Functional responses, functional covariates and the concurrent model

1. Predicting precipitation profiles from temperature curves

- Precipitation is much harder to predict than temperature.
- It comes in two main forms:
  - *Drizzle*: Large low pressure zones drop moisture over many hours or days.
  - *Storms*: Convective, short violent storms with a lot of precipitation in a hurry, and spatially localized.
- Precipitation tends to be seasonal; more in the spring and fall than in the summer and winter.

A model

- We can assume that climate zone is important.
- We will predict log precipitation; logging stabilizes variance and eliminates the positivity constraint.
- We will use the difference $\texttt{TempRes}_{mg}(t)$ between a temperature profile and the mean for the climate zone as a function covariate.
- We can extend the functional ANOVA model to

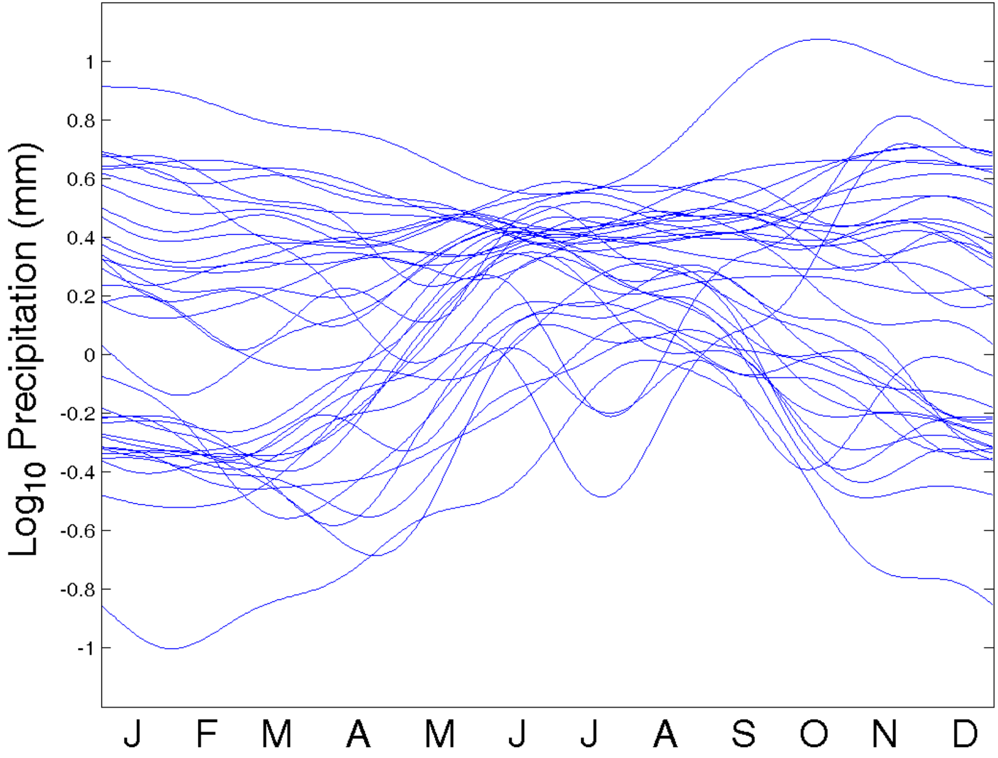
$$\log[\texttt{Prec}_{mg}(t)] = \mu(t) + \alpha_g(t) + \texttt{TempRes}_{mg}(t)\beta(t) + \epsilon_{mg}(t)$$

- We call this model *concurrent* because it assumes that the temperature today affects today's precipitation.

The functional data

- Where precipitation was recorded as 0 mm, we changed it to 0.05 mm, half the minimum positive value.
- We used 11 Fourier series basis functions for precipitation with no roughness penalty.
- We used 21 Fourier series basis functions for temperature with no roughness penalty.

Log precipitation profiles



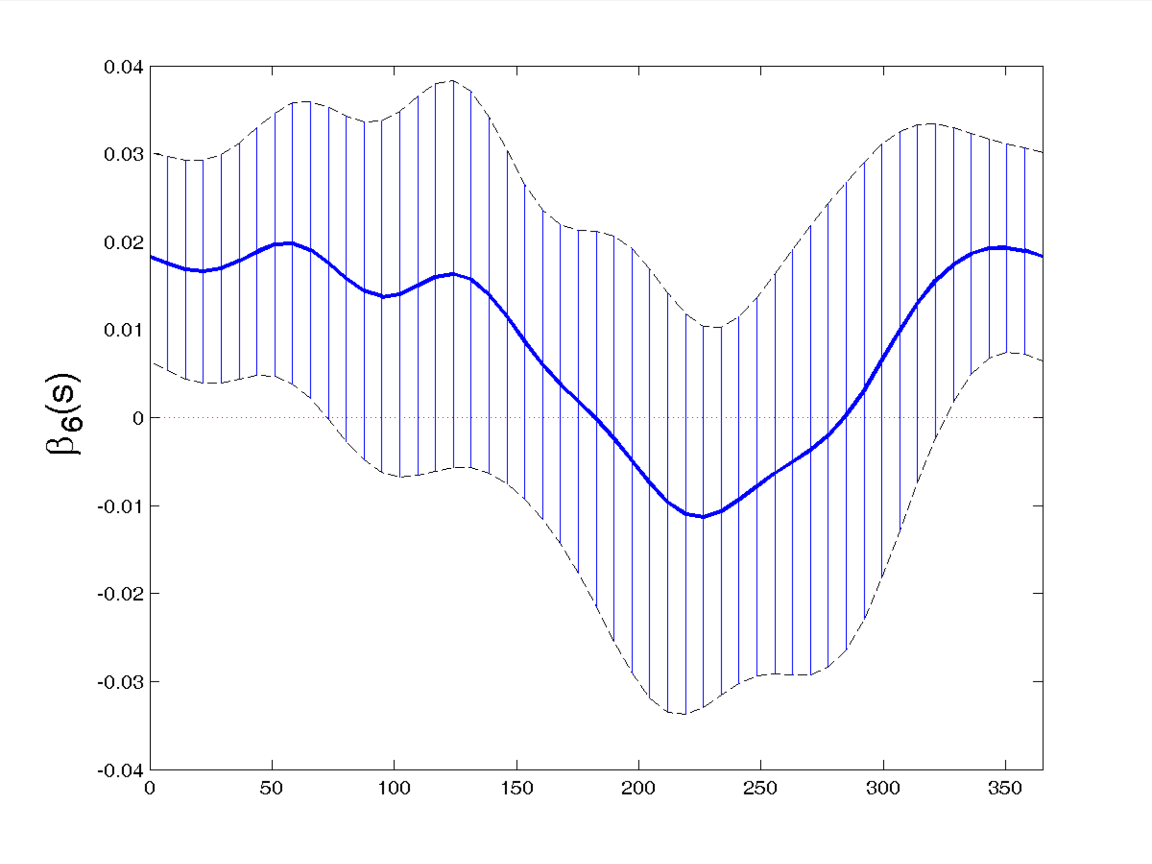
The fitting criterion and some results

- The fitting criterion was the unpenalized error sum of squares

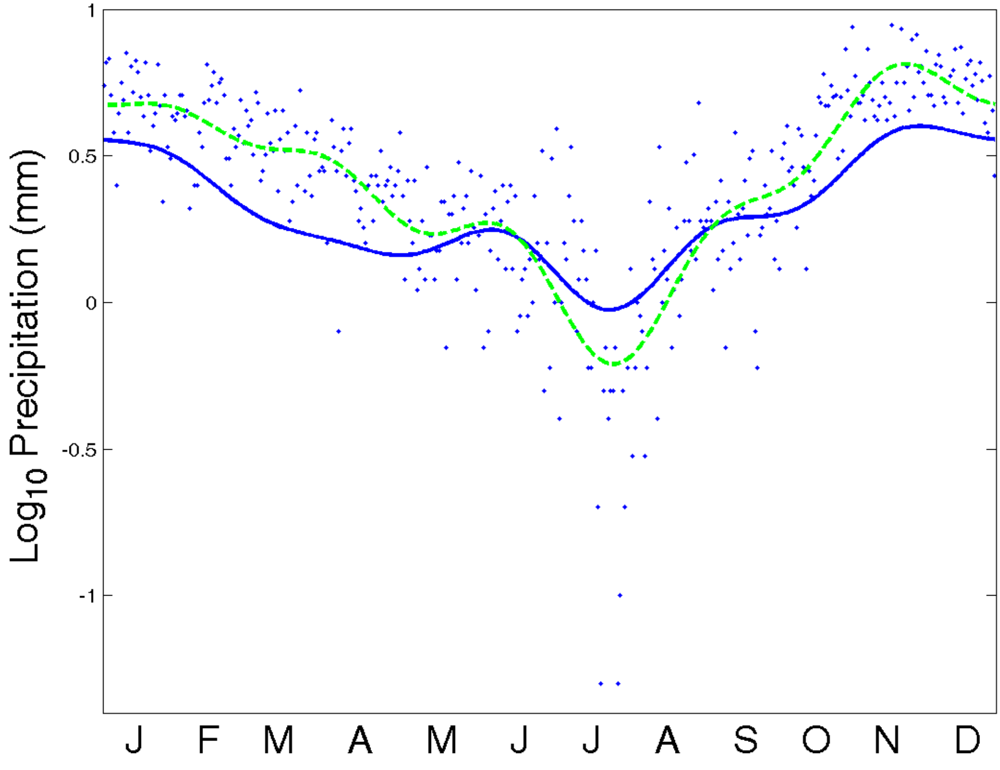
$$\texttt{LMSSE}(\mu, \alpha_g, \beta) = \int \sum_{k,g}^{N} [\texttt{LogPrec}_{kg}(t) - \mu(t) - \alpha_g(t) - \texttt{TempRes}_{kg}(t)\beta(t)]^2 \, dt$$

- The resulting root–mean–squared–residual was 0.19 mm.
- When we dropped $\texttt{TempRes}(t)$ from the model, this increased to 0.20 mm.
- As we see in the following plot, the only place where temperature appears to make a contribution is in mid–winter.

The estimated regression function $\beta_6(t)$



The fit to Vancouver’s data



A probe for the winter effect

- The confidence limits are point–wise; we need a measure of the temperature influence accumulated over the winter months.
- Here is a probe that works:

$$\int_0^{365} \cos[2\pi(t - 64.5)/365]\beta_6(t)\,dt = 2.32\ ,$$

- The estimated standard error of this probe is 0.77, giving a t-ratio of 3.0.
- It appears that elevated temperatures in mid-winter go along with increased precipitation.

2. Fitting the concurrent model

- Here is a general statement of the current functional/functional model:

$$y_i(t) = \sum_{j=1}^{q} z_{ij}(t)\beta_j(t) + \epsilon_i(t) \; .$$

- or in matrix notation:

$$\mathbf{y}(t) = \mathbf{Z}(t)\boldsymbol{\beta}(t) + \boldsymbol{\epsilon}(t) \; ,$$

- We will use a penalized error sum of squares criterion:

$$\texttt{LMSSE}(\boldsymbol{\beta}) = \int [\mathbf{y}(t) - \mathbf{Z}(t)\boldsymbol{\beta}(t)]'[\mathbf{y}(t) - \mathbf{Z}(t)\boldsymbol{\beta}(t)]\, dt + \sum_{j}^{p} \lambda_j \int [L_j \beta_j(t)]^2 \, dt \quad .$$

- where L is the harmonic acceleration operator

The basis function expansions for $\beta_j(s)$

- Let regression function $\beta_j(s)$ have the expansion

$$\beta_j(s) = \mathbf{b}_j' \boldsymbol{\theta}_j(s)$$

in terms of K_j basis functions $\theta_{jk}(s)$.

- Some of the independent variables can be scalar; in this case the basis for their $\beta_j(s)$'s is the constant basis;

$$\theta_{j1}(s) = 1$$

- Defining $K_\beta = \sum_j^q K_j$, we construct vector $\mathbf{b}$ of length K_β by stacking the coefficient vectors vertically, that is,

$$\mathbf{b} = (b_1', b_2', \ldots, b_q')' \ .$$

- Now assemble q by K_β matrix function $\boldsymbol{\Theta}$ as follows:

$$\boldsymbol{\Theta} = \begin{bmatrix} \boldsymbol{\theta}_1' & 0 & \cdots & 0 \\ 0 & \boldsymbol{\theta}_2' & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{\theta}_q' \end{bmatrix} \ .$$

- We can now express our model as

$$\mathbf{y}(t) = \mathbf{Z}(t)\boldsymbol{\Theta}(t)\mathbf{b} + \boldsymbol{\epsilon}(t) \ .$$

- We also need to arrange the order K_j roughness penalty matrices

$$\lambda_j \mathbf{R}_j = \lambda_j \int L\boldsymbol{\theta}_j(t) L\boldsymbol{\theta}'_j(t)\, dt$$

into the symmetric block diagonal matrix $\mathbf{R}$ of order K_β:

$$\mathbf{R} = \begin{bmatrix} \lambda_1 \mathbf{R}_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 \mathbf{R}_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_q \mathbf{R}_q \end{bmatrix}. \qquad (1)$$

The normal equations

- $$[\int \mathbf{\Theta}'(t)\mathbf{Z}'(t)\mathbf{Z}(t)\mathbf{\Theta}(t)\,dt + \mathbf{R}]\mathbf{b} \;=\; [\int \mathbf{\Theta}'(t)\mathbf{Z}'(t)\mathbf{y}(t)\,dt]$$

- The numerical integration in these equations is not as difficult as it seems. The scalar functions

$$\omega_{j\ell}(t) = \sum_i^N z_{ij}(t) z_{i\ell}(t)$$

play the role of *weighting functions* for the functional inner products

$$\int \boldsymbol{\theta}_j(t)\boldsymbol{\theta}'_\ell(t)\omega_{j\ell}(t)\,dt, j, \ell = 1, \ldots, q\;.$$