We would like to predict the highest ozone concentration during a day at one city, say Atlanta.

Ground Level Ozone Data



Triangulation of the USA with EPA Locations



Challenges

- Current methods use time series at one location over a period of time to make a prediction. However, these methods cannot take into consideration neighboring ozone concentrations.
- The locations of the EPA stations are scattered around.
    - There are too few data points in some places and relatively too many in other places.
    - The size of data set is large and nonuniform.
    - The changes of the data values are volatile. Low level ozone has more drastic behavior than that of the temperature over the U.S..

Research Strategy

- Assume that the ozone concentration over $\mathcal{D}$ is a random surface, X, with known values at N locations.
- We consider the highest ozone concentration in one day for one city to be a functional, $f(X)$, of the ozone concentration surface, X, over $\mathcal{D}$ on the previous day.
- We use the bivariate spline space $S^r_d(\triangle)$ with smoothness $r > 0$ and degree $d > r$ over a triangulation, $\triangle$, of $\mathcal{D}$ to approximate X using the given measurement values at the N locations.

The Spline Space: $S_d^r(\Delta)$

Definition

Let $\triangle$ be a triangulation of a domain containing

$$\{(x_i, y_i), i = 1, \cdots, N\}.$$

We define the spline space of smoothness r and degree d over $\triangle$ by,

$$S_d^r(\triangle) = \{s \in C^r(\Omega), s|_t \in \mathbf{P}_d, t \in \triangle\}$$

the spline space of smoothness r and degree d over $\triangle$.

The Size of a Triangulation

Definition

For triangulation $\triangle$ the size of the triangulation, denoted $|\triangle|$, is the length of the longest edge in the $\triangle$

Spline Approximation Order

Theorem

Suppose that $d \geq 3r + 2$ and $\triangle$ be a triangulation. Then there exists a quasi-interpolatory operator $Qf \in S_d^r(\triangle)$ mapping any $f \in L_1(\mathcal{D})$ into $S_d^r(\triangle)$ such that Qf achieves the optimal approximation order: if $f \in W_p^{m+1}(\mathcal{D})$,

$$\|D_1^\alpha D_2^\beta (Qf - f)\|_{L_p(\mathcal{D})} \leq C|\triangle|^{m+1-\alpha-\beta}|f|_{m+1,p,\mathcal{D}}$$

for all $\alpha + \beta \leq m + 1$ with $0 \leq m \leq d$.

(See [Lai and Schumaker'07].)

Penalized Least Squares Method:

Let $\{(x_i, y_i, f(x_i, y_i)),\ i = 1, \ldots, N\}$ be a scattered data set where *N is a relatively large integer*. Then the penalized least squares method is to find $s_f \in \mathcal{S}$ such that for a positive weight $\rho > 0$

$$P_\rho(s_f) := \min_{s \in \mathcal{S}} P_\rho(s)$$

where

$$P_\rho(s) := \sum_{i=1}^{N} |s(x_i, y_i) - f(x_i, y_i)|^2 + \rho E(s)$$

and

$$E(s) := \sum_{T \in \triangle} \int_T \left(s_{xx}^2 + 2s_{xy}^2 + s_{yy}^2\right) dxdy.$$

Approximation of Functional

Suppose that $X(\omega) \in L_2(\mathcal{D})$ is a random L_2 function for each event $\omega \in \Omega$.

Assume $f(X)$ is a bounded and continuous functional of X. Then by the Riesz Representation theorem we know that

$$f(X) = \langle g, X \rangle$$

for some function g.

That is, we approximate the functional f by a linear functional:

$$Y = \langle \alpha, X \rangle + \epsilon,$$

where $E(\epsilon) = 0$.

Autoregressive Process

Let α be the solution of the following:

$$\alpha = \arg \min_{\beta \in L_2(\mathcal{D})} E\left[\left(f(X) - \epsilon - \langle \beta, X \rangle \right)^2 \right]. \quad (1)$$

To solve the above problem, we want to write the equation in terms of the Covariance and Cross Covariance functions.

Covariance:

$$\Gamma x(s) = \int_{\mathcal{D}} \mathcal{E}(X(s)X(t))x(t)dt, \quad \forall x \in H.$$

Cross Covariance:

$$\Delta x = \int_{\mathcal{D}} \mathcal{E}(X(s)Y)x(s)ds, \quad \forall x \in H.$$

Autoregressive Process

Theorem

The solution to the minimization problem (1) is

$$\Gamma g = \Delta. \tag{2}$$

Autoregressive Process

Proof.

We start with the minimization problem (1). If g is the solution, the following function

$$F(r) = \mathcal{E}\left[(Y - \langle g + rf, X\rangle)^2\right] \tag{3}$$

achieves the minimum when $r = 0$ for any function $f \in H$. By taking the derivative with respect to r, we have

$$\begin{aligned} F'(r) &= 2\mathcal{E}\left[(Y - \langle g + rf, X\rangle)(-\langle f, X\rangle)\right] \\ &= \mathcal{E}\left[(Y - \langle g, X\rangle)(\langle f, X\rangle)\right] \\ &= 0. \end{aligned}$$

Evaluating at $r = 0$ yields

$$\mathcal{E}\left[(\langle g, X\rangle)(\langle f, X\rangle)\right] \;=\; \mathcal{E}\left[Y(\langle f, X\rangle)\right]. \tag{4}$$

Since the equation (4) holds for all $f \in H$, we have

$$\int_{t\in\mathcal{D}}\int_{s\in\mathcal{D}} g(s)\mathcal{E}[X(s)X(t)]f(t)dsdt = \int_{t\in\mathcal{D}} f(t)\mathcal{E}[X(t)Y]dt. \tag{5}$$

Now, let Γ be the standard covariance operator of the H-valued random variables X and take Δ to be the cross-covariance of X and Y. Then we have

$$(\Gamma g)(t) = \int_{s\in\mathcal{D}} \mathcal{E}(X(s)X(t))g(s)ds, \quad \forall g \in H.$$

$$\langle \Delta, f\rangle = \int_{t\in\mathcal{D}} \mathcal{E}(X(t)Y)f(t)dt \quad \forall f \in H. \tag{6}$$

Now, (5) can be denoted by

$$\Gamma g = \Delta.$$

Covariance

Let λ_j for $j = 1, 2, \ldots$ be the eigenvalues of Γ in decreasing order and let the $v_j \in H$ be the eigenfunctions of Γ associated with the corresponding λ_j. If the v_j form an orthonormal basis for H then we can express

$$\Gamma = \sum_j \lambda_j v_j$$

and for any $g \in H$ we have

$$g = \sum_j \langle g, v_j \rangle v_j.$$

Since Γ is a symmetric operator, we use (5) to have

$$\lambda_j \langle g, v_j \rangle = \langle g, \lambda_j v_j \rangle = \langle g, \Gamma v_j \rangle = \langle \Gamma g, v_j \rangle = \langle \Delta, v_j \rangle = \langle \mathcal{E}(X(t)Y), v_j \rangle,$$

If $\lambda_j > 0$ then

$$\langle g, v_j \rangle = \langle \mathcal{E}(X(\cdot)Y), v_j \rangle / \lambda_j.$$

Thus, we get the expansion for g:

$$g = \sum_{j=1}^{\infty} \frac{\langle \mathcal{E}(X(\cdot)Y), v_j \rangle}{\lambda_j} v_j.$$

Note that the function g is in H if and only if

$$\sum_{j=1}^{\infty} \left(\frac{\langle \mathcal{E}(X(\cdot)Y), v_j \rangle}{\lambda_j} \right)^2 < +\infty.$$

In general, we do not know if Γ is invertible or not.

- Let $\mathcal{N}(\Gamma) = \{x \in H, \Gamma x = 0\}$ and suppose that $\mathcal{N}(\Gamma) \neq \emptyset$. Then g can not be uniquely determined.
- However, g can be determined in $\mathcal{N}(\Gamma)^{\perp}$.
- Let $H_k = \text{span}\{v_1, \cdots, v_k\} \subset \mathcal{N}(\Gamma)^{\perp}$ and $\mathcal{P}_k$ be the orthogonal projection operator from H to H_k.

When $\lambda_k > 0$, $\mathcal{P}_k\Gamma\mathcal{P}_k$ is invertible.

$$\mathcal{P}_k\Gamma\mathcal{P}_k g = \sum_{j=1}^{k} \lambda_j \langle v_j, g\rangle v_j.$$

Thus, for all $f \in H$,

$$P_k f = \sum_{j=1}^{k} \langle f, v_j\rangle v_j$$

and (5) yields

$$\langle \mathcal{P}_k\Gamma\mathcal{P}_k g, P_k f\rangle = \langle \Delta, P_k f\rangle.$$

$$\sum_{j=1}^{k} \lambda_j \langle v_j, g \rangle \langle v_j, f \rangle = \sum_{j=1}^{k} \langle f, v_j \rangle \langle \Delta, v_j \rangle$$

for all $f \in H$ and it follows that

$$\langle v_j, g \rangle = \frac{1}{\lambda_j} \langle \Delta, v_j \rangle$$

for $j = 1, \ldots, k$. Hence, we have

$$g_k = \sum_{j=1}^{k} \frac{1}{\lambda_j} \langle \Delta, v_j \rangle v_j$$

is the approximation of g in H_k.

Empirical Estimators of Γ and Δ:

For random samples X_i, $i = 1, \ldots, n$ in H and Y_i another random sample dependant on X_i we have

$$\Gamma_n = \frac{1}{n} \sum_{i=1}^{n} \langle X_i, x \rangle X_i$$

and

$$\Delta_n = \frac{1}{n} \sum_{i=1}^{n} \langle X_i, x \rangle Y_i.$$

Γ_n is a compact operator mapping H to H and thus can be expanded in terms of its eigenfunctions $\hat{v}_j$ for $j = 1, 2, \ldots,$

$$\Gamma_n = \sum_{i=1}^{\infty} \hat{\lambda}_j \langle \hat{v}_j, x \rangle \hat{v}_j$$

and our model in the empirical case is given by

$$\Delta_n x = \langle g_n, \Gamma_n x \rangle.$$

Principal Component Regression Estimator of g_n:

$$\hat{g}_{PCR} = \sum_{j=1}^{k_n} \frac{\Delta_n(\hat{v}_j)}{\hat{\lambda}_j} \hat{v}_j.$$

The Smooth Principal Component Regression Estimator of g_n denoted $\hat{g}_{SPCR}$ the solution to:

$$\hat{g}_{SPCR} = \min_{f \in S_d^r(\triangle)} \int |\hat{g}_{PCR}(s) - f(s)|^2 ds.$$

Discrete Observations

In practice, we do not observe the continuous random surface X_i but we only observe the random surface at design points $s_k \in \mathcal{D},\ k = 1, \ldots, N$:

$$\{z_{i,k},\ k = 1, \ldots, N\}.$$

So we approximate X_i by taking the penalized least squares fit, S_{X_i}.

$$\widetilde{\Gamma_n}(x) = \frac{1}{n}\sum_{i=1}^{n}\langle S_{X_i}, x\rangle S_{X_i} = \sum_{j=1}^{m}\widetilde{\lambda}_j\langle\widetilde{v}_j, x\rangle\widetilde{v}_j,$$

where $\widetilde{\lambda}_j$ and $\widetilde{v}_j$ are a pair of eigenvalue and eigenvector of $\widetilde{\Gamma_n}$.

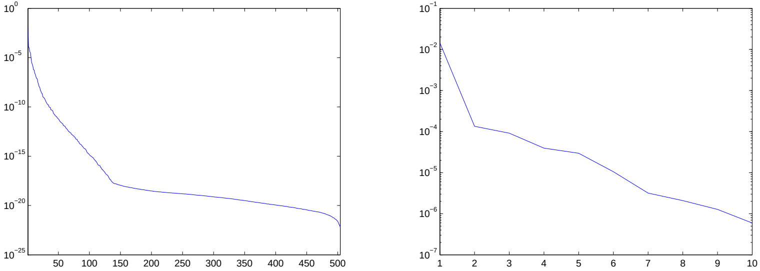
$$\widetilde{\Delta_n}(x) = \frac{1}{n}\sum_{i=1}^{n}\langle S_{X_i}, x\rangle Y_i,$$

Similarly, $\widetilde{\Delta_n}(x) = \langle g_n, \widetilde{\Gamma_n} x\rangle$ for some $g_n \in H$.

Assume that the first k_n largest eigenvalues $\widetilde{\lambda}_j,\ j = 1, \cdots, k_n$ are nonzero. Then the principal component regression estimator of g_n is

$$\widetilde{g}_{PCR} = \sum_{j=1}^{k_n} \frac{\Delta_n(\widetilde{v}_j)}{\widetilde{\lambda}_j} \widetilde{v}_j. \tag{7}$$

The Art of Choosing the Number of Eigenvalues



Brute Force Overview

- $\alpha = \arg\min_{\beta \in H} E\left[(f(X) - \epsilon - \langle \beta, X \rangle)^2\right]$
- $S_\alpha = \arg\min_{\beta \in S_d^r(\triangle)} E\left[(f(X) - \epsilon - \langle \beta, X \rangle)^2\right]$
- $\widehat{S_{\alpha,n}} = \arg\min_{\beta \in S_d^r(\triangle)} \frac{1}{n} \sum_{i=1}^{n} (f(X_i) - \epsilon_i - \langle \beta, X_i \rangle)^2$
- $\alpha_D = \arg\min_{\beta \in H} E\left[(f(X) - \epsilon - \langle \beta, S_X \rangle)^2\right]$
- $S_{\alpha_D} = \arg\min_{\beta \in S_d^r(\triangle)} E\left[(f(X) - \epsilon - \langle \beta, S_X \rangle)^2\right]$
- $\widetilde{S_{\alpha_D,n}} = \arg\min_{\beta \in S_d^r(\triangle)} \frac{1}{n} \sum_{i=1}^{n} (f(X_i) - \epsilon_i - \langle \beta, S_{X_i} \rangle)^2$

We want to find a solution solving the following minimization problem (1):

$$\alpha = \arg\min_{\beta \in H} E\left[(f(X) - \epsilon - \langle \beta, X \rangle)^2\right].$$

Since bivariate spline space $S_d^r(\triangle)$ can be dense in a Hilbert space, H, as $|\triangle| \to 0$, we look for $S_\alpha \in S_d^r(\triangle)$ of α such that

$$S_\alpha = \arg\min_{\beta \in S_d^r(\triangle)} E\left[(f(X) - \epsilon - \langle \beta, X \rangle)^2\right]. \tag{8}$$

Theorem

Suppose that only the zero spline in $S_d^r(\triangle)$ is orthogonal to the collection $\mathcal{X} = \{X(s),\ s \in \mathcal{D}\} \subset H$. Then the minimization problem (8) has a unique solution in $S_d^r(\triangle)$.

We rewrite this as a linear system $A\mathbf{c} = \mathbf{b}$, where

$$A = E(\langle \phi_i, X\rangle\langle \phi_j, X\rangle) \text{ for } i, j = 1, \ldots, m$$

$$\mathbf{b} = E((f(X) + \epsilon)\langle \phi_j, X\rangle) = E(f(X)\langle \phi_j, X\rangle) \text{ for } j = 1, \ldots, m$$

The coefficient vector of S_α is the vector $\mathbf{c} = (c_1, \ldots, c_m)^T$ which satisfies $A\mathbf{c} = \mathbf{b}$.

Theorem

Suppose that $E(\|X\|^2) < \infty$ *and suppose* $\alpha \in C^r(\mathcal{D})$ *for* $r \geq 1$*. Then the spline solution* S_α *approximates* α *in the following sense:*

$$E((\langle \alpha - S_\alpha, X \rangle)^2) \leq C|\triangle|^{2r} E(\|X\|^2),$$

where $|\triangle|$ *is the maximal length of the edges of* $\triangle$*.*

Empirical Estimate

Given a sample of independent and identically distributed (i.i.d) random surfaces $X_i, \ i = 1, \ldots, n$. The empirical estimate $\widehat{S_{\alpha,n}} \in S_d^r(\triangle)$ is the solution of

$$\widehat{S_{\alpha,n}} = \arg \min_{\beta \in S_d^r(\triangle)} \frac{1}{n} \sum_{i=1}^{n} (f(X_i) + \epsilon_i - \langle \beta, X_i \rangle)^2. \tag{9}$$

Empirical Estimate

Theorem

Suppose that only the zero spline function in the spline space $S_d^r(\triangle)$ is perpendicular to the subspace span$\{X_1, \ldots, X_n\}$ except on an event whose probability p_n goes to zero as $n \longrightarrow +\infty$. Then, with probability $1 - p_n$, there exists a unique $\widehat{S_{\alpha,n}} \in S_d^r(\triangle)$ minimizing (9).

Write $\widehat{S_{\alpha,n}} = \sum_{i=1}^{m} c_{n,i}\phi_i$. and rewrite the minimization problem as $\widehat{A}_n \mathbf{c}_n = \widehat{b}_n$ where

$$\widehat{A}_n = \frac{1}{n}\sum_{\ell=1}^{n}\langle \phi_i, X_\ell\rangle\langle \phi_j, X_\ell\rangle \text{ for } i,j = 1,\ldots,m$$

$$\widehat{b}_n = \frac{1}{n}\sum_{\ell=1}^{n} f(X_\ell)\langle \phi_j, X_\ell\rangle + \frac{1}{n}\sum_{\ell=1}^{n}\langle \phi_j, \epsilon_j X_\ell\rangle \text{ for } j = 1,\ldots,m.$$

The coefficient vector of $\widehat{S_{\alpha,n}}$ is the vector $\mathbf{c}_n = (c_{n,i}, i = 1,\ldots,m)$ that satisfies $\widehat{A}_n \mathbf{c}_n = \widehat{b}_n$.

Convergence

Theorem

Suppose that X_ℓ, $\ell = 1, \cdots, n$ are i.i.d. and $\|X_\ell\|$ are bounded a.s. Suppose ϵ_ℓ are also bounded almost surely and $f(X)$ is a bounded linear functional. Then $\widehat{S_{\alpha,n}}$ converges to S_α in probability with convergence rate

$$P\left(\frac{\|S_\alpha - \widetilde{S_{\alpha,n}}\|}{\|S_\alpha\|} \geq \delta\right) \leq 4m^2 \exp\left(-\frac{n\gamma^2\delta^2}{32\kappa(A)^2 m^2 M^2}\right)$$
$$+2m \exp\left(-\frac{n\gamma^2\delta^2}{32\kappa(A)^2 M_b^2}\right) + 2m \exp\left(-\frac{n\gamma^2\delta^2}{32\kappa(A)^2 M_\epsilon^2}\right).$$

Discrete Observations

We are given observations of X over some designed points $s_k, k = 1, \dots, N$ over $\mathcal{D}$.

Let S_X be the discrete or penalized least square fit of X.

We consider α_D that solves the minimization problem:

$$\alpha_D = \arg\min_{\beta \in H} E\left[(f(X) - \epsilon - \langle \beta, S_X \rangle)^2\right]. \tag{10}$$

Also we look for an approximation $S_{\alpha_D} \in S_d^r(\triangle)$ of α_D such that

$$S_{\alpha_D} = \arg\min_{\beta \in S_d^r(\triangle)} E\left[(f(X) - \epsilon - \langle \beta, S_X \rangle)^2\right]. \tag{11}$$

α_D approximates α

The function

$$F(\beta) = E\left[(f(X) - \epsilon - \langle \beta, X \rangle)^2\right]$$

is strictly convex as well as

$$F_D(\beta) = E\left[(f(X) - \epsilon - \langle \beta, S_X \rangle)^2\right].$$

Thus, the minimizer is unique for each of these functionals.

S_X approximates X as $|\triangle| \to 0$.

Thus, $F_D(\beta)$ approximates $F(\beta)$ for each β since both $F(\beta)$ and $F_D(\beta)$ are continuous.

It follows that α_D approximates α.

Write $S_{\alpha_D} = \sum_{j=1}^{m} c_{D,j}\phi_j$ and express the minimization problem as $A_D\mathbf{c}_D = \mathbf{b}_D$ where

$$A_D = E(\langle \phi_i, S_X\rangle\langle \phi_j, S_X\rangle) \text{ for } i, j = 1, \ldots, m$$

$$\mathbf{b}_D = E((f(X) - \epsilon)\langle \phi_j, S_X\rangle) \text{ for } j = 1, \ldots, m$$

The coefficient vector of S_{α_D} is the vector $\mathbf{c}_D = (c_{D,1}, \ldots, c_{D,m})^T$ satisfies $A_D\mathbf{c}_D = \mathbf{b}_D$.

S_{α_D} approximates α_D

Theorem

Suppose that $E(\|X\|^2) < \infty$ and suppose $\alpha \in C^r(\mathcal{D})$ for $r \geq 1$. Then the solution S_{α_D} from the minimization problem (11) approximates α_D in the following sense:

$$E((\langle \alpha_D - S_{\alpha_D}, S_X \rangle)^2) \leq C|\triangle|^{2r}$$

for a constant C dependent on $E(\|X\|^2)$.

Empirical Estimate

Next we consider the empirical estimate of S_{α_D} based on discrete observations of random surfaces $X_i, i = 1, \ldots, n$.

The empirical estimate $\widetilde{S_{\alpha_D,n}} \in S_d^r(\triangle)$ is the solution of

$$\widetilde{S_{\alpha_D,n}} = \arg \min_{\beta \in S_d^r(\triangle)} \frac{1}{n} \sum_{i=1}^{n} (f(X_i) - \epsilon_i - \langle \beta, S_{X_i} \rangle)^2. \qquad (12)$$

Write $\widetilde{S_{\alpha_D,n}} = \sum_{i=1}^{m} \widetilde{c_{n,i}}\phi_i$ and rewrite the minimization problem as $\widetilde{A_n}\widetilde{\mathbf{c}_n} = \widetilde{b_n}$, where

$$\widetilde{A_n} = \frac{1}{n}\sum_{\ell=1}^{n}\langle \phi_i, S_{X_\ell}\rangle\langle \phi_j, S_{X_\ell}\rangle \text{ for } i,j = 1,\ldots,m$$

$$\widetilde{b_n} = \frac{1}{n}\sum_{\ell=1}^{n} f(X_\ell)\langle \phi_j, S_{X_\ell}\rangle + \frac{1}{n}\sum_{\ell=1}^{n}\langle \phi_j, \epsilon_\ell S_{X_\ell}\rangle \text{ for } j = 1,\ldots,m.$$

The coefficient vector of $\widetilde{S_{\alpha_D,n}}$ is the vector $\widetilde{\mathbf{c}_n} = (\widetilde{c_{n,i}}, i = 1,\ldots,m)$ that satisfies $\widetilde{A_n}\widetilde{\mathbf{c}_n} = \widetilde{b_n}$.

Computational Method

- For a time series over a regional domain we use a bivariate spline to approximate a surface over a bounded region. We call the resulting spline S_X it is an approximation of a functional random variable X.
- We also collect the desired quantity at location of interest. This yields another real random variable $Y = f(X)$.
- Next we compute:

$$\hat{A} = E(\langle \phi_i, S_X \rangle \langle \phi_i, S_X \rangle) = \frac{1}{n} \sum_{l=1}^{n} \langle \phi_i, S_{X_l} \rangle \langle \phi_j, S_{X_l} \rangle$$

$$\hat{b} = E\left((f(X) - \epsilon) \langle \phi_i, S_X \rangle \right) = \frac{1}{n} \sum_{l=1}^{n} f(X_l) \langle \phi_j, S_{X_l} \rangle$$

Computational Method

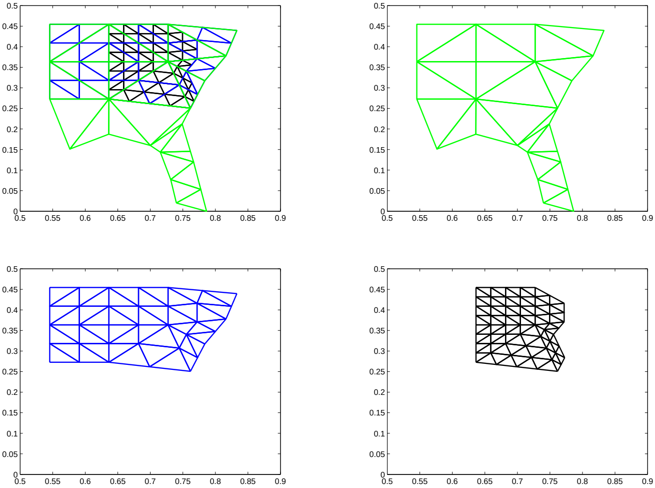
Now we would like to solve $\hat{A}c = \hat{b}$ for c which will yield a coefficient vector for a spline S_α such that

$$f(X) \approx \langle S_\alpha, S_X \rangle.$$

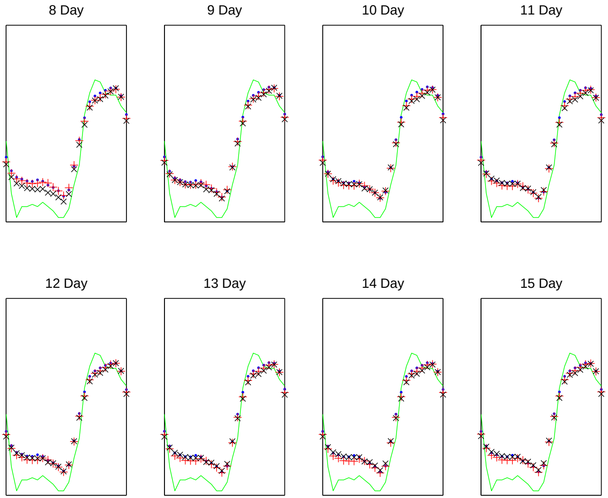
Then we implement the two options we have explored for finding the coefficients of S_α:

- Solve by Principle Component Analysis: $\widetilde{g}_{PCR}$
- Solve by Brute Force: $\widetilde{S_{\alpha_D,n}}$

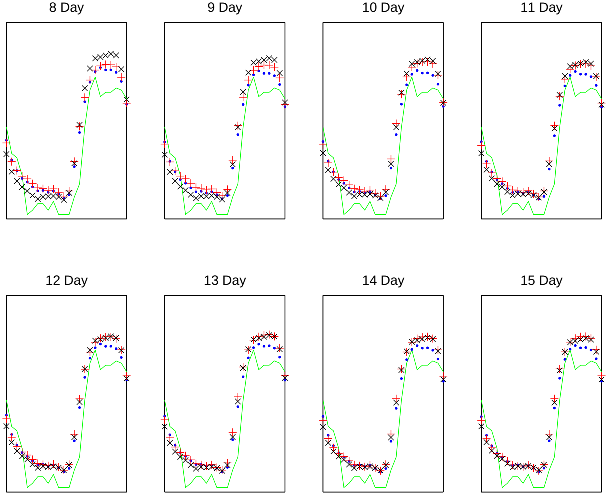
Southeast Triangulations



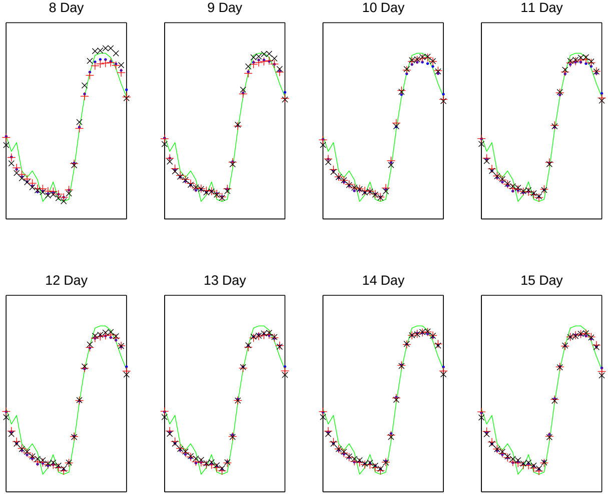
Autoregressive: Sept 1 - Atlanta



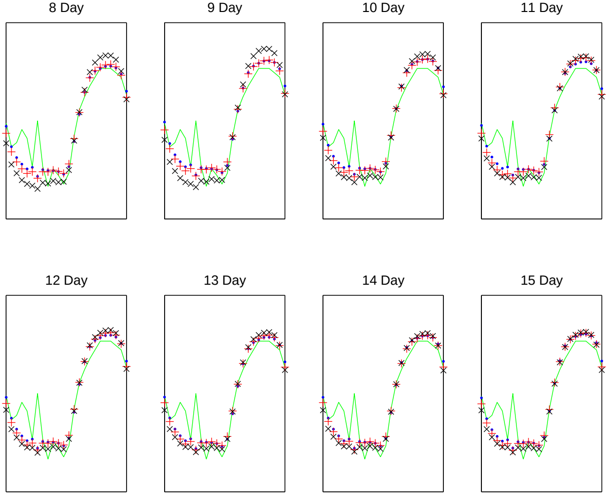
Autoregressive: Sept 2 - Atlanta



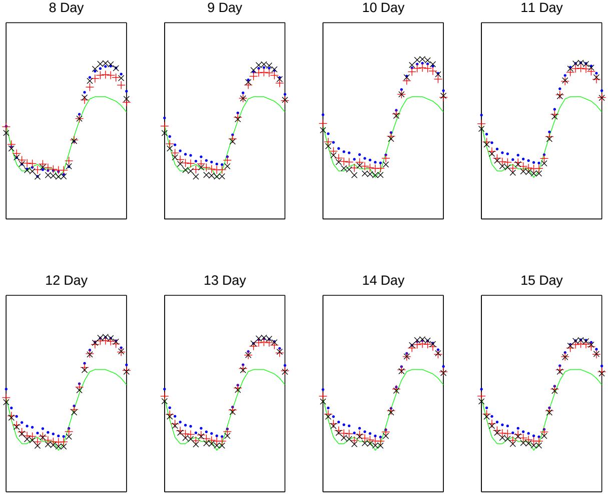
Autoregressive: Sept 3 - Atlanta



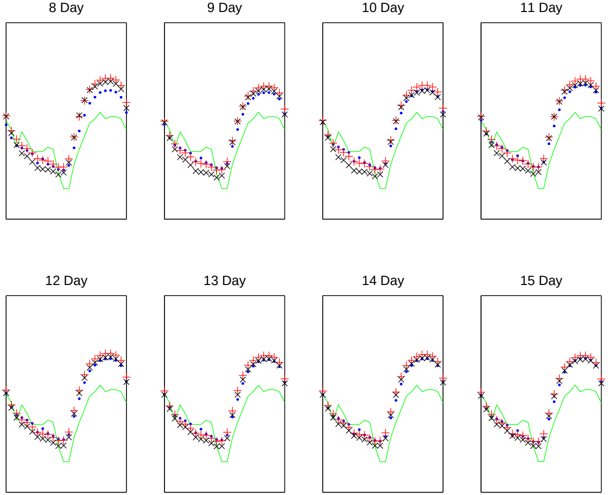
Autoregressive: Sept 4 - Atlanta



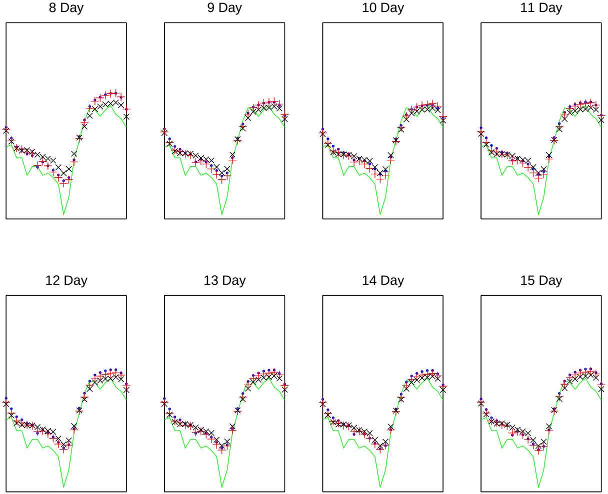
Autoregressive: Sept 5 - Atlanta



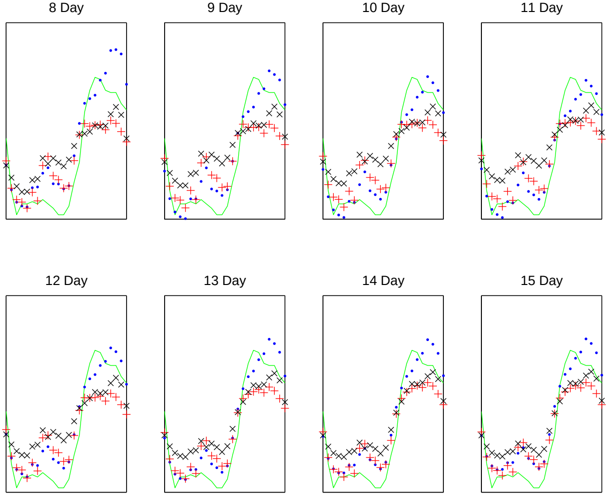
Autoregressive: Sept 6 - Atlanta



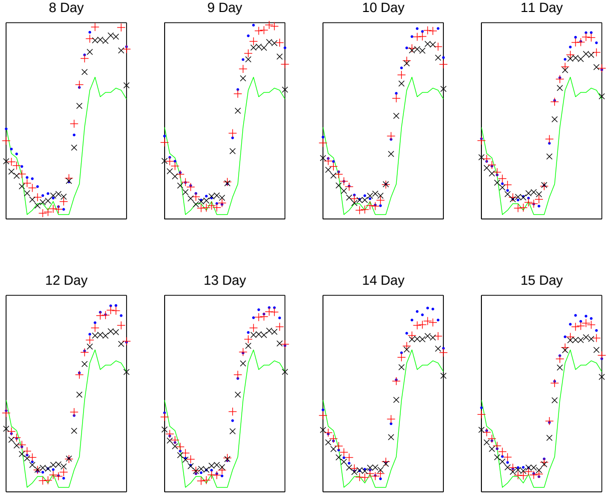
Autoregressive: Sept 7 - Atlanta



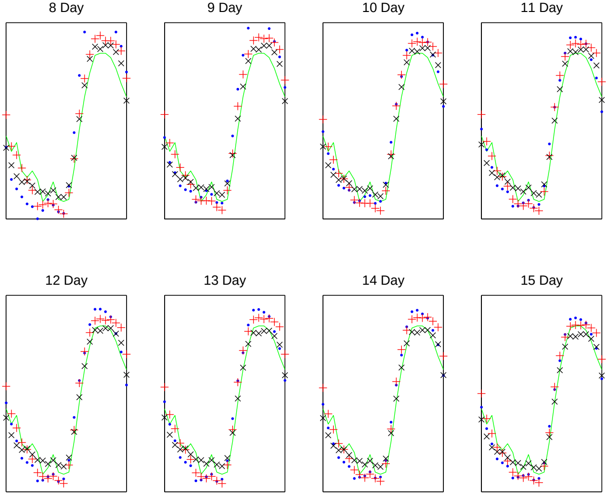
Brute Force: Sept 1 - Atlanta



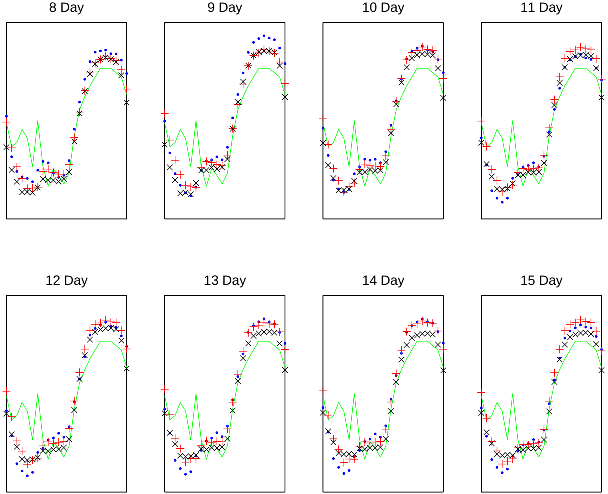
Brute Force: Sept 2 - Atlanta



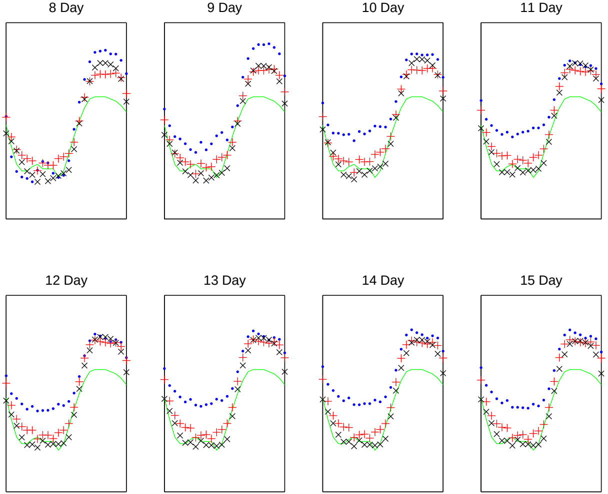
Brute Force: Sept 3 - Atlanta



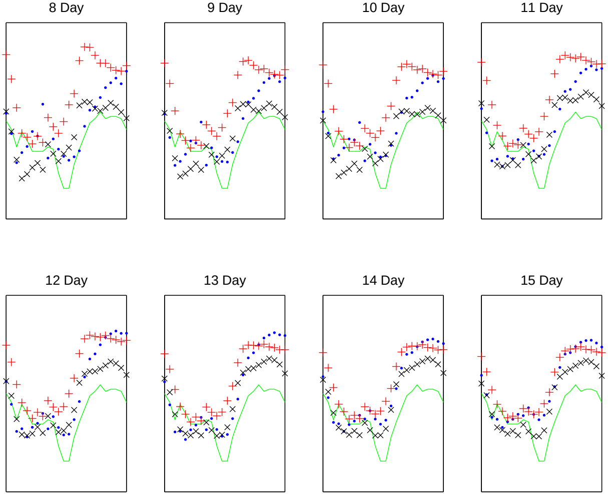
Brute Force: Sept 4 - Atlanta



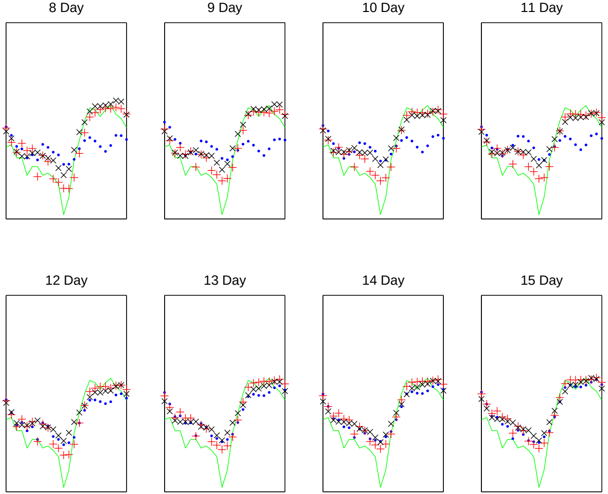
Brute Force: Sept 5 - Atlanta



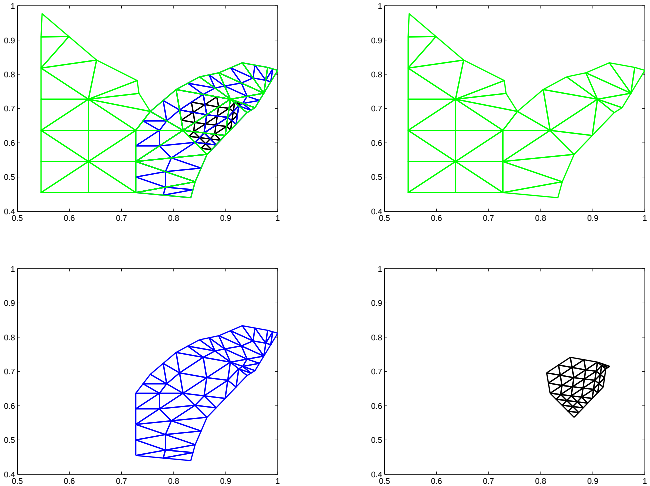
Brute Force: Sept 6 - Atlanta



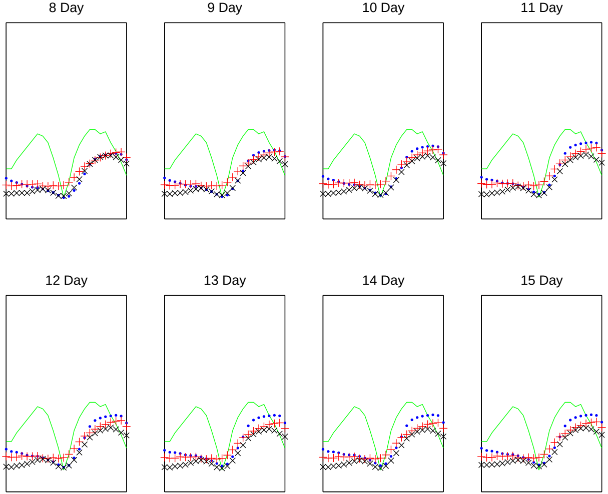
Brute Force: Sept 7 - Atlanta



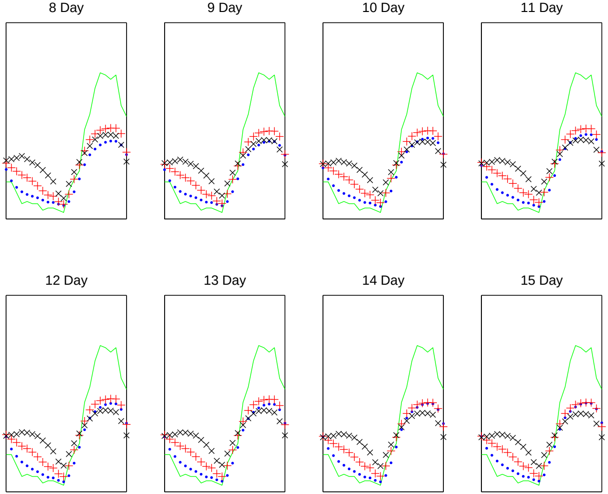
Northeast Triangulations



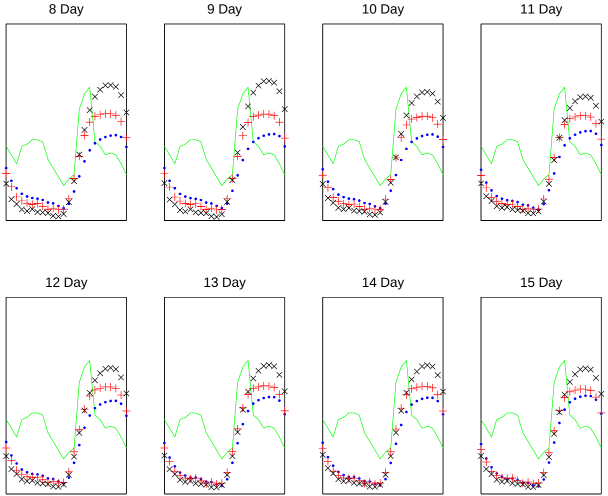
Autoregressive: Sept 1 - Boston



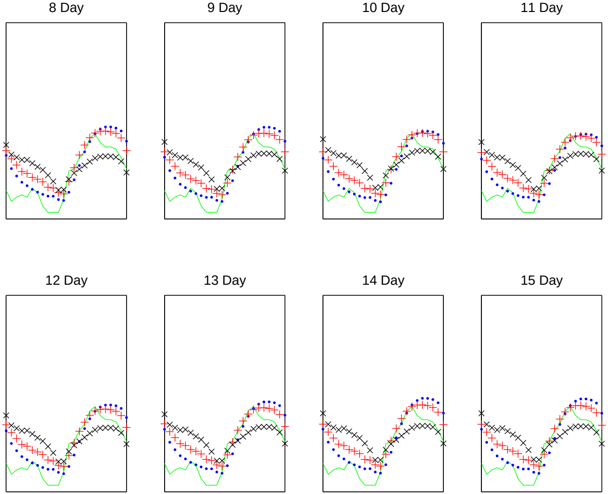
Autoregressive: Sept 2 - Boston



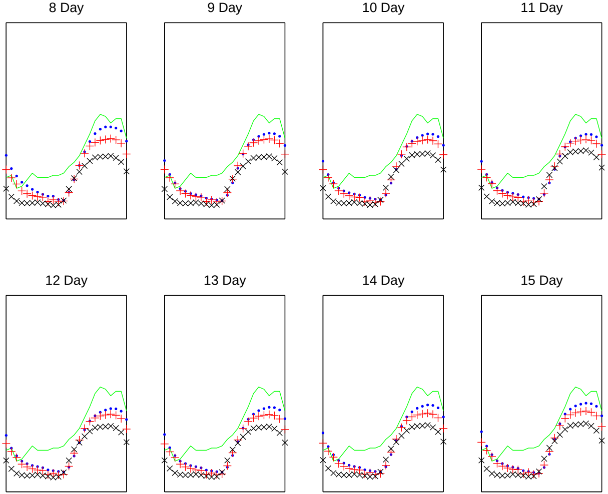
Autoregressive: Sept 3 - Boston



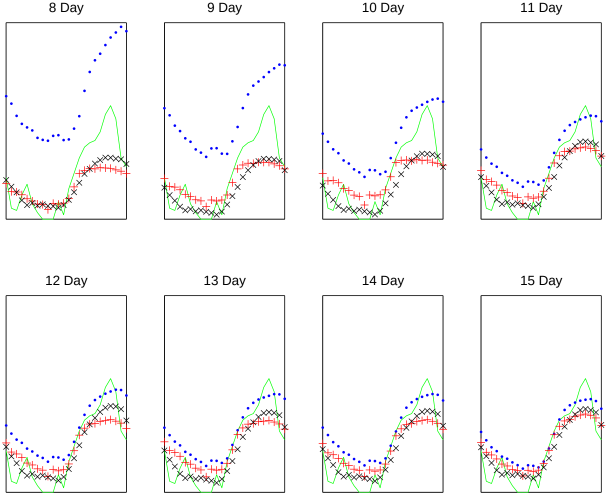
Autoregressive: Sept 4 - Boston



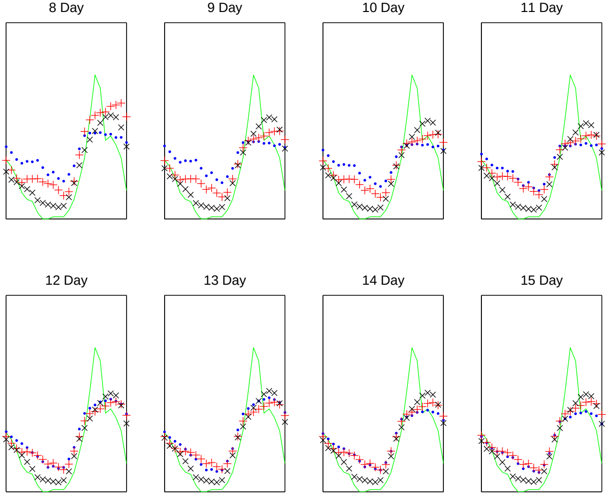
Autoregressive: Sept 5 - Boston



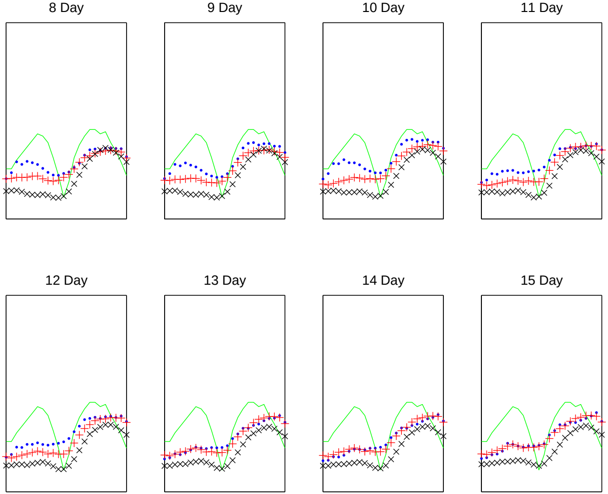
Autoregressive: Sept 6 - Boston



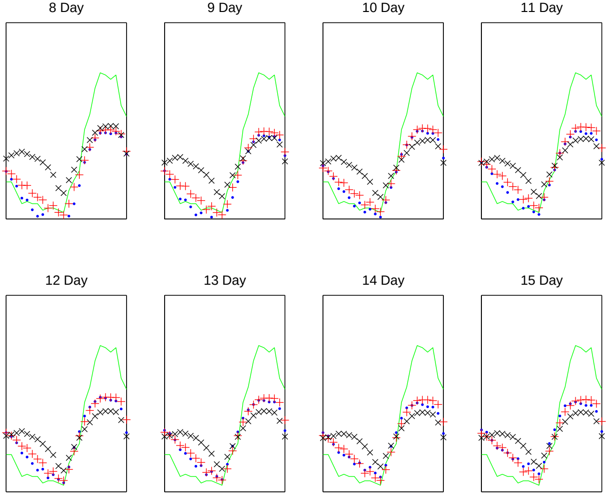
Autoregressive: Sept 7 - Boston



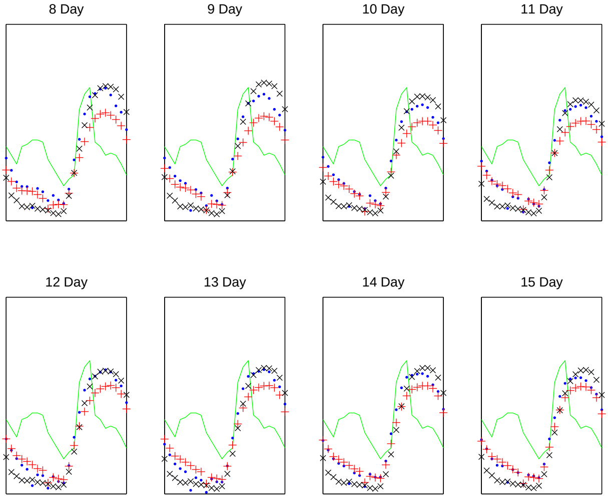
Brute Force: Sept 1 - Boston



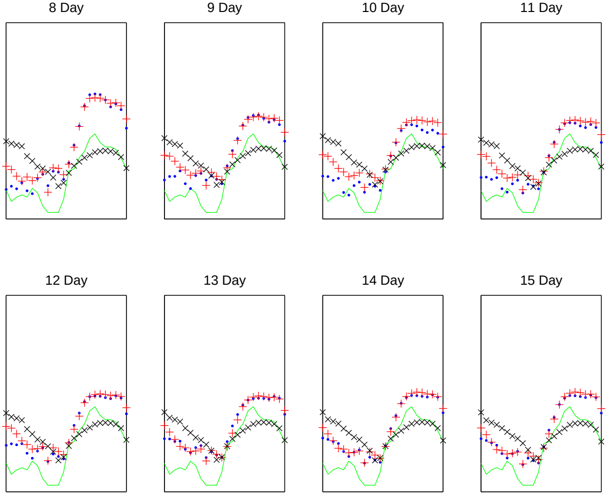
Brute Force: Sept 2 - Boston



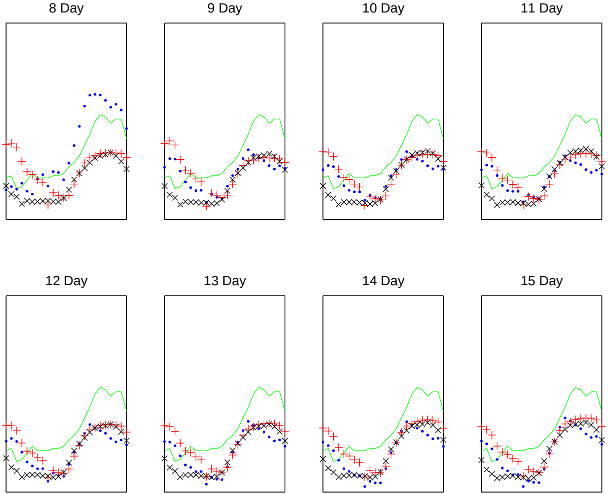
Brute Force: Sept 3 - Boston



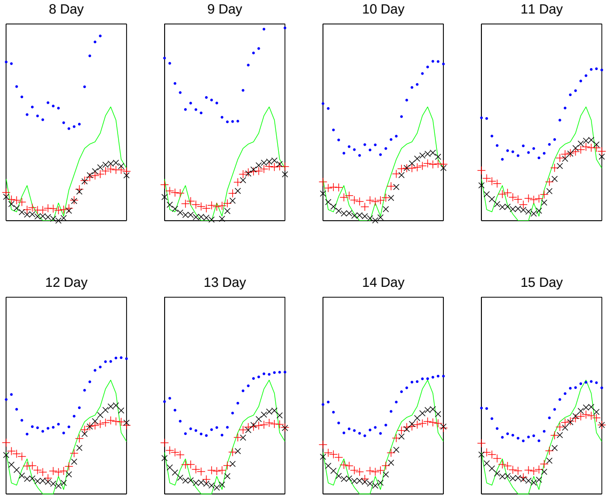
Brute Force: Sept 4 - Boston



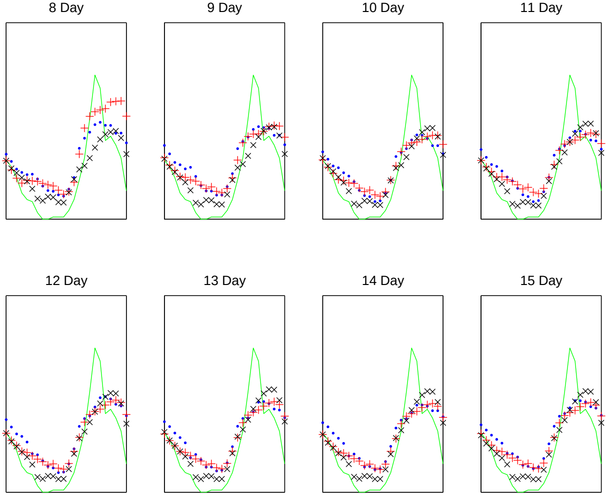
Brute Force: Sept 5 - Boston



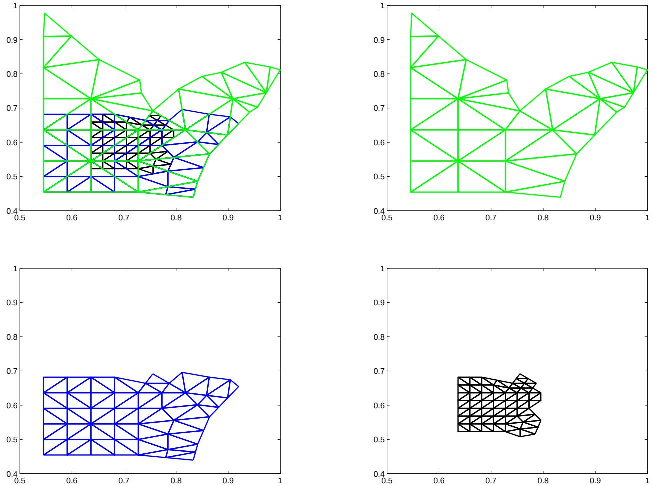
Brute Force: Sept 6 - Boston



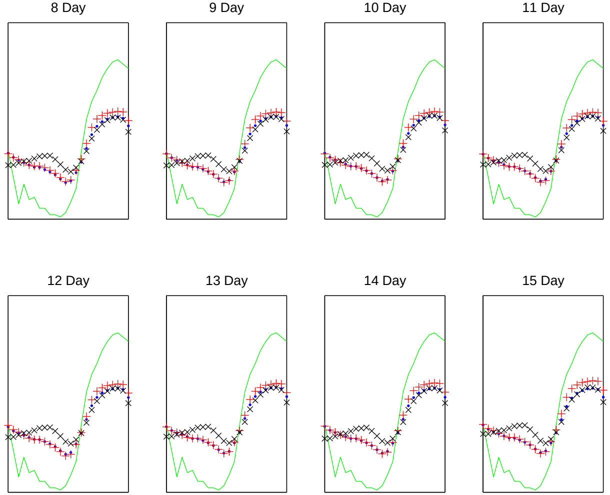
Brute Force: Sept 7 - Boston



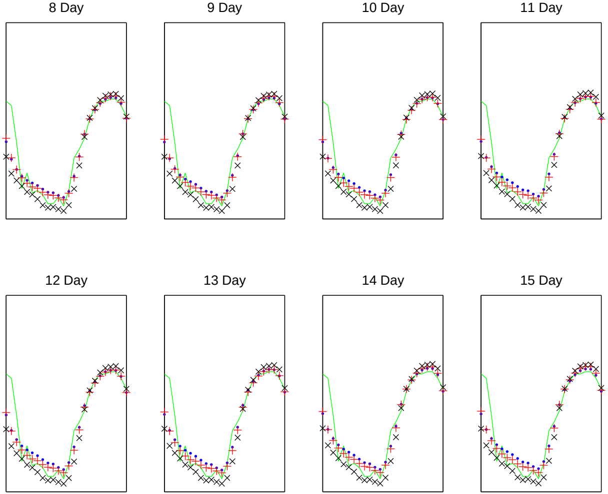
Northeast Triangulations



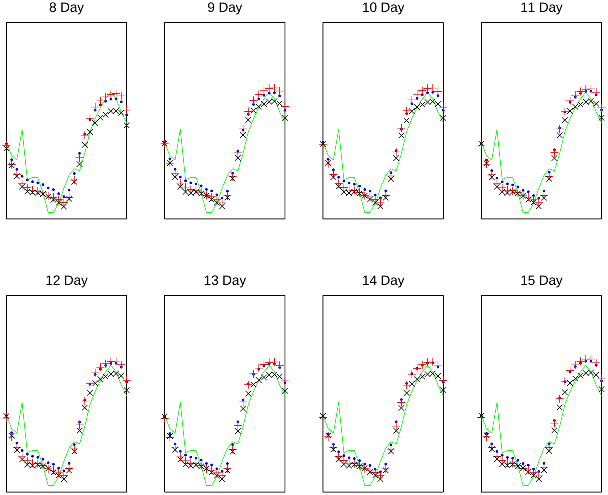
Autoregressive: Sept 1 - Cincinnati



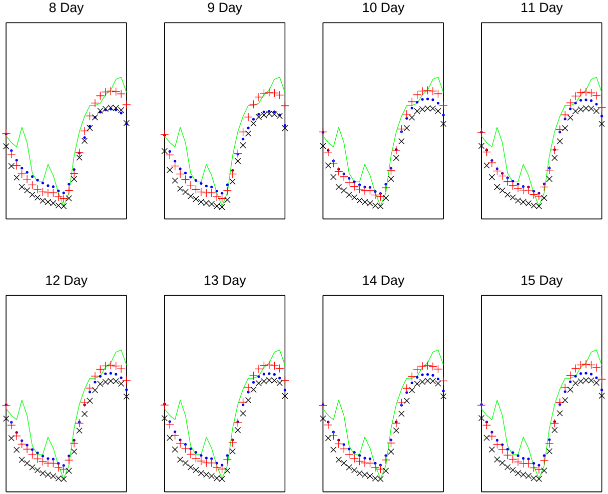
Autoregressive: Sept 2 - Cincinnati



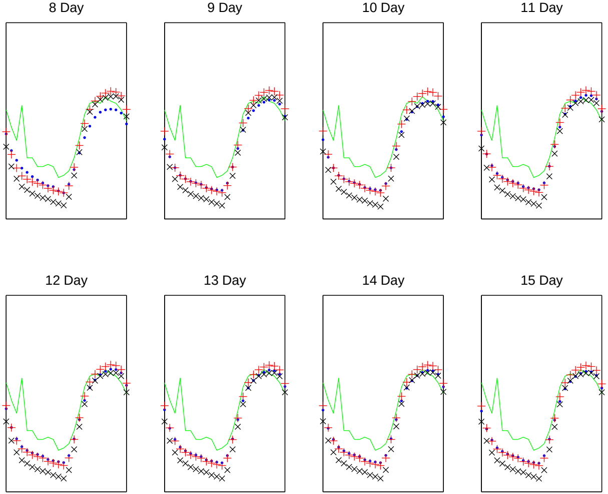
Autoregressive: Sept 3 - Cincinnati



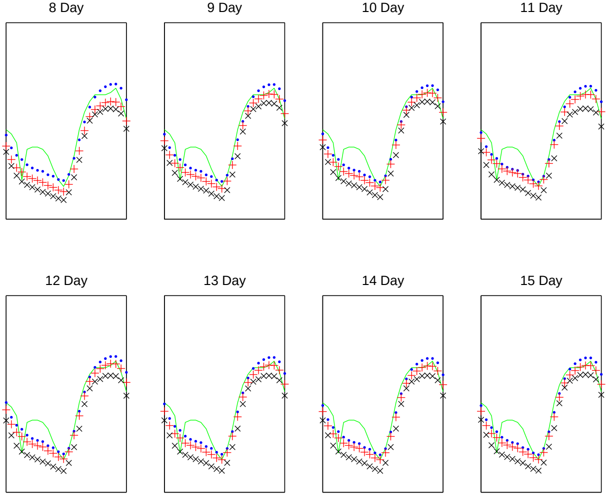
Autoregressive: Sept 4 - Cincinnati



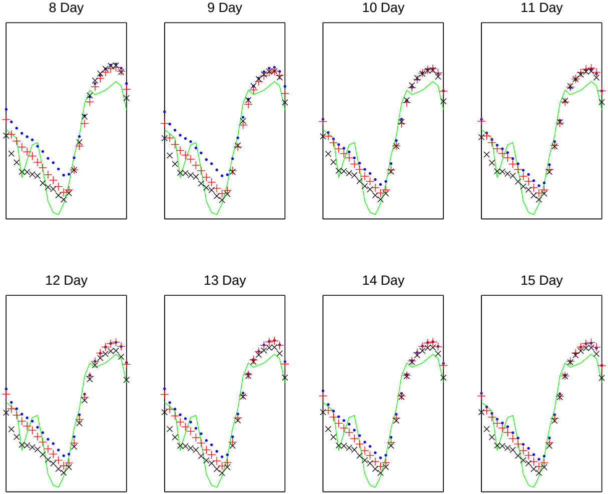
Autoregressive: Sept 5 - Cincinnati



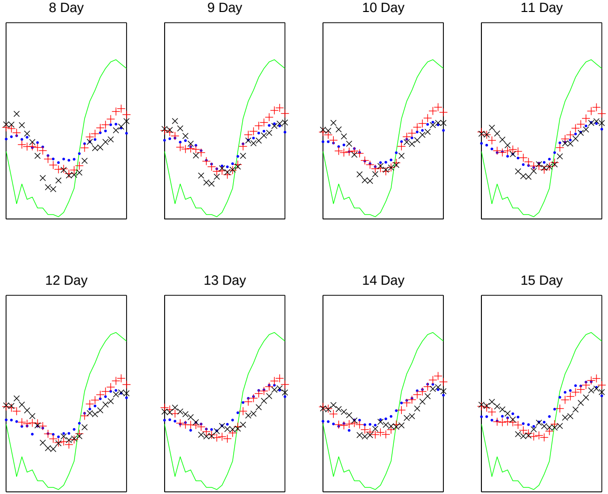
Autoregressive: Sept 6 - Cincinnati



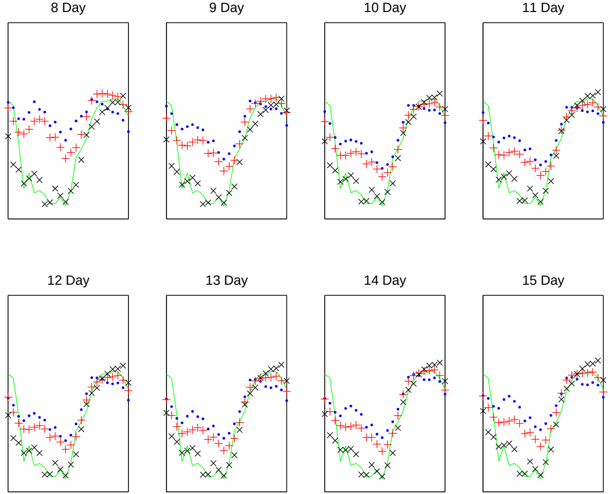
Autoregressive: Sept 7 - Cincinnati



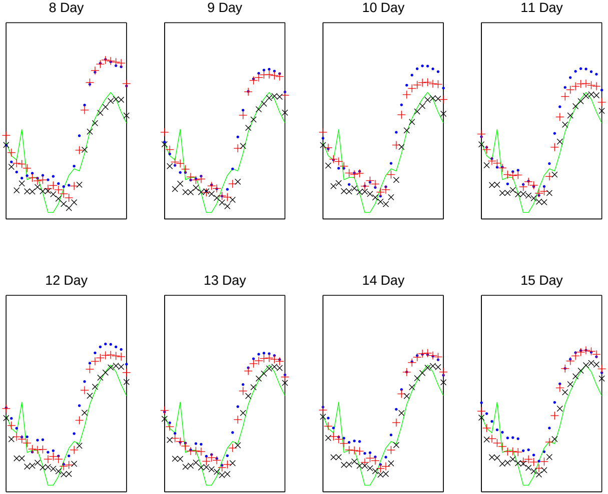
Brute Force: Sept 1 - Cincinnati



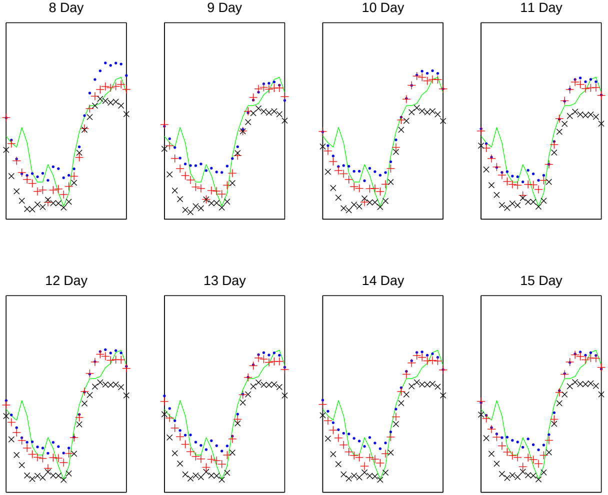
Brute Force: Sept 2 - Cincinnati



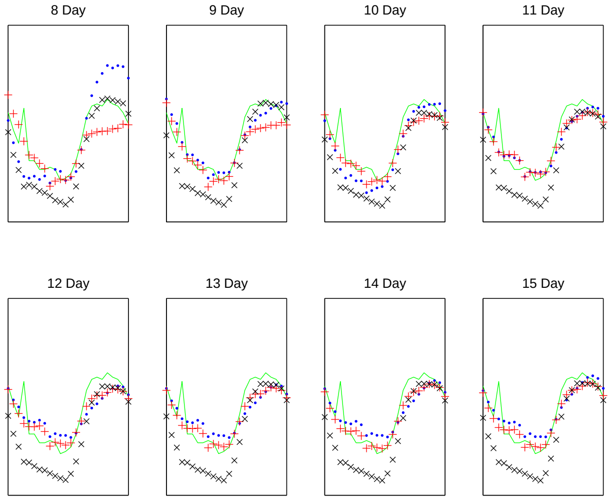
Brute Force: Sept 3 - Cincinnati



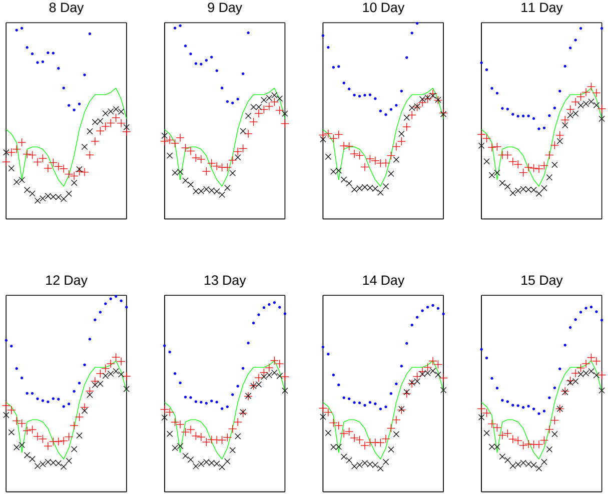
Brute Force: Sept 4 - Cincinnati



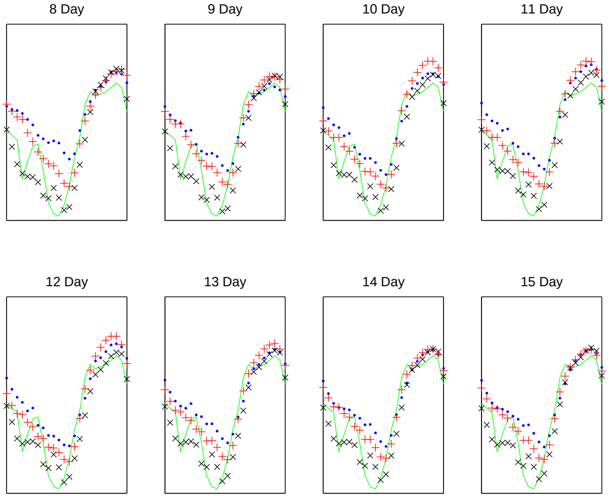
Brute Force: Sept 5 - Cincinnati



Brute Force: Sept 6 - Cincinnati



Brute Force: Sept 7 - Cincinnati



Summary of Numerical Results

- Advantage of Brute Force approach
  - No need to decide how many eigenvalues to use
- Disadvantage of Brute Force approach
  - Sometimes it is not as accurate as Autoregressive approach
- Disadvantage of Autoregressive approach
  - Need to learn how many principal eigenvalues to use
    - Atlanta - 2 eigenvalues
    - Indianapolis - 2 eigenvalues
    - Boston - 7 eigenvalues
- Location in the triangulation important - Example: Boston
- Brute Force improves as $|\triangle| \to 0$

To predict values for locations where there are no measurements, we consider the case when the explanatory and response variables are both random surfaces given by the model

$$Y(s) = G(s,t) * X(t) = \int_{\mathcal{D}} G(s,t)X(t)dt. \qquad (13)$$

where $\mathcal{D} \subset \mathbb{R}^2$ and $G(s,t) \in H \times H$.

We assume we are given a function F such that $F = G * X$ for some function G.

The objective is to recover the function G.

The function, G, is given by solving the following minimization problem:

$$G = arg \min_{\beta \in H \times H} E\left(\int_{s \in \mathcal{D}} \left(F(X)(s) - \beta(s, \cdot) * X(\cdot)\right)^2 ds \right). \quad (14)$$

We find a solution in $S_d^r(\triangle) \times S_d^r(\triangle)$ that is dense in $H \times H$ and approximation to G, in the finite dimensional space $S_G \in S_d^r(\triangle) \times S_d^r(\triangle)$, such that

$$S_G = arg \min_{\beta \in S_d^r(\triangle) \times S_d^r(\triangle)} E\left(\int_{s \in \mathcal{D}} \left(F(X)(s) - \beta(s, \cdot) * X(\cdot)\right)^2 ds \right). \quad (15)$$

Optimal Approximation Theorem for Tensor Product

Theorem

Let $F(s,t) \in H \times H$ and let $Q_F(s,t)$ be the special case of quasi-interpolatory spline defined for the tensor product of two splines:

$$Q_F = \sum_{T'} \sum_{i'+j'+k'=d} \sum_{T} \sum_{i+j+k=d} F(\xi_{ijk}, \xi'_{i'j'k'}) B^T_{ijk}(t) B^{T'}_{i'j'k'}(s). \tag{16}$$

Then

$$|F(s,t) - Q_F(s,t)| \leq 2|\triangle|^{d+1} \max\{|F(s,\cdot)|_{d+1}, |F(\cdot,t)|_{d+1}\} \tag{17}$$

Theorem

Suppose that only the zero spline in $S_d^r(\triangle) \times S_d^r(\triangle)$ is orthogonal to the collection $\mathcal{X} \in H \times H$. Then solution to (15) has a unique solution in $S_d^r(\triangle) \times S_d^r(\triangle)$.

Write $S_G = \sum_{i=1}^{m} c\phi_i$ and rewrite the minimization problem as $Ac = b$, where

$$A = E\left(\int_{\mathcal{D}} (\phi_i(s)\langle X, \phi_j\rangle \phi_l(s)\langle X, \phi_k\rangle)\, ds\right)$$

$$b = E\left(\int_{\mathcal{D}} (F(X)(s)\phi_l(s)\langle X, \phi_k\rangle)\, ds\right).$$

The coefficient vector of S_G is the vector $\mathbf{c} = (c_i,\ i = 1, \ldots, m)$ that satisfies $Ac = b$.

Theorem

Suppose that $E(\|X\|^2) < \infty$ and $G \in \mathcal{C}^{d+1}(\mathcal{D})$. Then the solution S_G from the minimization problem (15) approximates G in the following sense

$$E\left(\int ((G - S_G) * X)^2 ds\right) \leq \left(2|\triangle|^{d+1} \max\{|G(s,\cdot)|_{d+1}, |G(\cdot,t)|_{d+1}\}\right)^2 E\left(\|X\|^2\right).$$

Empirical Estimate

Now we consider the empirical estimate of S_G. Let $X_\ell,\ \ell = 1, \ldots, n$ a sample of be random surfaces in $\mathcal{X}$ such that only the zero spline function in the space $S_d^r(\triangle) \times S_d^r(\triangle)$ is orthogonal to the subspace spanned by $\{X_1, \ldots, X_n\}$. Then the empirical estimate of S_G is

$$\widehat{S_{G,n}} = arg \min_{\beta \in S_d^r(\triangle) \times S_d^r(\triangle)} \frac{1}{n} \sum_{l=1}^{n} \left(\int_{s \in \mathcal{D}} (F(X_\ell(\cdot))(s) - \beta(s, \cdot) * X_\ell)^2 \, ds \right). \tag{18}$$

Theorem

Suppose that only the zero spline function in the space $S^r_d(\triangle) \times S^r_d(\triangle)$ is orthogonal to the subspace spanned by $\{X_1, \ldots, X_n\}$. Then solution to (18) has a unique solution in $S^r_d(\triangle) \times S^r_d(\triangle)$.

Discrete Observations

For applications, we will only know X over some given points in the domain $\mathcal{D}$. That is, we will have observations of X over designated points s_k, $k = 1, \ldots, N$ in $\mathcal{D}$. Let S_X be the penalized least squares fit of X over the designated points s_k. We consider G_D that solves the following minimization problem:

$$G_D = arg \min_{\beta \in H \times H} E\left(\sum_{k=1}^{N} (F(X)(s_k) - \beta(s_k, \cdot) * S_X(\cdot))^2\right). \quad (19)$$

We also look for an approximation $S_{G_D} \in S_d^r(\triangle) \times S_d^r(\triangle)$ of G_D such that

$$S_{G_D} = arg \min_{\beta \in S_d^r(\triangle) \times S_d^r(\triangle)} E\left(\sum_{k=1}^{N} (F(X)(s_k) - \beta(s_k, \cdot) * S_X(\cdot))^2\right). \quad (20)$$

Write $S_{G_D} = \sum_{i=1}^{m} c\phi_i$ and rewrite the minimization problem as $A_D c_D = b_D$, where

$$A_D = E\left(\sum_{k=1}^{N} (\phi_i(s_k)\langle S_X, \phi_j\rangle \phi_l(s_k)\langle S_X, \phi_m\rangle)\, ds\right)$$

$$b_D = E\left(\sum_{k=1}^{N} (F(X)(s)\phi_l(s)\langle S_X, \phi_m\rangle)\, ds\right)$$

The coefficient vector of S_G is the vector $\mathbf{c}_D = (c_i,\ i = 1, \ldots, m)$ that satisfies $A_D c_D = b_D$.

S_{G_D} approximates G_D

Theorem

Suppose that $E(\|S_X\|^2) < \infty$ and $G_D \in \mathcal{C}^{d+1}(\mathcal{D})$. Then the solution S_{G_D} from the minimization problem (20) approximates G_D in the following sense

$$E\left(\sum_{k=1}^{N}((G_D - S_{G_D}) * S_X)^2 ds\right)$$
$$\leq C\left(2|\triangle|^{d+1}\max\{|G(s,\cdot)|_{d+1}, |G(\cdot,t)|_{d+1}\}\right)^2.$$

where C is a constant dependent on $E(\|X\|^2)$.

Empirical Estimate

Let $X_\ell,\ \ell = 1, \ldots, n$ a sample of be random surfaces in $\mathcal{X}$ such that only the zero spline function in the space $S_d^r(\triangle) \times S_d^r(\triangle)$ is orthogonal to the subspace spanned by $\{X_1, \ldots, X_n\}$. Then the empirical estimate of S_{G_D} is

$$\widetilde{S_{G_D,n}} = arg \min_{\beta \in S_d^r(\triangle) \times S_d^r(\triangle)} \frac{1}{n} \sum_{\ell=1}^{n} \sum_{k=1}^{N} \left(F(X_\ell(\cdot))(s_k) - \beta(s_k, \cdot) * S_{X_\ell}(\cdot) \right)^2 . \tag{21}$$

Theorem

Suppose that only the zero spline in $S_d^r(\triangle)$ is orthogonal to the collection $X \in H$. Then solution to (21) has a unique solution in $S_d^r(\triangle) \times S_d^r(\triangle)$.

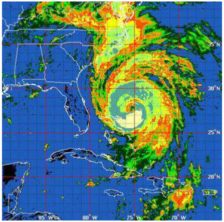
Write $\widetilde{S_{G_D}, n} = \sum_{i=1}^{m} c\phi_i$ and rewrite the minimization problem as $\tilde{A}\tilde{c} = \tilde{b}$, where

$$\tilde{A} = \frac{1}{n} \sum_{\ell=1}^{n} \sum_{k=1}^{N} \phi_i(s_k) \langle S_{X_\ell}, \phi_j \rangle \phi_l(s_k) \langle S_{X_\ell}, \phi_t \rangle$$

$$\tilde{b} = \frac{1}{n} \sum_{\ell=1}^{n} \sum_{k=1}^{N} F(X_\ell)(s_k) \phi_l(s_k) \langle S_{X_\ell}, \phi_t \rangle$$

The coefficient vector of S_G is the vector $\tilde{\mathbf{c}}_D = (c_i,\ i = 1, \ldots, m)$ that satisfies $A_D c_D = b_D$.

- Apply methods to other spatial temporal time series
- Apply for methods to predict values for locations without data
  - Use Autoregressive or Brute Force Method to predict several locations and fit a surface to the prediction
  - Implement Brute Force Extension
- Include covariants in the models
- Increase prediction from one day out to two or more days out



Thank You! Questions?

Awanou, G. and Lai, M. J. and Wenston, P., The multivariate spline method for numerical solution of partial differential equations and scattered data interpolation, in Wavelets and Splines: Athens 2005, G. Chen and M. J. Lai (eds), Nashboro Press, 2006, 24–74.

Cardot, H. and Ferraty, F. and Sarda, P., Spline estimators for the functional linear model, Stat. Sin., 2003, 13, 571-591.

Chui, C. K. and Lai, M. J., Multivariate vertex splines and finite elements, J. Approx. Theory, 60(1990), pp245-343.

Ettinger, B., Guillas, S. and Lai, M. J., *Bivariate Splines for Ozone Concentration Prediction, under preparation, 2007.*

Ferraty, F. and Vieu, P., *Nonparametric Functional Data Analsysis: Theory and Practice*, Springer-Verlag, London, 2006

von Golitschek, M., Lai, M. J., Schumaker,L. L., *Error bounds for minimal energy bivariate polynomial splines*, Numer. Math. 93(2002), 315–331.

von Golitschek, M., and Schumaker, L. L., *Bounds on projections onto bivariate polynomial spline spaces with stable local bases*, Const. Approx. 18(2002), 241–254.

von Golitschek, M., and Schumaker, L. L., *Penalized least squares fitting*, Serdica 18 (2002), 1001–1020.

Guillas, S. and M. J. Lai, Bivariate Splines for Spatial Functional Regression Models, submitted, 2008.

Lai, M. J. and Schumaker,L. L., *Spline Functions over Triangulation*, Cambridge University Press, Cambridge, U.K., 2007.