Weakly dependent functional data

Piotr Kokoszka

Utah State University

Joint work with

Siegfried Hörmann

University of Utah

Outline

Examples of functional time series

L^4–m– approximability

Convergence of Eigenvalues and Eigenfunctions

Estimation of the long–run variance

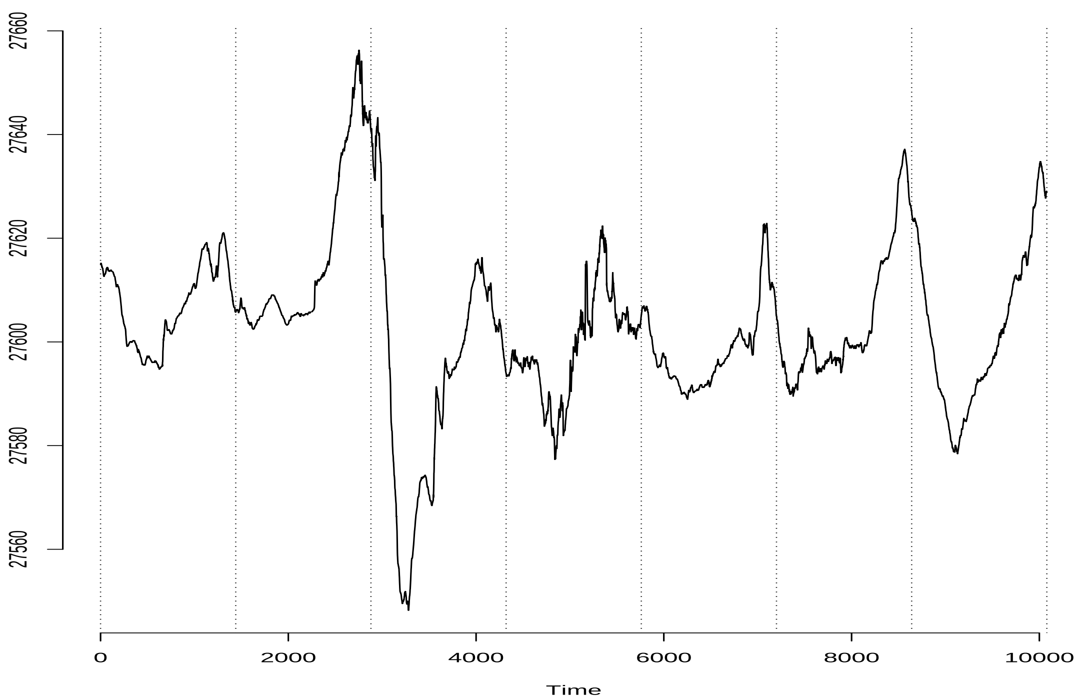
Change point detection

Functional linear model with dependent dependent regressors

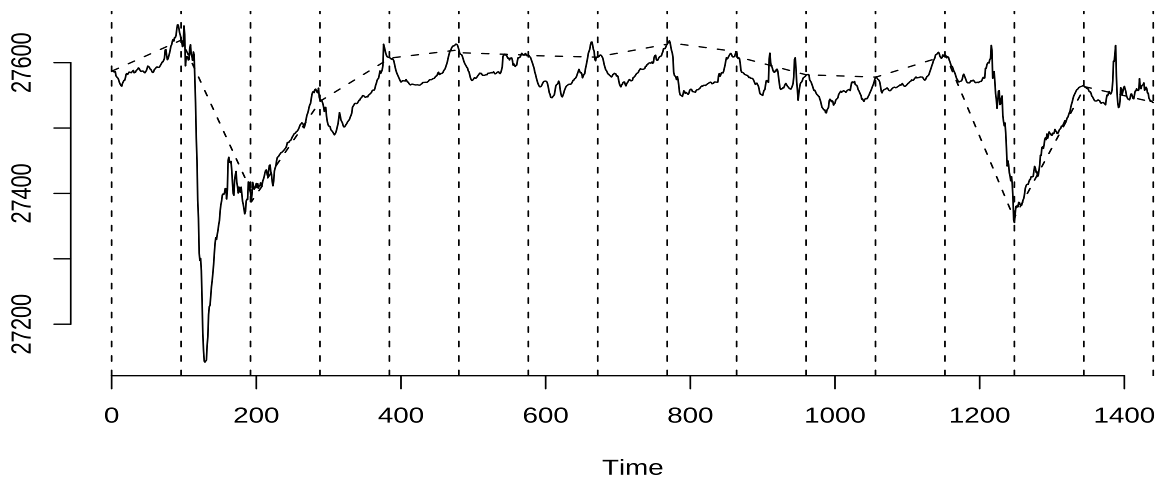
Seven functional time series observations

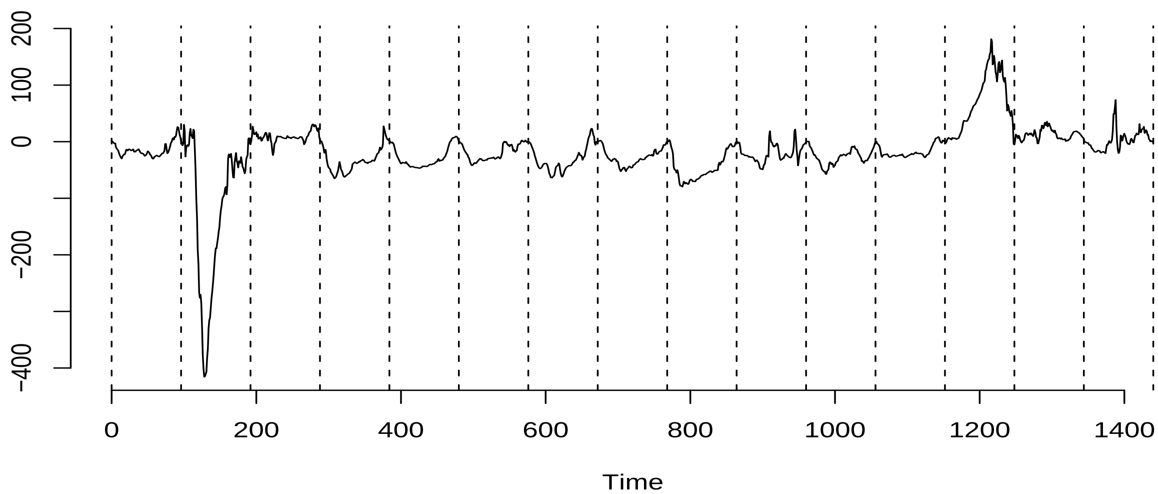
(1440 measurements per day)

The horizontal component of the magnetic field measured in one minute resolution at Honolulu magnetic observatory from 1/1/2001 00:00 UT to 1/7/2001 24:00 UT.



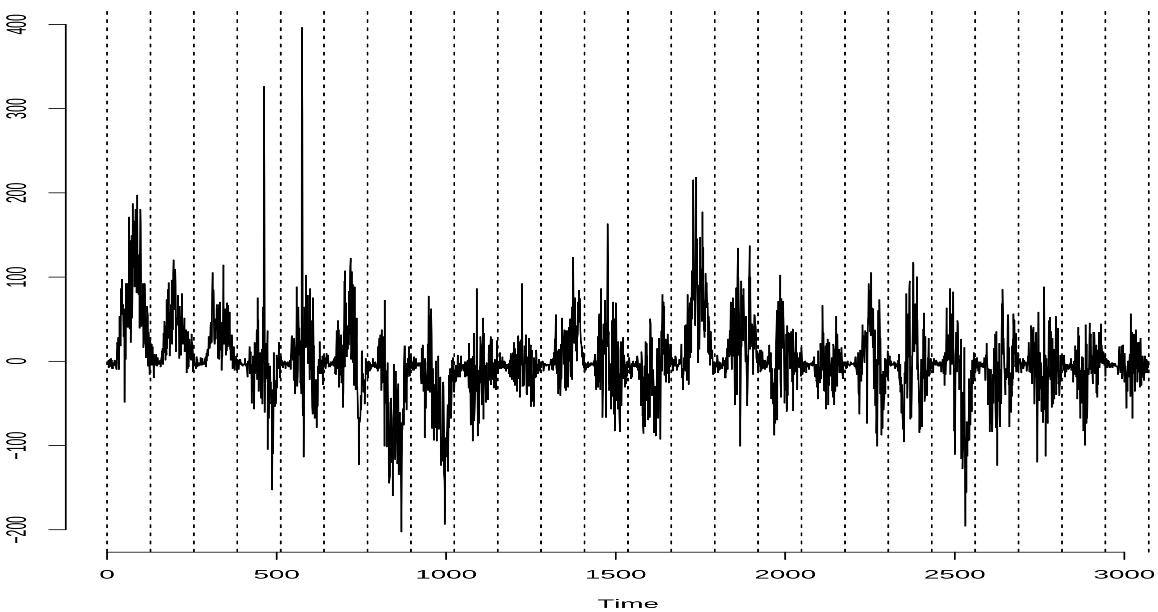
Functional time series can be transformed to stationary series (removal of trend, differencing, ad hoc)





Autoregressive dependence:

Three weeks of a time series derived from credit card transaction data. The vertical dotted lines separate days.



Definition

The sequence $\{X_n\}$ of functions in L^2 is

L^4–m–*approximable*

if it admits the representation

$$X_n = f(\varepsilon_n, \varepsilon_{n-1}, \ldots),$$

where the ε_i are iid elements in a measurable space S, and f is a measurable function $f : S^\infty \to L^2$.

Let $\{\varepsilon_i'\}$ be an independent copy of $\{\varepsilon_i\}$, set

$$X_n^{(m)} = f(\varepsilon_n, \varepsilon_{n-1}, \ldots, \varepsilon_{n-m+1}, \varepsilon'_{n-m}, \varepsilon'_{n-m-1}, \ldots)$$

Main Condition:

$$\sum_{m=1}^{\infty} \left(E||X_m - X_m^{(m)}||^4 \right)^{1/4} < \infty.$$

Discussion

The main condition depends only on

$$\nu_4\left(X_0 - X_0^{(m)}\right) = \left(E||X_0 - X_0^{(m)}||^4\right)^{1/4}$$

$\{\varepsilon_k^{(n)}, k \in Z\}$ independent copy of $\{\varepsilon_k, k \in Z\}$

Sequences $\{\varepsilon_k^{(n)}, k \in Z\}$ are independent

$$X_n^{(m)} = f(\varepsilon_n, \varepsilon_{n-1}, \ldots, \varepsilon_{n-m+1}, \varepsilon_{n-m}^{(n)}, \varepsilon_{n-m-1}^{(n)}, \ldots)$$

For each m, the sequence $\left\{X_n^{(m)}, n \in Z\right\}$ is m dependent.

$$X_0^{(2)} = f(\varepsilon_0, \varepsilon_{-1}, \varepsilon_{-2}^{(0)}, \varepsilon_{-3}^{(0)}, \ldots)$$

$$X_2^{(2)} = f(\varepsilon_2, \varepsilon_1, \varepsilon_0^{(2)}, \varepsilon_{-1}^{(2)}, \ldots)$$

An alternative definition

$$X_n^{(m)} = f^{(m)}(\varepsilon_n, \varepsilon_{n-1}, \dots, \varepsilon_{n-m+1})$$

For each m, the sequence $\left\{X_n^{(m)}, n \in Z\right\}$ is m dependent.

This $X_n^{(m)}$ does not have the same distribution as X_n (extra line in proofs.)

Functions $f^{(m)}$ very natural in examples, e.g.

$$f^{(m)}(\varepsilon_n, \varepsilon_{n-1}, \dots, \varepsilon_{n-m+1})$$

$$= f(\varepsilon_n, \varepsilon_{n-1}, \dots, \varepsilon_{n-m+1}, 0, 0, \dots)$$

This definition produces a (theoretically) narrower class.

Example: AR(1)

$$X_n(t) = \int \psi(t,s) X_{n-1}(s) ds + \varepsilon_n(t),$$

$$X_n = \sum_{j=0}^{\infty} \Psi^j(\varepsilon_{n-j}), \quad (||\Psi|| < 1)$$

$$X_n^{(m)} = \sum_{j=0}^{m-1} \Psi^j(\varepsilon_{n-j}) + \sum_{j=m}^{\infty} \Psi^j(\varepsilon_{n-j}^{(n)}).$$

or

$$X_n^{(m)} = \sum_{j=0}^{m-1} \Psi^j(\varepsilon_{n-j})$$

$$\nu_4(X_m - X_m^{(m)}) \le (2) \sum_{j=m}^{\infty} ||\Psi||^j \nu_4(\varepsilon_0)$$

$$\sum_{m=1}^{\infty} \nu_4(X_m - X_m^{(m)}) \le O(1) \sum_{m=1}^{\infty} ||\Psi||^m < \infty,$$

Example: bilinear model

$$X_n(t) = U_n Y_n(t)$$

$\{U_n\}$ independent of $\{Y_n\}$.

$$X_n^{(m)}(t) = U_n^{(m)} Y_n^{(m)}(t)$$

Example: Functional ARCH

$$y_k(t) = \varepsilon_k(t)\sigma_k(t),$$

where

$$\sigma_k^2(t) = \delta(t) + \int_0^1 \beta(t,s)\sigma_{k-1}^2(s)\varepsilon_{k-1}^2(s)ds$$

Bounds on Eigenfunctions and Eigenvalues

Theorem:

$$NE||\widehat{C} - C||_{\mathcal{S}}^2 \leq U_X,$$

where

$$U_X = \nu_4^4(X) + 4\sqrt{2}\ \nu_4^3(X) \sum_{r=1}^{\infty} \nu_4(X_r - X_r^{(r)}).$$

Corollary:

$$NE\left[|\lambda_j - \widehat{\lambda}_j|^2\right] \leq U_X.$$

Set $\widehat{c}_j = \text{sign}(< \widehat{v}_j, v_j >)$

$$NE\left[||\widehat{c}_j \widehat{v}_j - v_j||^2\right] \leq \frac{8U_X}{\alpha_j^2},$$

where $\alpha_1 = \lambda_1 - \lambda_2$ and

$$\alpha_j = \min(\lambda_{j-1} - \lambda_j, \lambda_j - \lambda_{j+1}),\ 2 \leq j \leq d.$$

Long–run variance

introduction: scalar case.

$$N\text{Var}[\bar{X}_N] \to \sum_{j=-\infty}^{\infty} \gamma_j, \quad \gamma_j = \text{Cov}(X_0, X_j)$$

If $\{X_n\}$ is L^2–m–approximable,
then $\sum_{j=-\infty}^{\infty} |\gamma_j| < \infty$

Proof:

$$\text{Cov}(X_0, X_j) = \text{Cov}(X_0, X_j - X_j^{(j)}) + \text{Cov}(X_0, X_j^{(j)}).$$

$$X_0 = f(\varepsilon_0, \varepsilon_{-1}, \ldots), \quad X_j^{(j)} = f^{(j)}(\varepsilon_j, \varepsilon_{j-1}, \ldots, \varepsilon_1),$$

are independent

$$|\gamma_j| \le [EX_0^2]^{1/2}[E(X_j - X_j^{(j)})^2]^{1/2}.$$

Kernel estimator:

$$\hat{\sigma} = \sum_{|j| \le q} \omega_q(j)\hat{\gamma}_j,$$

Consistency requires cumulant conditions:

$$\sup_h \sum_{r=-\infty}^{\infty} \sum_{s=-\infty}^{\infty} |\kappa(h, r, s)| < \infty.$$

$$\kappa(h, r, s) = E[(X_0 - \mu)(X_h - \mu)(X_r - \mu)(X_s - \mu)]$$

$$-(\gamma_h \gamma_{r-s} + \gamma_r \gamma_{h-s} + \gamma_r \gamma_{h-r}),$$

For an L^4–m–approximable sequence

$$\sup_{k,l \ge 0} \sum_{r=1}^{\infty} \left| \mathrm{Cov}\left(X_0(X_k - X_k^{(k)}), X_r^{(r)} X_{r+\ell}^{(r+\ell)} \right) \right| < \infty.$$

$$X_k = \sum_{j=1}^{\infty} c_j \varepsilon_{k-j}; \qquad X_k^{(k)} = \sum_{j=1}^{k-1} c_j \varepsilon_{k-j}$$

$X_0 \left(X_k - X_k^{(k)} \right)$ depends on $\{\varepsilon_j, j \le 0\}$

$X_r^{(r)} X_{r+\ell}^{(r+\ell)}$ depends on $\{\varepsilon_j, 1 \le j \le r + \ell\}$.

Condition holds trivially

Approximability: $\sum_{m=1}^{\infty} \sum_{j=m}^{\infty} |c_j| < \infty$

Long–run variance for functional time series

Project functions on principal components

The resulting sequence of vectors has a long run covariance matrix

It inherits L^4–m–approximability

kernel estimator is consistent under the condition

$$N^{-1} \sum_{k,l=0}^{q(N)} \sum_{r=1}^{N-1}$$

$$\max_{1 \le i,j \le d} \left| \mathrm{Cov}\left(X_{i0}(X_{jk} - X_{jk}^{(k)}), X_{ir}^{(r)} X_{j,r+\ell}^{(r+\ell)} \right) \right| \to 0.$$

X_{jk} – projection of kth observation on jth principal component

Change point detection

H_0 :

$$X_i(t) = \mu(t) + Y_i(t), \quad EY_i(t) = 0.$$

H_A :

$$X_i(t) = \begin{cases} \mu_1(t) + Y_i(t), & 1 \le i \le k^*, \\ \mu_2(t) + Y_i(t), & k^* < i \le N, \end{cases}$$

Test statistic is based on

$$\int \left\{ \sum_{1 \le i \le k} X_i(t) - \frac{k}{N} \sum_{1 \le i \le N} X_i(t) \right\} \widehat{v}_\ell(t) dt$$

$$= \sum_{1 \le i \le k} \widehat{\eta}_{\ell i} - \frac{k}{N} \sum_{1 \le i \le N} \widehat{\eta}_{\ell i}.$$

For functional time series it is different than for iid data (long–run variance).

Required asymptotic properties can be established under L^4–m–approximability.

Functional linear model with dependent regressors

$$Y_n(t) = \int \psi(t,s) X_n(s) + \varepsilon_n.$$

Idea of Yao, Müller, Wang:

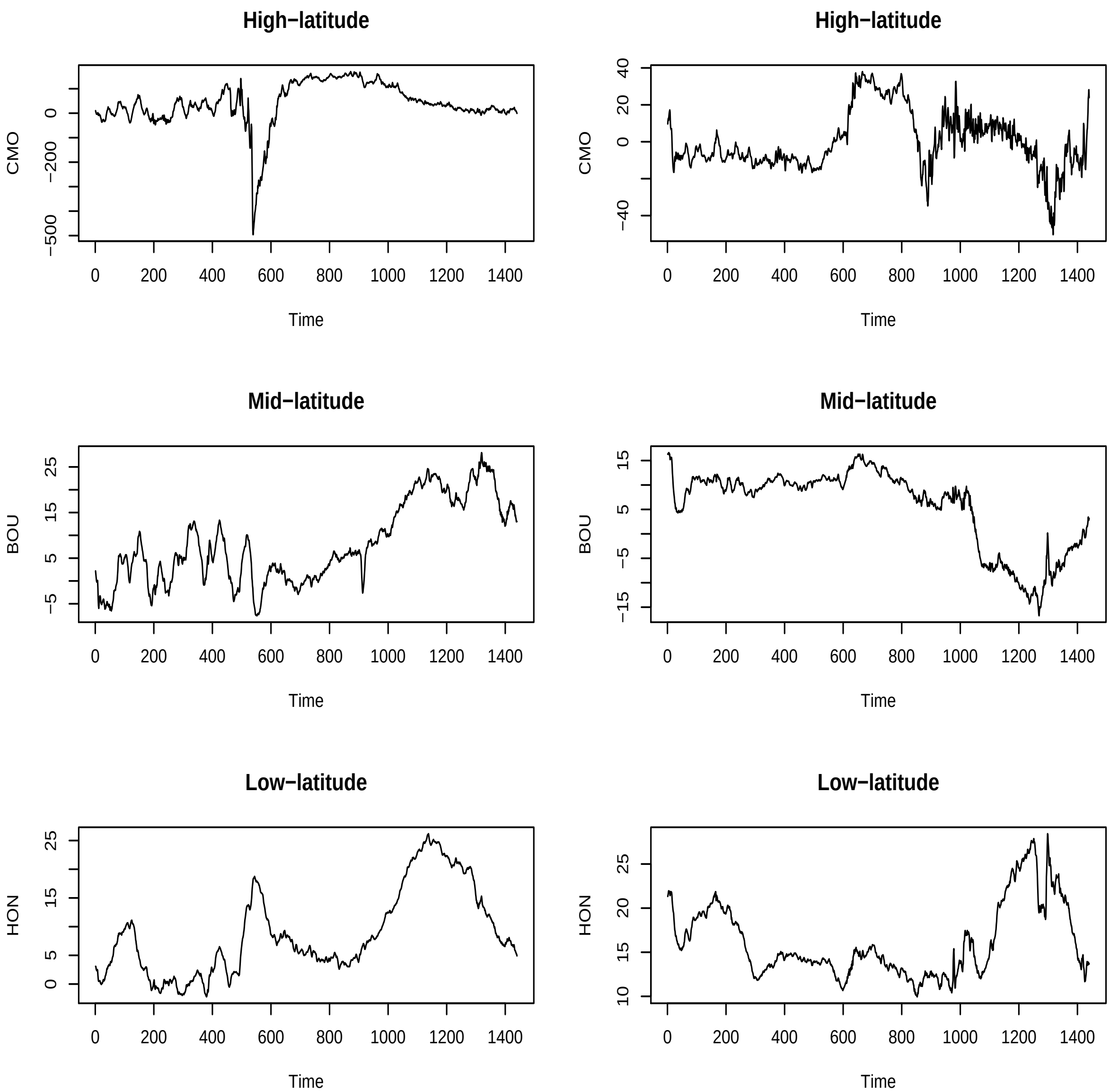
$$X(s) = \sum_{i=1}^{\infty} \xi_i v_i(s), \qquad Y(t) = \sum_{j=1}^{\infty} \zeta_j u_j(t),$$

$$\psi(t,s) = \sum_{k=1}^{\infty} \sum_{\ell=1}^{\infty} \frac{E[\xi_\ell \zeta_k]}{E[\xi_\ell^2]} u_k(t) v_\ell(s).$$

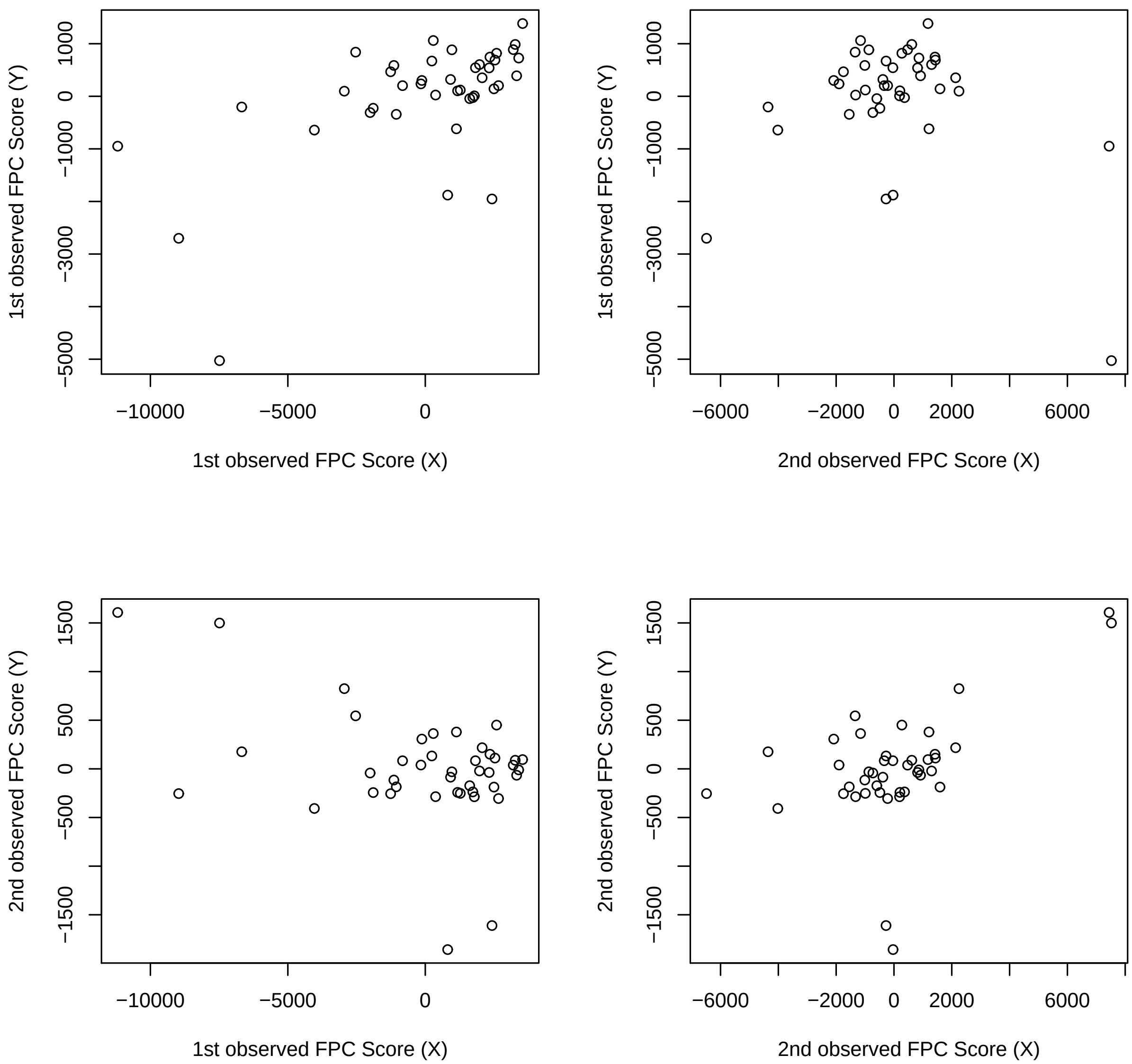
$$\widehat{\psi}_{KL}(t,s) = \sum_{k=1}^{K} \sum_{\ell=1}^{L} \widehat{\lambda}_\ell^{-1} \widehat{\sigma}_{\ell k} \widehat{u}_k(t) \widehat{v}_\ell(s),$$

YMW focus on *smoothing*, assumptions (2+ pages) deal mostly with smoothing parameters,
K and L depend on the smoothing bandwidths, integrals of FT of kernels, and the rate of decay of the eigenfunctions.

Examples of Magnetometer data



Horizontal intensities of the magnetic field measured at a high-, mid- and low-latitude stations during a sub-storm (left column) and a quiet day (right column). Note the different vertical scales for high-latitude records.



Functional predictor-response plots of FPC scores of response functions versus FPC of explanatory functions for magnetometer data (CMO vs THY0)

For functional time series data, the focus is on *dependence*.

We assume that the X_n are L^4–m–approximable.

λ_j , γ_j eigenvalues corresponding to v_j u_j.

Recall

$$\alpha_j = \min(\lambda_{j-1} - \lambda_j, \lambda_j - \lambda_{j+1}),\ 2 \leq j \leq d.$$

Define α'_j correspondingly for γ_j

Set $g_L = \min\{\lambda_j | 1 \leq j \leq L\}$,
$h_L = \min\{\alpha_j \wedge \alpha'_j | 1 \leq j \leq L\}$.

Assumption for consistency:

$$L = o(N^{1/12}(g_L h_L)^{1/2}).$$

The same for K.