Longitudinal Scalar-on-Functions Regression with Application to Tractography Data

Jan Gertheiss

Department of Statistics
LMU Munich

(Joint work with Jeff Goldsmith, Ciprian Crainiceanu & Sonja Greven)

Bristol 2012

Outline

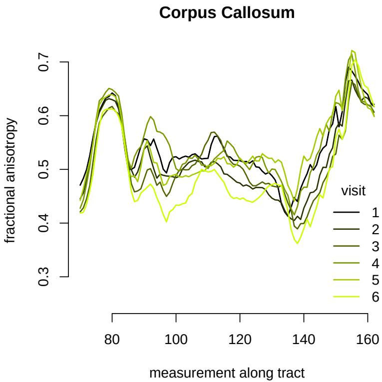
1. Introduction: The Tractography Data
2. Methods
   - Longitudinal Penalized Functional Regression (LPFR)
   - Longitudinal Functional Principal Components Regression (LFPCR)
3. Simulation Studies
4. Application

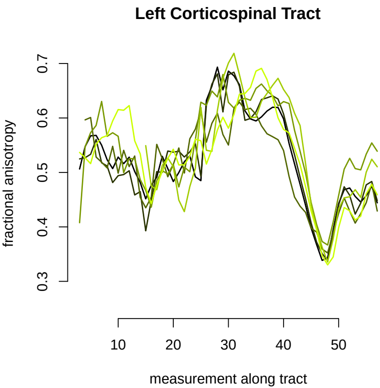
The Tractography Data

Neurological study on disease progression and corresponding changes in diffusion tensor images of the brain in multiple sclerosis (MS) patients.

- MS affects the central nervous system and in particular damages white matter tracts in the brain.
- Diffusion tensor imaging (DTI) allows the extraction of information on individual tracts.
- Different tracts are considered, such as corpus callosum or the corticospinal tract.
- Functional measurements of water diffusivity such as fractional anisotropy or magnetization transfer ratio.
- Relating changes in neuronal tract properties extracted from the diffusion tensor images to disability scores measured at each visit.
- We will use the disability score obtained from the 9 hole peg test.

The Tractography Data





LPFR

Longitudinal Penalized Functional Regression (Goldsmith et al., 2012)

Data of the form $(Y_{ij}, X_{ij1}, \ldots, X_{ijp}, Z_{ij0}, \ldots, Z_{ijg}, W_{ij1}(s), \ldots, W_{ijq}(s))$, where

- Y_{ij} is the response for individual i at visit j ($i = 1, \ldots, n$; $j = 1, \ldots, n_i$);
- $W_{ijm}(s)$ is the functional predictor over domain $\mathcal{D}_m$, $m = 1, \ldots, q$;
- $X_{ij} = (X_{ij1}, \ldots, X_{ijp})^\top$ and $Z_{ij} = (Z_{ij0}, \ldots, Z_{ijg})^\top$ are vectors of additional variables.

For $\mu_{ij} = E(Y_{ij} | b_i, X_{ij}, Z_{ij}, W_{ij1}, \ldots, W_{ijq})$:

$$\mu_{ij} = h(\eta_{ij}) \text{ and } \eta_{ij} = \alpha + \sum_{l=1}^{p} X_{ijl}\beta_l + \sum_{v=0}^{g} Z_{ijv} b_{vi} + \sum_{m=1}^{q} \int_{\mathcal{D}_m} W_{ijm}(s)\gamma_m(s)\, ds,$$

with fixed effects $\beta_1, \ldots, \beta_p$, and iid vectors of random effects $(b_{0i}, \ldots, b_{gi})^\top = b_i \sim N(0, \Gamma)$.

LPFR

Fitting Strategy:

- Dimension reduction/smoothing using simple functional principal components analysis (FPCA) of functional predictor curves $W_{ijm}(s)$.
- Coefficient functions $\gamma_m(s)$ are expressed using a flexible spline basis.
- Use mixed models formulation of penalized splines (see, e.g., Ruppert et al., 2003), e.g.

$$\gamma(s) = \delta_0 + \delta_1 s + \sum_{k=2}^{K} \delta_k (s - \kappa_k)_+,$$

  with random effects $(\delta_2, \ldots, \delta_K)^\top = u \sim N(0, \vartheta^2 I)$.
- The LPFR model can be represented as a generalized linear mixed model.

LPFR

Advantages:

- Random effects b_i account for correlation in the outcomes Y_{ij}.
- Smoothness in the coefficient functions is induced.
- General mixed model software can be used for fitting.
- The mixed model framework allows the construction of confidence intervals for estimated coefficient functions.
- Smoothing parameters that control the shape of the coefficient functions can be automatically estimated by ML or REML.

LPFR

Drawbacks:

Subject specific random intercepts (and slopes) to account for within-subject correlation of outcomes, but

LPFR does not explicitly account for the longitudinal structure of the functional predictor.

- FPCA decomposition used ignores the longitudinal structure of the observations and may miss important sources of variability.
- The term $\int_{\mathcal{D}_m} W_{ijm}(s)\gamma_m(s)\,ds$ does not separate the subject- and visit-level effects of the curve W_{ijm}.
- Separating effects essential if one is interested in the association between individual components of variability and the outcome.
- LPFR in particular problematic when one of the components is not actually associated with the outcome.

LFPCA

Longitudinal Functional Principal Components Analysis (Greven et al., 2010)

For subject i at visit j, measurement $W_{ij}(s)$ at location $s \in \mathcal{D}$ is modeled as

$$W_{ij}(s) = \eta(s, T_{ij}) + B_{i,0}(s) + T_{ij}B_{i,1}(s) + U_{ij}(s) + \varepsilon_{ij}(s).$$

- T_{ij} the time of visit j for subject i,
- $\eta(s, T)$ the overall smooth mean surface,
- mean zero and mutually uncorrelated random processes $B_i(s) = \{B_{i,0}(s), B_{i,1}(s)\}$, $U_{ij}(s)$ and $\varepsilon_{ij}(s)$,
- functional random intercept $B_{i,0}(s)$ and random slope $B_{i,0}(s)$, capturing *between-subject variation*,
- visit-specific functional deviation $U_{ij}(s)$ from the subject-specific functional trend, capturing *within-subject variation*,
- white noise error $\varepsilon_{ij}(s)$.

LFPCA

Longitudinal Functional Principal Components Analysis (Greven et al., 2010)

For subject i at visit j, measurement $W_{ij}(s)$ at location $s \in \mathcal{D}$ is modeled as

$$W_{ij}(s) = \eta(s, T_{ij}) + B_{i,0}(s) + T_{ij} B_{i,1}(s) + U_{ij}(s) + \varepsilon_{ij}(s).$$

- T_{ij} the time of visit j for subject i,
- $\eta(s, T)$ the overall smooth mean surface,
- mean zero and mutually uncorrelated random processes $B_i(s) = \{B_{i,0}(s), B_{i,1}(s)\}$, $U_{ij}(s)$ and $\varepsilon_{ij}(s)$,
- functional random intercept $B_{i,0}(s)$ and random slope $B_{i,0}(s)$, capturing *between-subject variation*,
- visit-specific functional deviation $U_{ij}(s)$ from the subject-specific functional trend, capturing *within-subject variation*,
- white noise error $\varepsilon_{ij}(s)$.

$\Rightarrow$ Decompose functional variation into three parts: *subject-specific variation* $B_i(s)$, *visit-specific variation* $U_{ij}(s)$, measurement error.

LFPCA

Longitudinal Functional Principal Components Analysis (Greven et al., 2010)

Karhunen-Loève expansion of the processes $B_i(s)$ and $U_{ij}(s)$ using eigenfunctions (ϕ_k^0, ϕ_k^1) and ϕ_r^U:

$$B_{i,0}(s) = \sum_{k=1}^{\infty} \xi_{ik}\phi_k^0(s),\; B_{i,1}(s) = \sum_{k=1}^{\infty} \xi_{ik}\phi_k^1(s),\; U_{ij}(s) = \sum_{r=1}^{\infty} \zeta_{ijr}\phi_r^U(s),$$

with principal component scores

$$\xi_{ik} = \int_{\mathcal{D}} B_{i,0}(s)\phi_k^0(s)\,ds + \int_{\mathcal{D}} B_{i,1}(s)\phi_k^1(s)\,ds,\; \zeta_{ijr} = \int_{\mathcal{D}} U_{ij}(s)\phi_r^U(s)\,ds,$$

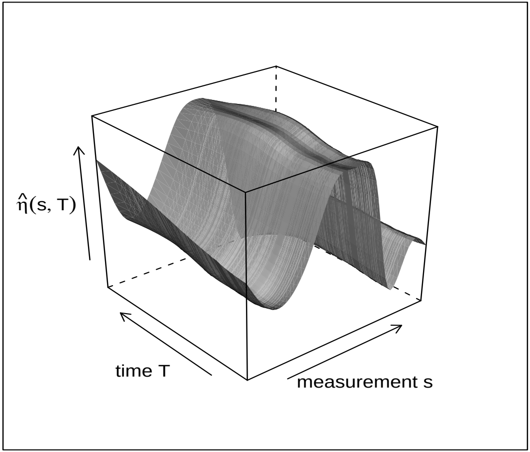
uncorrelated random variables with mean zero and variances λ_k and ν_r.

LFPCA uses a truncated version such that

$$W_{ij}(s) \approx \eta(s, T_{ij}) + \sum_{k=1}^{N_B} \xi_{ik}(\phi_k^0(s) + T_{ij}\phi_k^1(s)) + \sum_{r=1}^{N_U} \zeta_{ijr}\phi_r^U(s) + \varepsilon_{ij}(s).$$

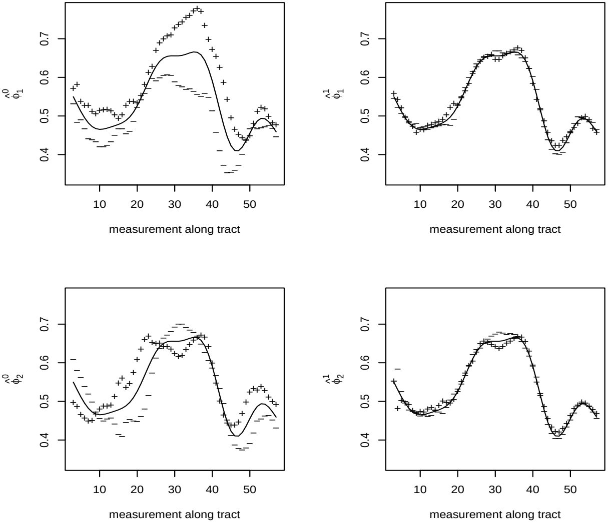
LFPCA

Illustration (mean function)



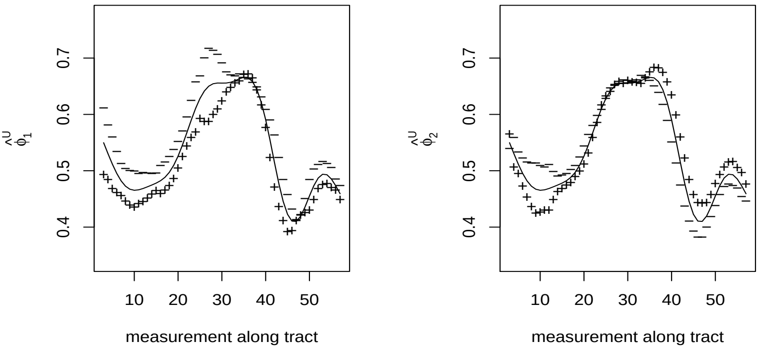
LFPCA

Illustration (B-process eigenfunctions)



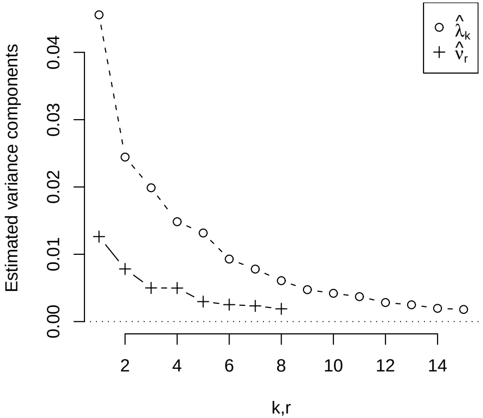
LFPCA

Illustration (U-process eigenfunctions)



LFPCA

Illustration (variance components)



LFPCR

Longitudinal Functional Principal Components Regression (Gertheiss et al., 2012)

Approach 1: Score LFPCR

$$\mu_{ij} = h(\eta_{ij}) \ \text{ and } \ \eta_{ij} = \alpha(T_{ij}) + b_i + \sum_{l=1}^{p} \beta_l X_{ijl} + \sum_{k=1}^{N_B} \theta_k \xi_{ik} + \sum_{r=1}^{N_U} \delta_r \zeta_{ijr}$$

- Y_{ij} is regressed on the scores ξ_{ik} and ζ_{ijr}.
- Since scores in the LFPCA model only refer to deviations from the mean surface $\eta(s, T_{ij})$, we include a time-varying intercept $\int_{\mathcal{D}} \varphi(s)\eta(s, T_{ij})\, ds = \alpha(T_{ij})$.
- We assume $b_i \sim N(0, \tau^2)$, additional random effects may be added as done with LPFR.
- Additional functional predictors result in additional LFPCA scores and can thus be easily included.
- Estimation via the generalized additive mixed models framework.

LFPCR

Longitudinal Functional Principal Components Regression (Gertheiss et al., 2012)

Approach 2: Smooth LFPCR

Reconstruct between-subject variation $B_i(s, T_{ij}) = B_{i,0}(s) + T_{ij} B_{i,1}(s)$ and within-subject variation $U_{ij}(s)$.

- $B_i(s, T_{ij})$ represents the systematic trend in subject i over time.
- $U_{ij}(s)$ denotes visit-specific deviations from this trend.

Both parts may be important as predictors.

- $B_i(s, T_{ij})$ may be more relevant if $U_{ij}(s)$ constitutes mostly measurement error.
- $U_{ij}(s)$ might be the more important component if curves that are *unusual* for subject i are highly predictive for the outcome Y_{ij}.

LFPCR

Longitudinal Functional Principal Components Regression (Gertheiss et al., 2012)

Functional covariates $B_i(s, T_{ij})$ and $U_{ij}(s)$ can now be used in a functional regression model:

$$\mu_{ij} = h(\eta_{ij}) \text{ and } \eta_{ij} = \alpha(T_{ij}) + b_i + \int_{\mathcal{D}} \gamma_B(s) B_i(s, T_{ij}) ds + \int_{\mathcal{D}} \gamma_U(s) U_{ij}(s) ds,$$

with $B_i(s, T_{ij}) = B_{i,0}(s) + T_{ij} B_{i,1}(s)$ and $B_{i,0}(s) = \sum_{k=1}^{N_B} \xi_{ik} \phi_k^0(s)$, $B_{i,1}(s) = \sum_{k=1}^{N_B} \xi_{ik} \phi_k^1(s)$, $U_{ij}(s) = \sum_{r=1}^{N_U} \zeta_{ijr} \phi_r^U(s)$.

- Additional functional predictors can be included as additional B- and U-processes resulting from LFPCA of these curves.
- Coefficient functions γ_B and γ_U are estimated using penalized spline expansions with tuning parameters estimated via ML or REML.

LFPCR

Comparing Score and Smooth LFPCR

$$
\begin{aligned}
\sum_{k=1}^{N_B} \theta_k \xi_{ik} + \sum_{r=1}^{N_U} \delta_r \zeta_{ijr} &= \sum_{k=1}^{N_B} \theta_k \left(\int_{\mathcal{D}} B_{i,0}(s) \phi_k^0(s) ds + \int_{\mathcal{D}} B_{i,1}(s) \phi_k^1(s) ds \right) \\
&+ \sum_{r=1}^{N_U} \delta_r \int_{\mathcal{D}} U_{ij}(s) \phi_r^U(s) ds \\
&= \int_{\mathcal{D}} B_{i,0}(s) \sum_{k=1}^{N_B} \theta_k \phi_k^0(s) ds + \int_{\mathcal{D}} B_{i,1}(s) \sum_{k=1}^{N_B} \theta_k \phi_k^1(s) ds \\
&+ \int_{\mathcal{D}} U_{ij}(s) \sum_{r=1}^{N_U} \delta_r \phi_r^U(s) ds.
\end{aligned}
$$

- Functional linear model with predictors $B_{i,0}(s)$, $B_{i,1}(s)$ and $U_{ij}(s)$.
- Coefficient functions restricted to spaces spanned by the first eigenfunctions.
- Alternative: estimate these coefficient functions directly. (third way to do LFPCR)

Assume that the LPFR model and the LFPCA decomposition hold.

$$\Rightarrow \int_{\mathcal{D}} W_{ij}(s)\gamma(s)ds = \alpha(T_{ij}) + \int_{\mathcal{D}} \gamma(s)B_i(s, T_{ij})ds + \int_{\mathcal{D}} \gamma(s)U_{ij}(s)ds + \tilde{\varepsilon}_{ij},$$

where $\tilde{\varepsilon}_{ij}$ is noise with mean zero and $\alpha(T_{ij}) = \int_{\mathcal{D}} \gamma(s)\eta(s, T_{ij})ds$.

- If $W_{ij}(s)$ is smooth, LPFR can be seen as a special case of LFPCR where $\gamma_B(s) = \gamma_U(s) = \gamma(s)$.
- If the LPFR model holds, and the LFPCA model is a reasonable approximation to the (functional) data generating process, smooth LFPCR will also be an adequate modeling approach.
- If the LFPCR model is correct and $\eta(s, T_{ij})$ is not relevant for the response, or $\gamma_B(s) \neq \gamma_U(s)$, LFPCR will outperform LPFR.

Simulation Studies

Setup

- Scenario 1: LPFR model true, predictors generated according to the LFPCA model.
- Scenario 2: LPFR model true, predictors directly simulated.

- Scenario 3: Score LFPCR true.
- Scenario 4: Smooth LFPCR true, only U-process relevant.
- Scenario 5: Smooth LFPCR true, only B-process relevant.

Simulation Studies

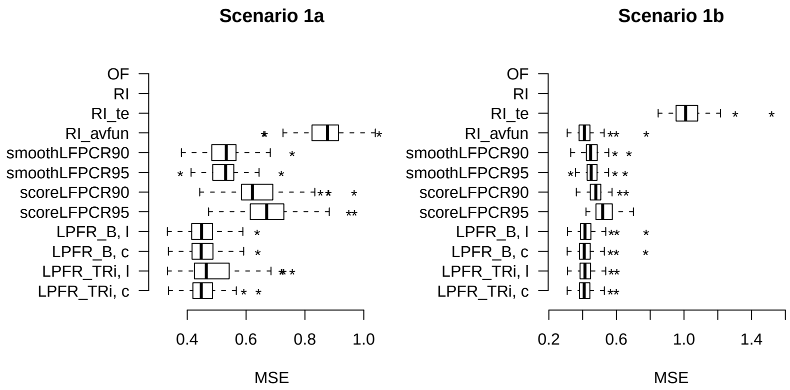
Setup

We consider the MSE $\frac{1}{n}\sum_{i,j}(\mu_{ij} - \hat{\mu}_{ij})^2$ and compare

- Score/Smooth LFPCR, each with 90% and 95% of the functional covariates' variance explained.
- LPFR with truncated power basis or with B-splines.
- Simple benchmark methods:
  - a saturated model with massive overfitting (OF), i.e., $\hat{\mu}_{ij} = y_{ij}$,
  - a simple random intercept model without any covariates (RI),
  - a random intercept model without covariates but including a smooth trend function $f(T_{ij})$, i.e., a model with time-effect (RI_te),
  - a random intercept model (with smooth time-trend) where each functional covariate is simply averaged (RI_avfun).

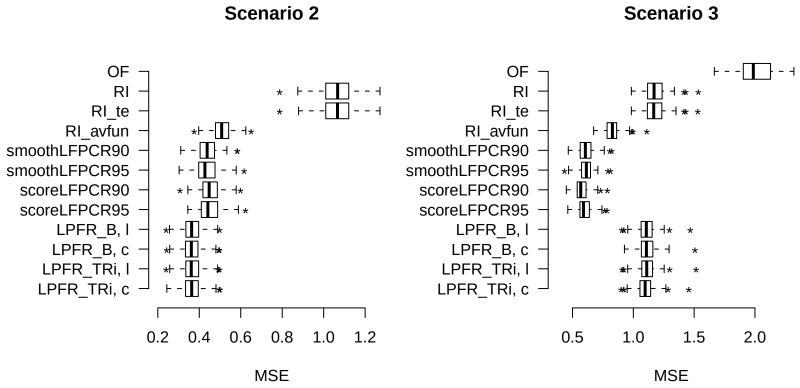
Simulation Studies

Results



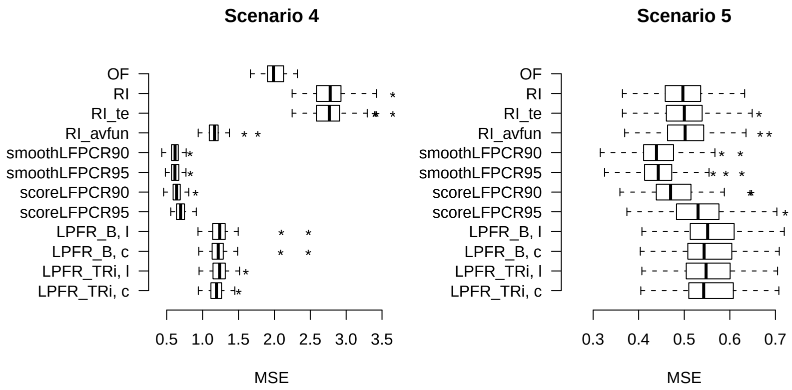
Simulation Studies

Results



Simulation Studies

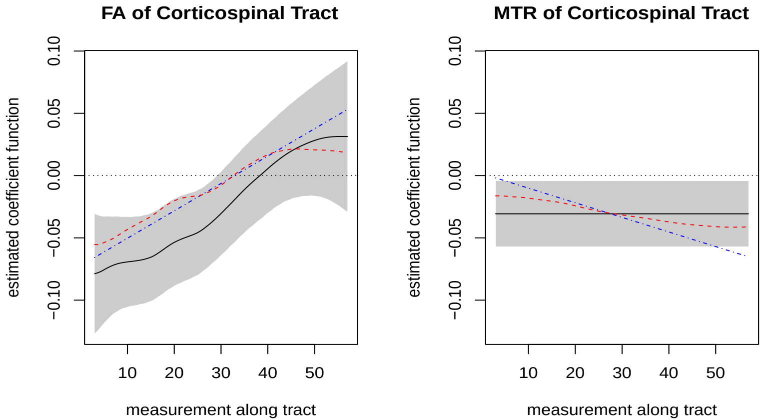
Results



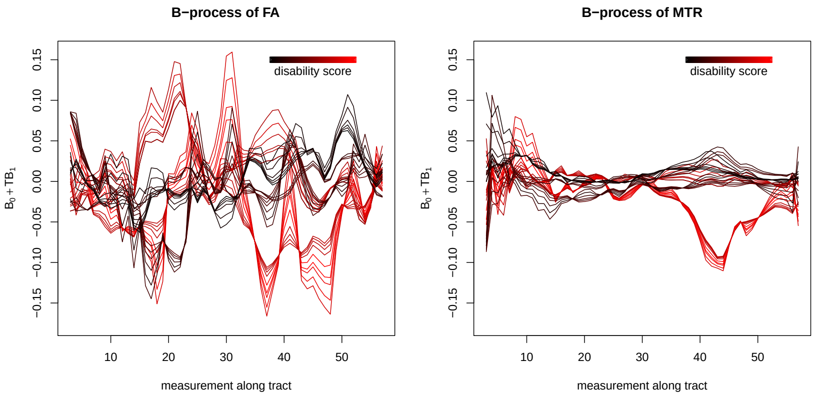
Application to Tractography Data

- Response: disability score obtained from the 9 hole peg test (*peg9*).
- We only found a clear dependence between measures along the corticospinal tract and *peg9*.
- Fractional anisotropy (FA) and the magnetization transfer ratio (MTR) along the corticospinal tract as functional predictors for *peg9*.
- Additional scalar covariates: sex and age, and a dummy variable indicating whether it's the patient's first visit or not.
- For modeling, we use LPFR and smooth LFPCR, both with Gamma response distribution and log-link.

LPFR

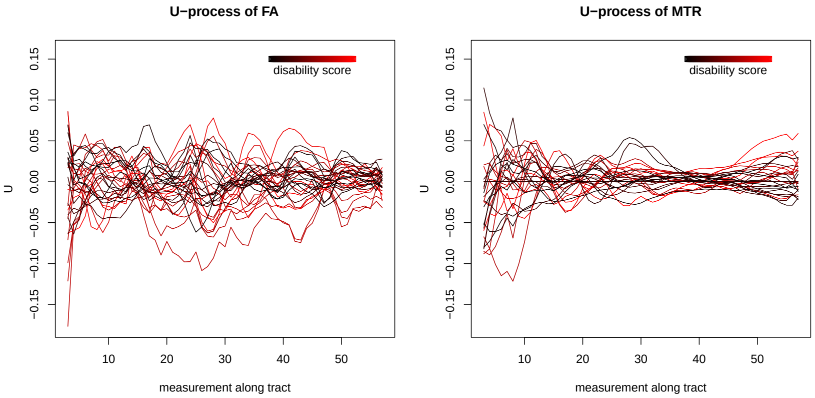

FA of Corticospinal Tract
MTR of Corticospinal Tract
estimated coefficient function
−0.10
−0.05
0.00
0.05
0.10
10
20
30
40
50
measurement along tract

LFPCR

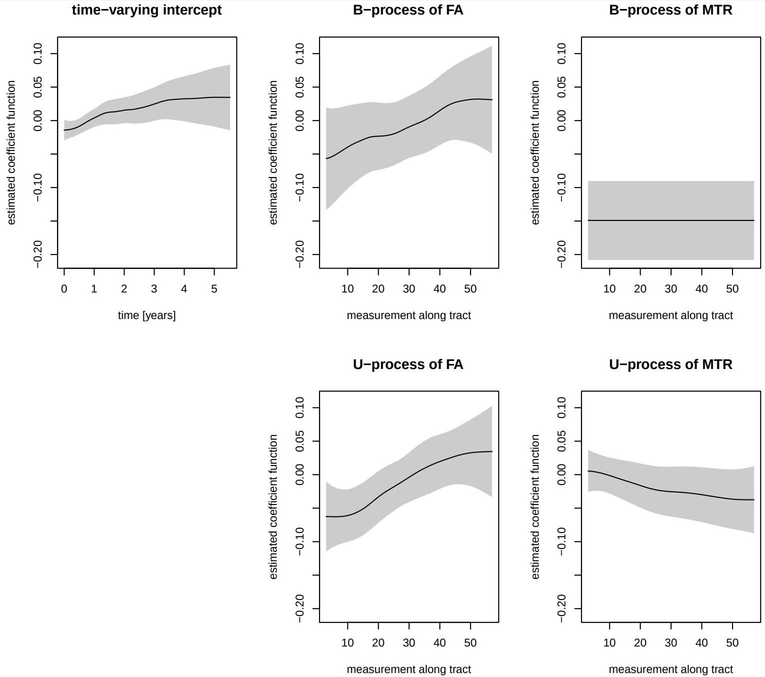


[Illustration for the first 5 patients]

LFPCR



[Illustration for the first 5 patients]



time−varying intercept
B−process of FA
B−process of MTR
U−process of FA
U−process of MTR
estimated coefficient function
−0.20
−0.10
0.00
0.05
0.10
time [years]
0
1
2
3
4
5
measurement along tract
10
20
30
40
50

Summary and Discussion

- We presented and compared different tools for scalar-on-function regression when observations are taken repeatedly over time.
- We proposed two novel versions of principal components regression for longitudinal functional data: score LFPCR and smooth LFPCR.
- LFPCR separates the influence of subject- and visit-specific variation in the functional predictors.
- Smooth LFPCR tends to perform better, and it yields nice interpretations.
- (Smooth) LFPCR is highly competitive to mixed models that use functional covariates directly (LPFR).
- Score LFPCR heavily depends on the number of principal components N_B and N_U.
- Thanks to implicit regularization when fitting coefficient functions, smooth LFPCR is robust against different choices of N_B and N_U.

References

Gertheiss, J., Goldsmith, J., Crainiceanu, C. and Greven, S. (2012). Longitudinal Scalar-on-Functions Regression with Application to Tractography Data. *Preprint*.

Goldsmith, J., Crainiceanu, C., Caffo, B. and Reich, D. (2012). Longitudinal Penalized Functional Regression for Cognitive Outcomes on Neuronal Tract Measurements. *Journal of the Royal Statistical Society C (Applied Statistics)* 61, 453–469.

Greven, S., Crainiceanu, C., Caffo, B. and Reich, D. (2010). Longitudinal Functional Principal Components Analysis. *Electronic Journal of Statistics* 4, 1022–1054.

Ruppert, D., Wand, M.P. and Carroll, R.J. (2003). *Semiparametric Regression*. Cambridge: Cambridge University Press.

THANK YOU!