Change point tests in functional factor models with application to yield curves

Piotr Kokoszka

Department of Statistics, Colorado State University

Joint work with

Lajos Horváth

University of Utah

Gabriel Young

Columbia University

Patrick Bardsley

University of Texas

Scalar change point model

$$X_i = \mu_i + u_i, \quad 1 \leq i \leq N, \qquad Eu_i = 0.$$

Null hypothesis:

$$H_0 : \quad \mu_1 = \mu_2 = \ldots = \mu_N.$$

H_A: the means change in some way: abruptly, gradually, epidemically ...

It is typically assumed that the errors u_i are homoskedastic, the series u_i is stationary.

There has been limited work for nonstationary u_i: Cavaliere, Taylor, Dalla, Giraitis, Phillips ...

CUSUM approach

CUSUM process:

$$Z_N(t) = N^{-1/2} \left(\sum_{\ell=1}^{[Nt]} X_\ell - \frac{[Nt]}{N} \sum_{\ell=1}^{N} X_\ell \right).$$

Functionals of this process are used as tests statistics.

If the u_i are iid, Z_N typically leads to optimal tests, it can be connected to the likelihood method.

There has been recent work for the u_i which are dependent and and exhibit change/points or irregular trends in variability.

Change points or trends in the variability of the u_i can confound change points in the mean of the X_i.

Assumptions on the u_i

$$u_i = u_{i,N} = a(i/N)e_i$$

$[0,1] \ni t \mapsto a(t)$ has regular variation.

$$N^{-1/2} \sum_{\ell=1}^{[Nt]} e_\ell \quad \xrightarrow{\mathcal{D}[0,1]} \quad \sigma W(t).$$

(no mixing needed). Set

$$b(t) = \sigma^2 \int_0^t a^2(u)du.$$

Then

$$N^{-1/2} \sum_{\ell=1}^{[Nt]} u_\ell \quad \xrightarrow{\mathcal{D}[0,1]} \quad W(b(t)).$$

Main difficulty: the function $b(\cdot)$ is unknown.

Approach to testing

Under H_0,

$$Z_N(t) \stackrel{\mathcal{D}[0,1]}{\longrightarrow} \Gamma(t), \text{ where } \Gamma(t) = W(b(t)) - tW(b(1)).$$

$$C(t,s) = E[\Gamma(t)\Gamma(s)] = b(t \wedge s) - tb(s) - sb(t) + tsb(1).$$

$b(t)$ is estimated by the LRV of $X_i, i \leq [Nt]$.
λ_i eigenvalues of $C(\cdot,\cdot)$

$$\int_0^1 Z_N^2(t)dt \stackrel{\mathcal{D}}{\to} \int_0^1 \Gamma^2(t)dt. \stackrel{\mathcal{D}}{=} \sum_{i=1}^{\infty} \lambda_i \xi_i^2.$$

Weighted functionals, e.g. Anderson–Darling, are possible.

Yield curves

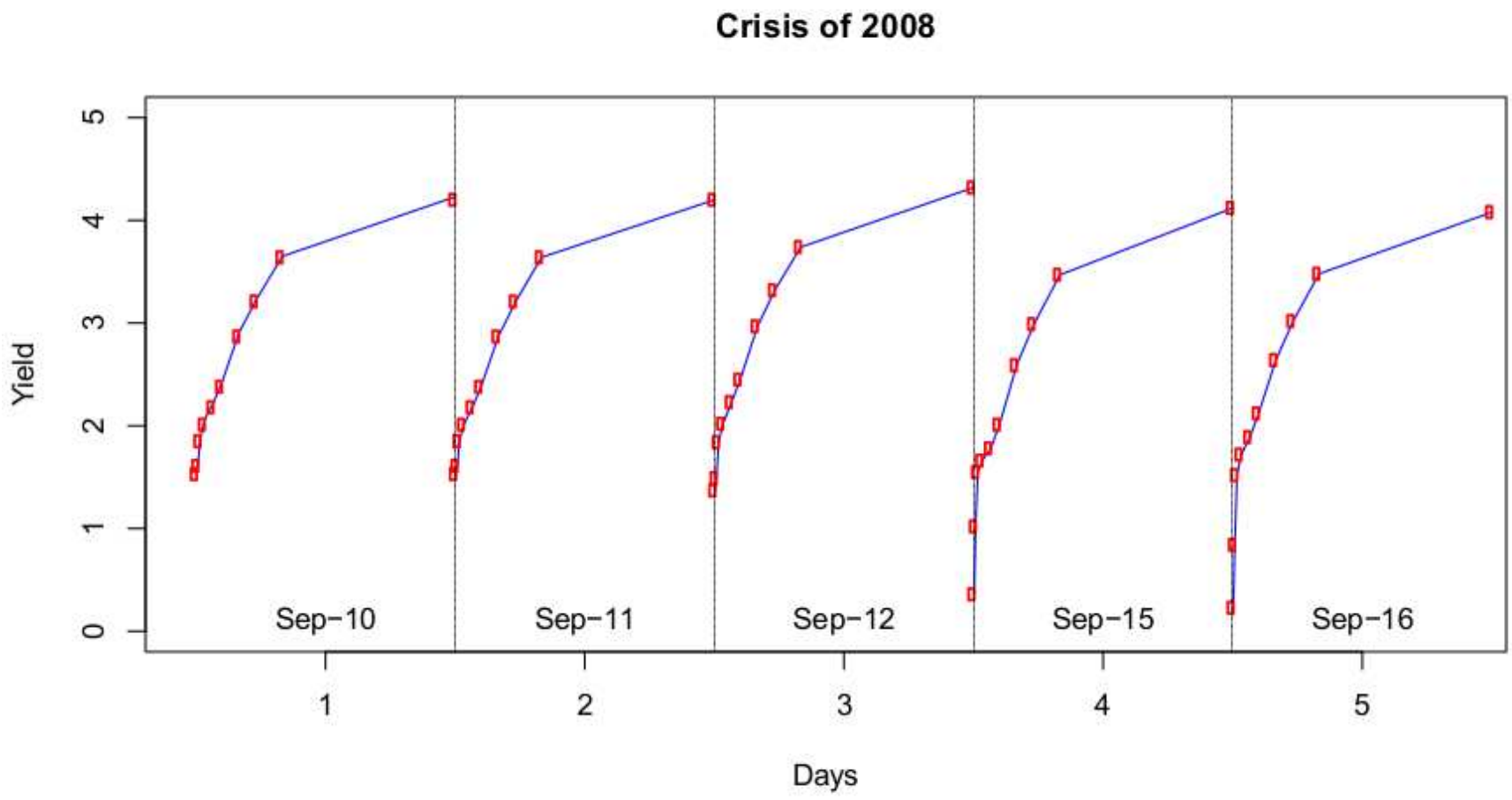
i - day (or month)

t_j - time to maturity (1 month, 3 months, ... , 30 years)
Fractions of bonds with continuous maturities are traded.
Our theory can be in discrete or continuous time t.

$X_i(t_j)$ - yield of a bond bought on day i and held for time t_j.
For the theory, we can use, the yield curves, $X_i(t), t \in [0, T]$.

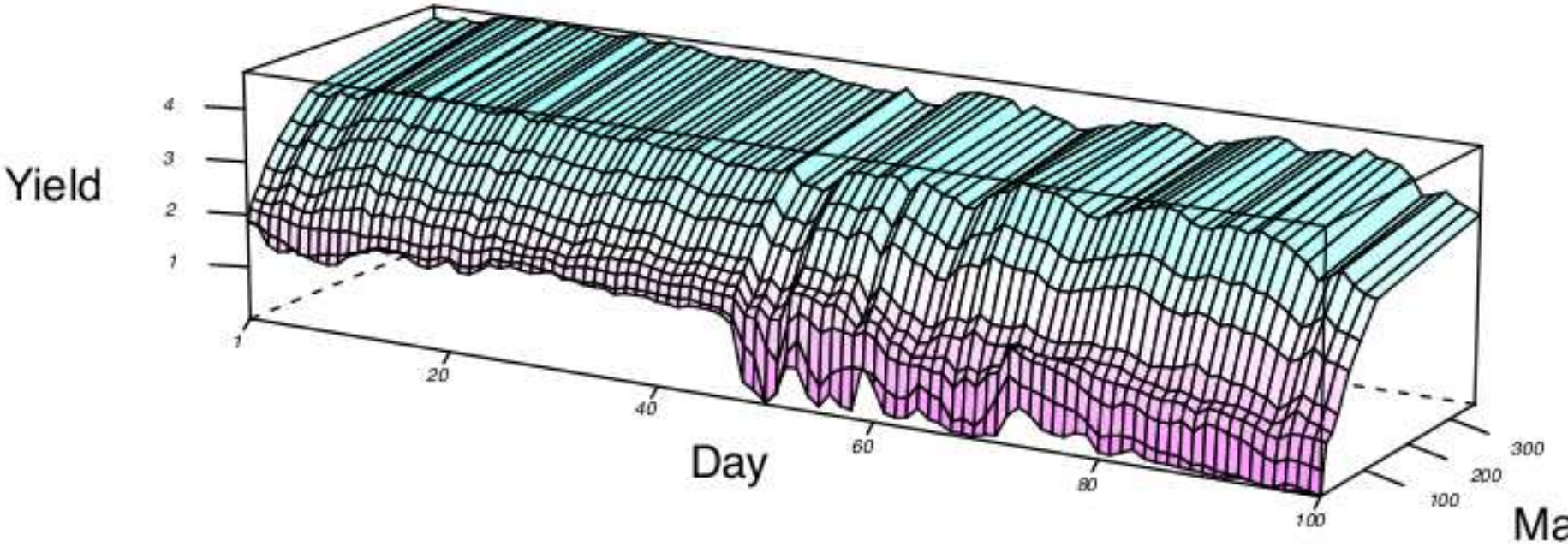
Yield curves

US yield curves over five business days around the Lehman Brothers bankruptcy filing on Sep 15 2008.



Yield curves

US yield curves over $N = 100$ business days; the central time point corresponds to the Lehman Brothers collapse.



Factor models

$$X_i(t_j) = \sum_{k=1}^{K} \beta_{i,k} f_k(t_j) + \varepsilon_i(t_j).$$

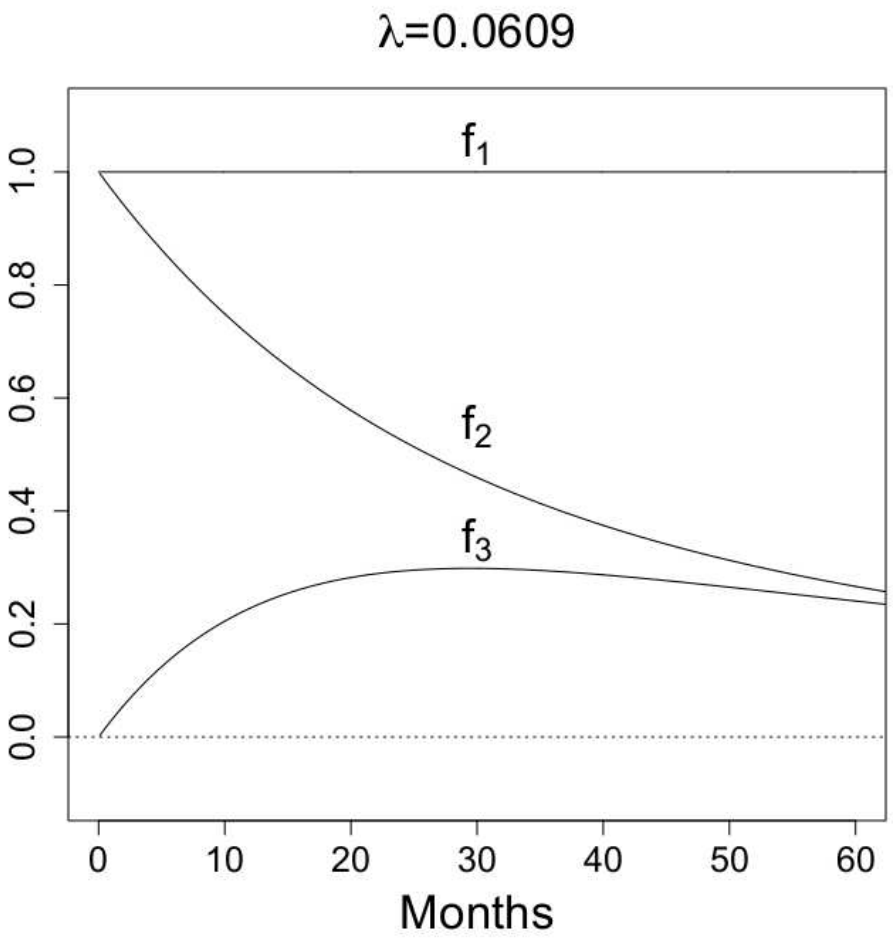
The f_k are deterministic functions,
treated as known in standard finance methodology

The K time series $\{\beta_{i,k}, \ = 1, 2, 3, \ldots\}$ are modeled as time series (modern dynamic models).

If $\beta_{i,k} \equiv \beta_k$, static model.

Nelson–Siegel factors

Nelson–Siegel factors $f_1(t,\lambda), f_2(t,\lambda), f_3(t,\lambda)$.



Functional vs. multivariate modeling

$\{X_i(t_j),\ i = 1, 2, 3, \ldots\}$ is a multivariate time series of dimension $J = 10$.

Why not model it as a VAR process?

The yield curves have a specific shape whose components have economic interpretation.

Factors lead to nontrivial and practically relevant dimension reduction and improved estimation and prediction.

Change point test in a functional factor model

$$X_i(t) = \sum_{k=1}^{K} \beta_{i,k} f_k(t) + \varepsilon_i(t), \quad 1 \le i \le N.$$

(Continuous time)

$$\beta_{i,k} = \mu_{i,k} + b_{i,k}, \quad Eb_{i,k} = 0.$$

$$\boldsymbol{\mu}_i = [\mu_{i,1}, \mu_{i,2}, \ldots, \mu_{i,K}]^\top$$

$$H_0: \quad \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \ldots = \boldsymbol{\mu}_N.$$

Change point vs. break point

Recall: H_0: mean structure does not change

Error functions: $\sum_{k=1}^{K} b_{i,k} f_k(t) + \varepsilon_i(t)$.
The distribution of the error functions can change at *known* points:

$$1 = i_0 < i_1 < i_2 < \ldots < i_M < i_{M+1} = N.$$

(Similar to prior information in Bayesian inference)
On each segment, error functions are realizations of potentially different weakly dependent processes in L^2.

Compare:
unknown *change points* in mean structure
known *break points* in error structure

Determining break points

Dates of substantial central bank intervention,

Dates of events of economic impact,

Exploratory analysis of the variability of the yield curves.

After the application of the test, some change points may be close to break points.

Detection through projections onto factors

Vectors of projections:

$$\mathbf{z}_i = [\langle X_i, f_1 \rangle, \ldots, \langle X_i, f_K \rangle]^\top$$

Introduce the *deterministic* matrix

$$\mathbf{C} = [\langle f_k, f_j \rangle, \ 1 \leq k, j \leq K].$$

and random vectors

$$\mathbf{b}_i = [b_{i,1}, b_{i,2}, \ldots, b_{i,K}]^\top, \qquad \varepsilon_i = [\langle \varepsilon_i, f_1 \rangle, \langle \varepsilon_i, f_2 \rangle, \ldots, \langle \varepsilon_i, f_K \rangle]^\top.$$

Then

$$\mathbf{z}_i = \mathbf{C}\boldsymbol{\mu}_i + \gamma_i, \quad \gamma_i = \mathbf{C}\mathbf{b}_i + \varepsilon_i,$$

Change in the vectors $\boldsymbol{\mu}_i$ is equivalent to a change in the $\mathbf{z}_i$ at the same change points.

Detection through projections onto factors

CUSUM process:

$$\boldsymbol{\alpha}_N(x) = N^{-1/2}\left(\sum_{i=1}^{[Nx]} \mathbf{z}_i - \frac{[Nx]}{N}\sum_{i=1}^{N} \mathbf{z}_i\right), \quad 0 \le x \le 1.$$

Even under H_0, the distribution of the $\mathbf{z}_i$ can change at break points.

$\left\{\gamma_i^{(m)}\right\}$ - model for the vector errors γ_i over the interval $(i_m, i_{m+1}]$.

Long–run covariance matrices:

$$\mathbf{V}_m = \sum_{\ell=-\infty}^{\infty} \mathrm{cov}(\gamma_i^{(m)}, \gamma_{i+\ell}^{(m)}).$$

Detection through projections onto factors

Limit distribution of the CUSUM process:

$$\boldsymbol{\alpha}_N \to \mathbf{G}^0, \quad \text{in } D^K([0,1]), \quad \mathbf{G}^0(x) = \mathbf{G}(x) - x\mathbf{G}(1).$$

$\{\mathbf{G}(x),\ x \in [0,1]\}$ is a mean zero $\mathbb{R}^K$–valued Gaussian process with covariances

$$E[\mathbf{G}(x)\mathbf{G}^\top(y)] = \sum_{j=1}^{m}(\theta_j - \theta_{j-1})\mathbf{V}_j + (x - \theta_m)\mathbf{V}_{m+1}, \ \ \theta_m \le x \le \theta_{m+1}, y \ge x$$

The covariances of the process $\mathbf{G}^0$ can be computed explicitly (a long formula).
Cramér–von–Mises functional

$$\int_0^1 \|\boldsymbol{\alpha}_N(x)\|^2\, dx \to \int_0^1 \left\|\mathbf{G}^0(x)\right\|^2 dx.$$

We must simulate the distribution of the right–hand side.

Detection through projections onto factors

1. Eigenvalue method (**Eigen**):

Estimate $\mathbf{R}(x,y) = E[\mathbf{G}^0(x)\mathbf{G}^0(y)^\top]$.
Use equality in distribution:

$$\int_0^1 \left\|\mathbf{G}^0(x)\right\|^2 dx \sim \sum_{j=1}^{\infty} \lambda_j Z_j^2.$$

The Z_j are independent standard normal;

$$\int_0^1 \mathbf{R}(x,y)\phi_j(y)dy = \lambda_j \phi_j(x).$$

Detection through projections onto factors

2. Direct simulation method (**Sim**):

Generate replications of the process $\mathbf{G}^0$. Enough to generate

$$\mathbf{G}(x) = \sum_{j=1}^{m} (\theta_j - \theta_{j-1})^{1/2} \mathbf{G}_j(1) + (\theta_{m+1} - \theta_m)^{1/2} \mathbf{G}_{m+1}\left(\frac{x - \theta_m}{\theta_{m+1} - \theta_m}\right),$$

$$x \in (\theta_m, \theta_{m+1}].$$

$E[\mathbf{G}_j(x)\mathbf{G}_j(y)^\top] = \min(x, y)\mathbf{V}_j$,
$\mathbf{G}_j(x) = \mathbf{L}_j \mathbf{W}(x), \quad \mathbf{L}_j \mathbf{L}_j^\top = \mathbf{V}_j$,
$\mathbf{W} = [W_1, W_2, \ldots, W_K]^\top$,
K independent standard Wiener processes.

Nonparametric functional approach

Motivation: the factor model can be rewritten as

$$X_i(t) = \tau_i(t) + \eta_i(t), \quad 1 \le i \le N,$$

where

$$\tau_i(t) = \sum_{k=1}^{K} \mu_{i,k} f_k(t), \qquad \eta_i(t) = \sum_{k=1}^{K} b_{ik} f_k(t) + \varepsilon_i(t).$$

The first line is a valid model without the second line.

We can test

$$H_0: \quad \tau_1 = \tau_2 = \ldots = \tau_N.$$

A version of the **Eigen** method can be derived,

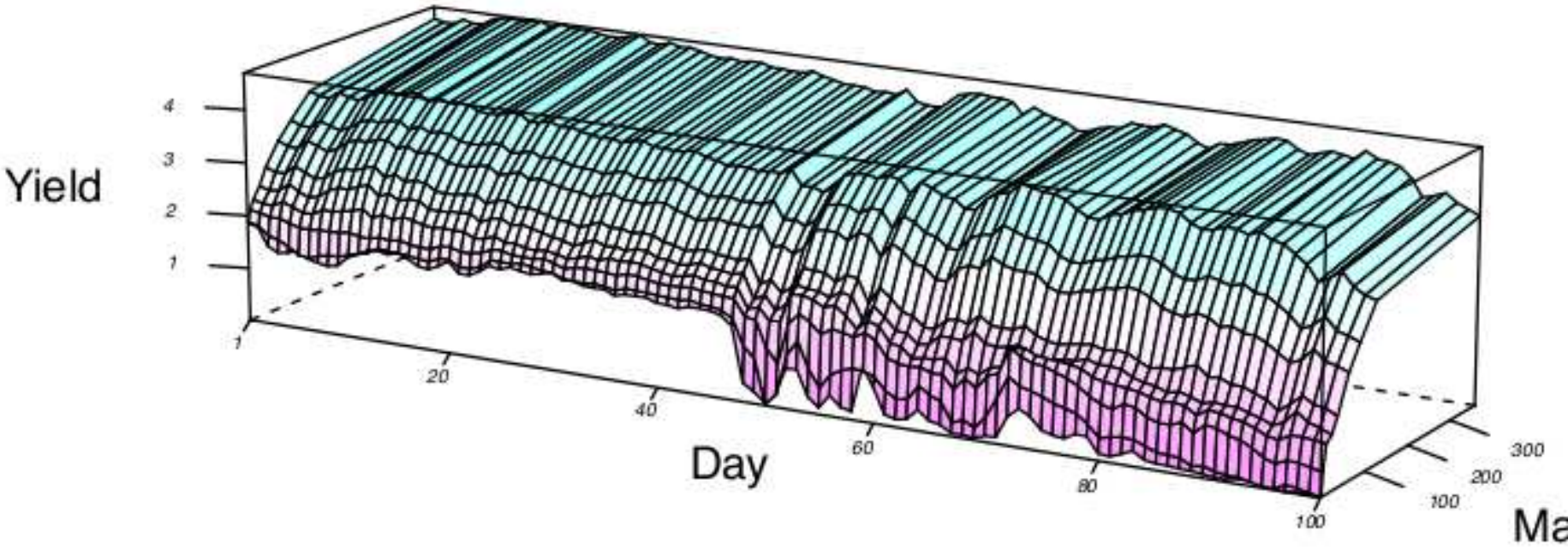
It has to be custom coded, one needs to solve

$$\iint U^0(x, y; t, s)\varphi_j(y, s)dyds = \lambda_j \varphi_j(x, t).$$

The 4D kernel is expressed in terms suitably defined LRV kernels.

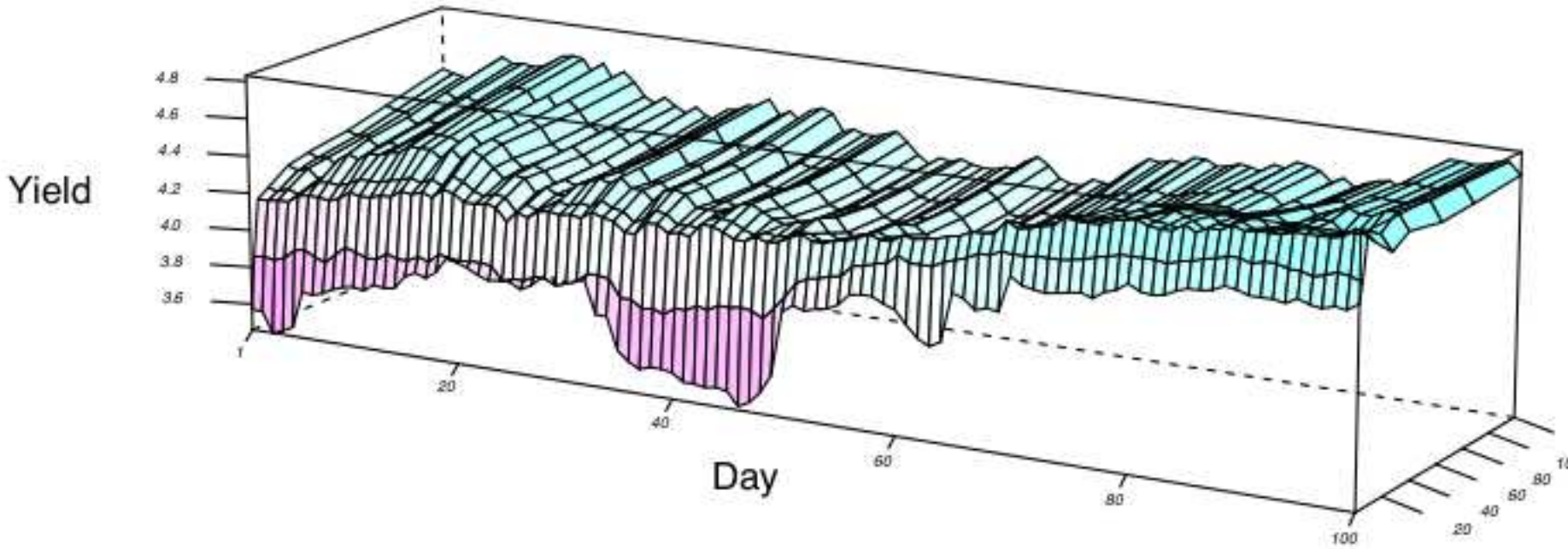
Application to yield curves

Case (1)



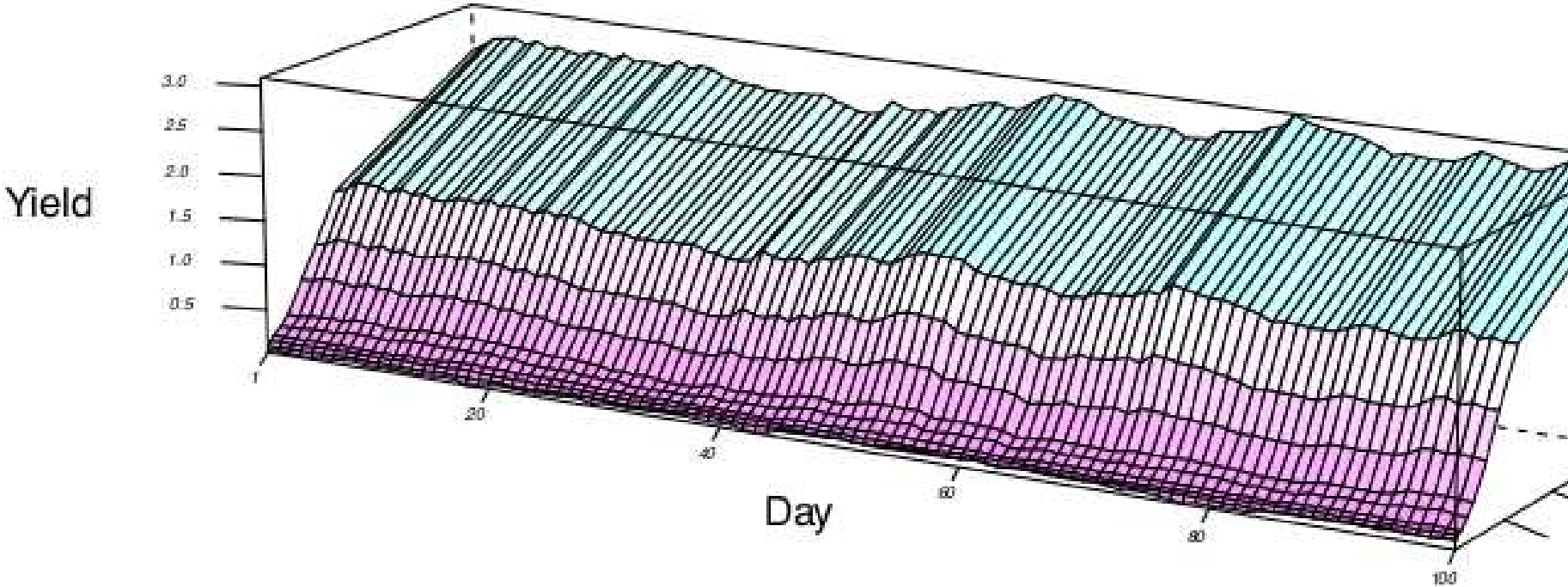
Application to yield curves

Case (2)



Application to yield curves

Case (3)



Application to yield curves

Sampling Period	Method	Break Point	P–value
(1) 03/20/2008 – 03/19/2009	ProjSim	yes	1.5%
	ProjEigen	yes	1.7%
	NFEigen	yes	0.1%
	ProjSim	no	87.9%
	ProjEigen	no	85.2%
	NFEigen	no	26.2%
(2) 06/30/2005 - 06/29/2006	ProjSim	yes	0.1%
	ProjEigen	yes	0.0%
	NFEigen	yes	0.2%
	ProjSim	no	56.7%
	ProjEigen	no	50.5%
	NFEigen	no	57.2%
(3) 02/16/2012 – 02/14/2013	ProjSim	yes	68.1%
	ProjEigen	yes	66.9%
	NFEigen	yes	55.8%
	ProjSim	no	80.4%
	ProjEigen	no	77.7%
	NFEigen	no	76.3%

Conclusions

Conclusions based on similar data examples and simulations:

1. If a possible break point in the error structure is not taken into account in any of the testing procedures, an existing change point in the mean structure may not to be detected.
2. Break points can be misplaced, ± 50 days is fine.
3. The tests are generally well calibrated if $N = 500$. If $N = 250$, all tests have a tendency to overreject by some 2 percent.
4. If the intelligible factors (Lengwiller and Lenz (2010) are used, the empirical size of projSim test improves at the 5% nominal level, but the size of the ProjEigen test deteriorates.
5. We recommend ProjSim with intelligible factors and NFEigen.

The paper contains complete asymptotic theory and details of numerical implementation.