Dynamic Prediction with fPCA

2022-11-3

Dynamic prediction

- With observations up to t_m, predict outcomes (or probabilities of outcome) after that time point
- Prediction updates with new observations

Functional Concurrent Regression (FCR)

- Goal: to predict future track based on observed track
- For a subject i, we observe a function over t

$$Y_i(t) = f_0(t) + b_i(t) + \epsilon_i(t)$$

- We usually observe Y_i on a series of discrete t_{ij}

$$Y_{ij} = f_0(t_{ij}) + b_i(t_{ij}) + \epsilon_{ij}, \quad j = 1...J_i$$

where $\epsilon_{ij} \sim N(0, \sigma_{\epsilon^2})$.

- Subject-specific random effect

$$b_i(t) = \Sigma_{k=1}^{c} u_{ik} B_k(t)$$

where $\boldsymbol{u_i} \sim N(0, \Gamma)$

Dynamic prediction

- When there is no covariate in the model, this is essentially a fPCA problem.
  - $\boldsymbol{B}$ is a matrix of eigenfunctions
  - $\boldsymbol{u}$ is a matrix of PC scores/loadings
- Use fPCA to estimate f_0, Γ and σ_ϵ
- For a new subject with observations up to t_m, estimate its score:

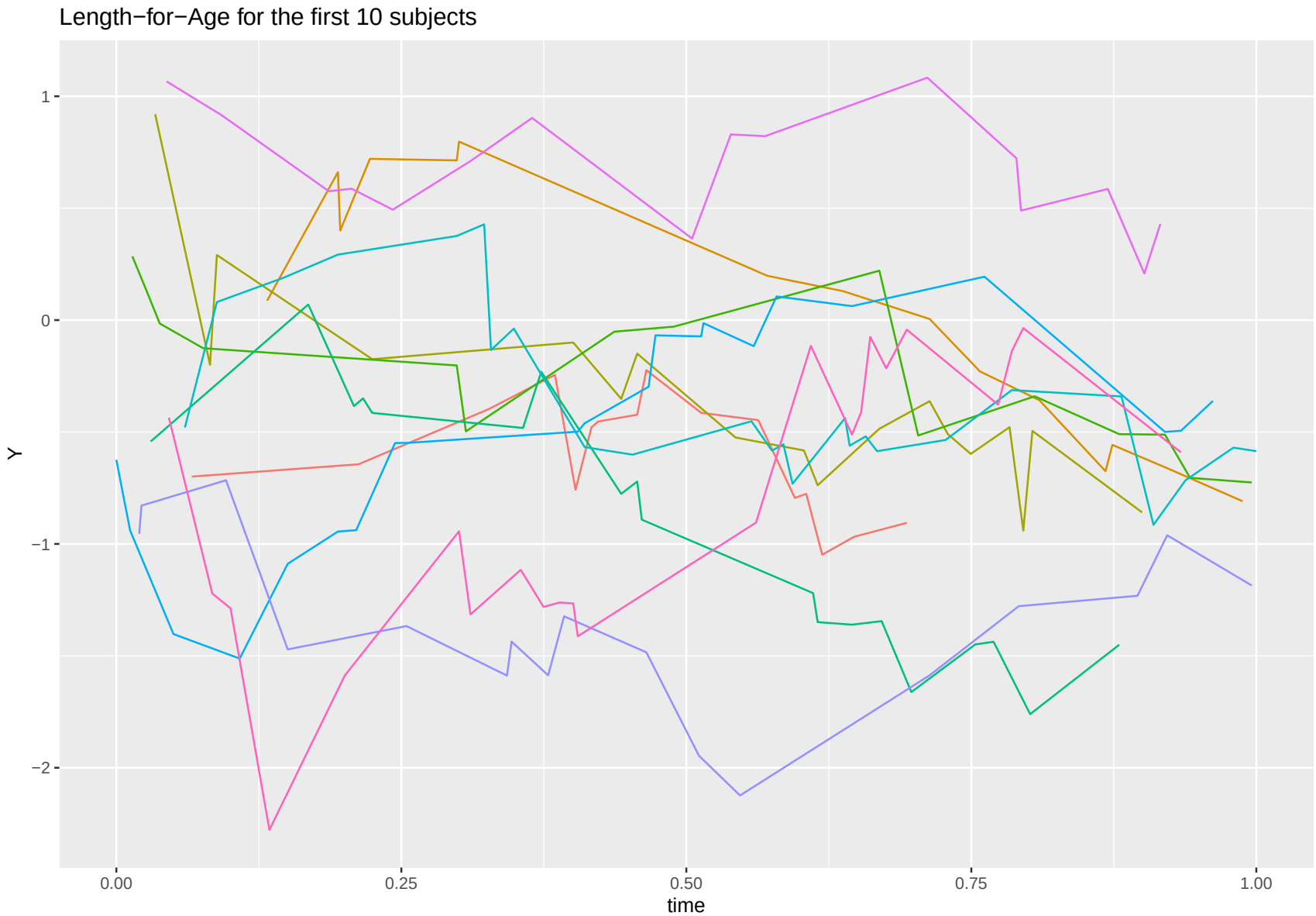
$$\hat{\boldsymbol{u}} = E(\boldsymbol{u}|\boldsymbol{y}) = \hat{\boldsymbol{\Gamma}}\boldsymbol{B}^{\boldsymbol{T}}(\boldsymbol{B}\hat{\boldsymbol{\Gamma}}\boldsymbol{B}^{\boldsymbol{T}} + \hat{\sigma_\epsilon}^2\boldsymbol{I_m})$$

With the estimated score, we can predict its outcome in following time points

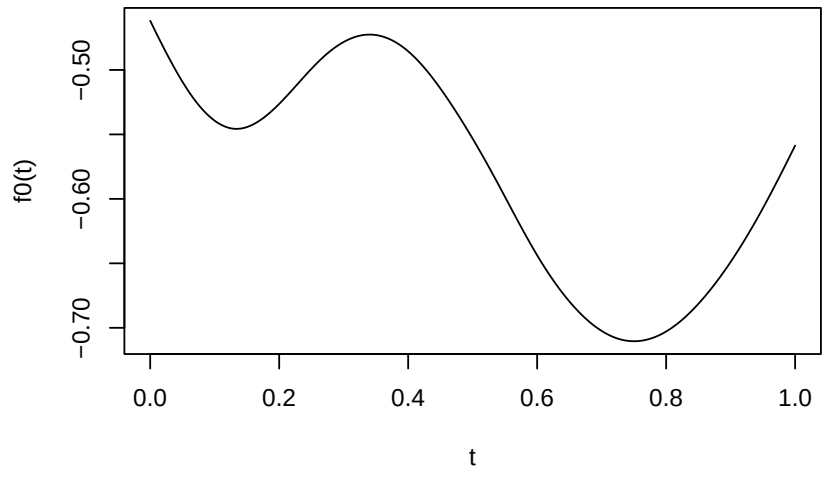
$$\hat{\boldsymbol{Y}} = \hat{\boldsymbol{f}}_0 + \boldsymbol{B}^T\hat{\boldsymbol{u}}$$

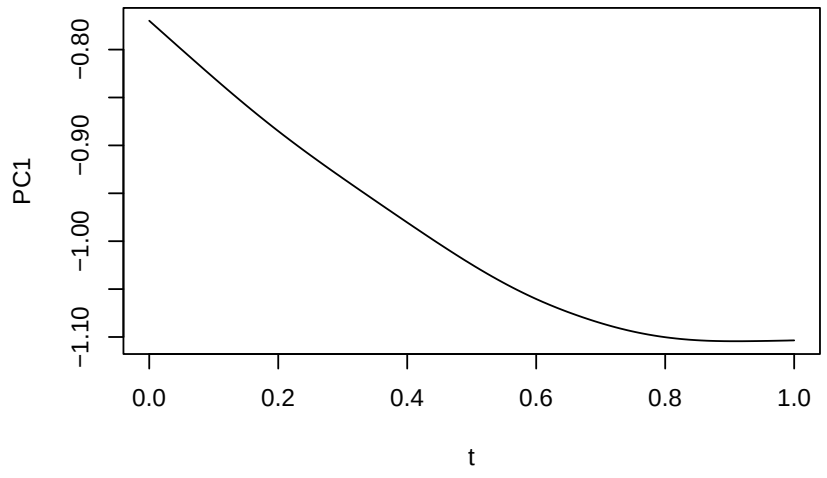
Simulated child growth data

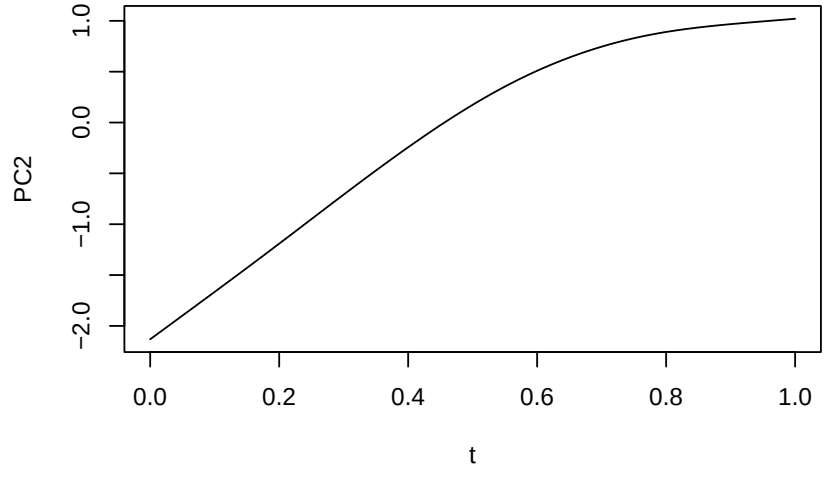
- Predict length-for-age, observed with noise

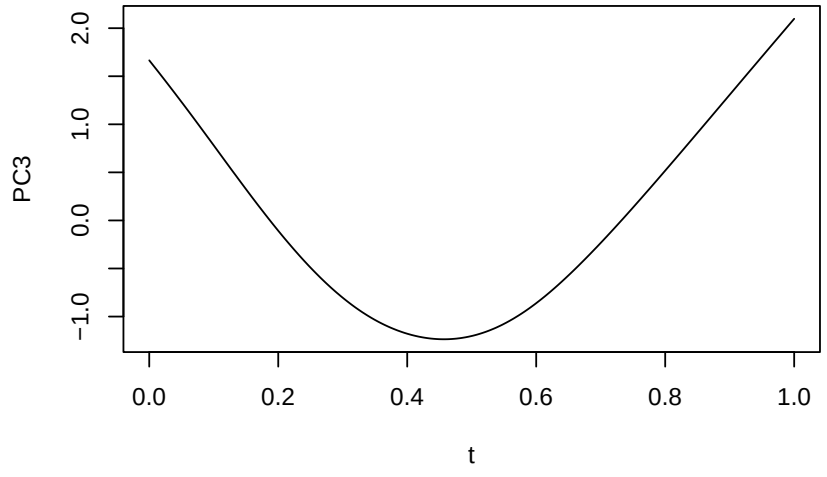


FPCA on observed LAZ

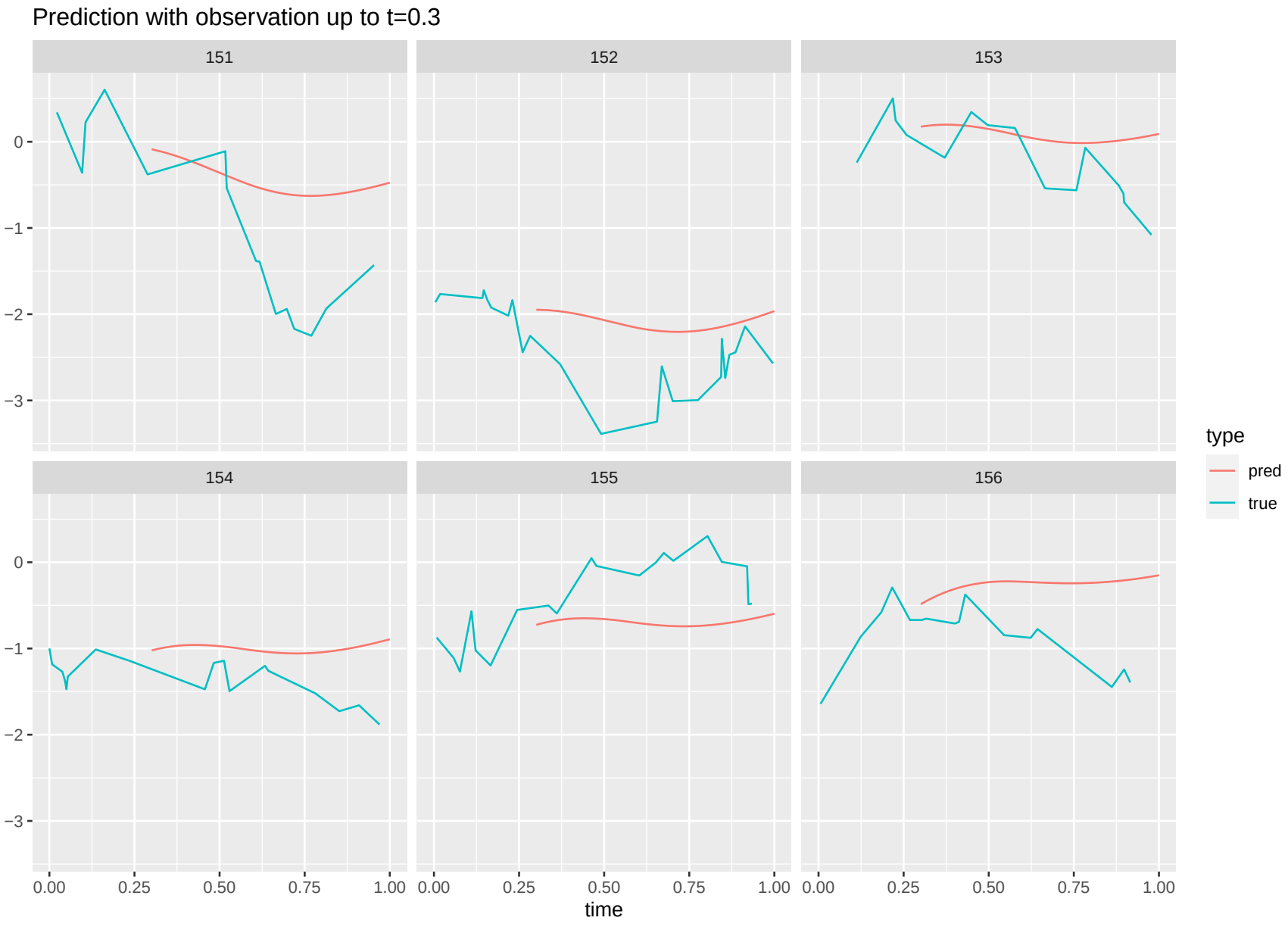




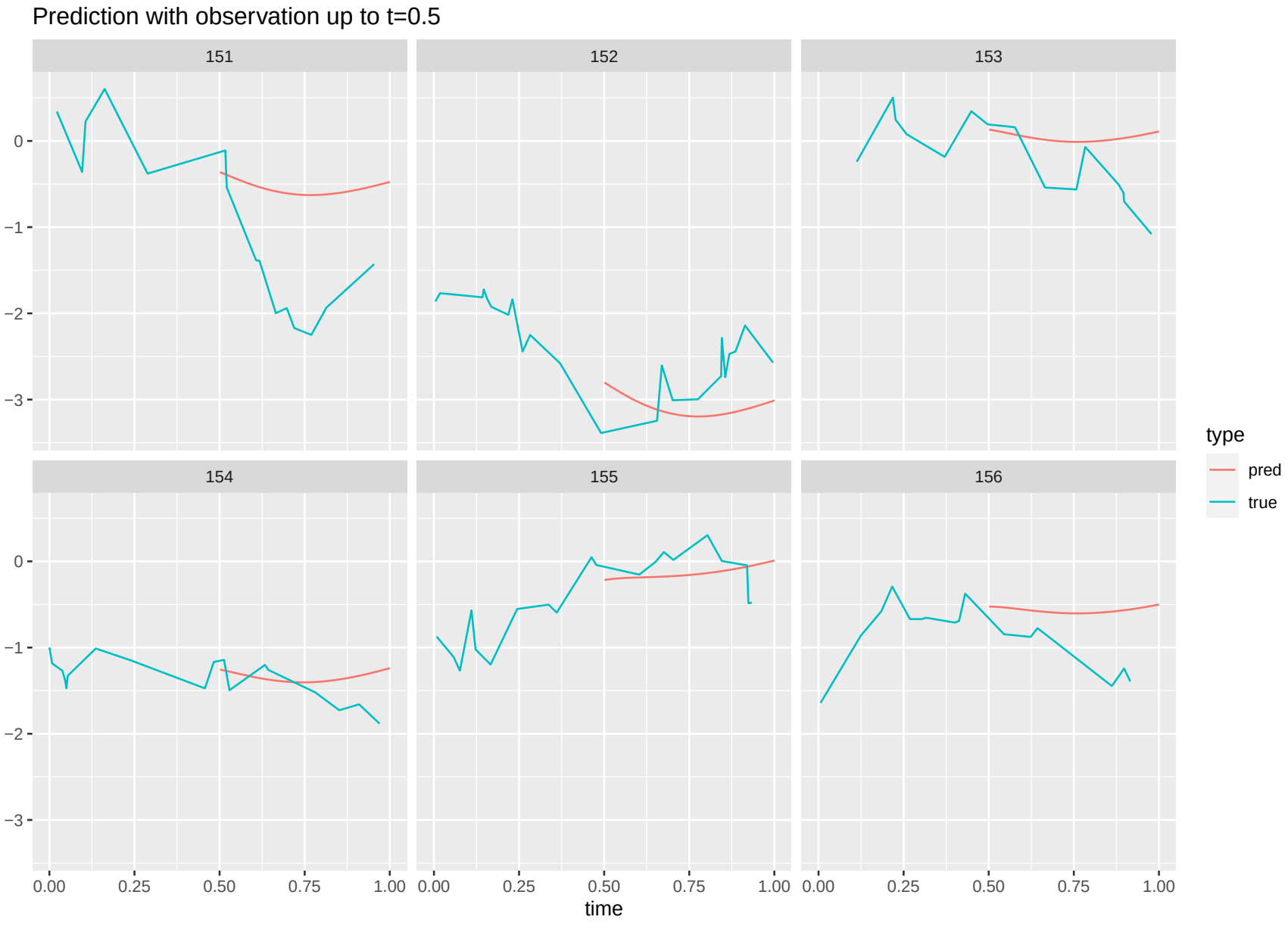




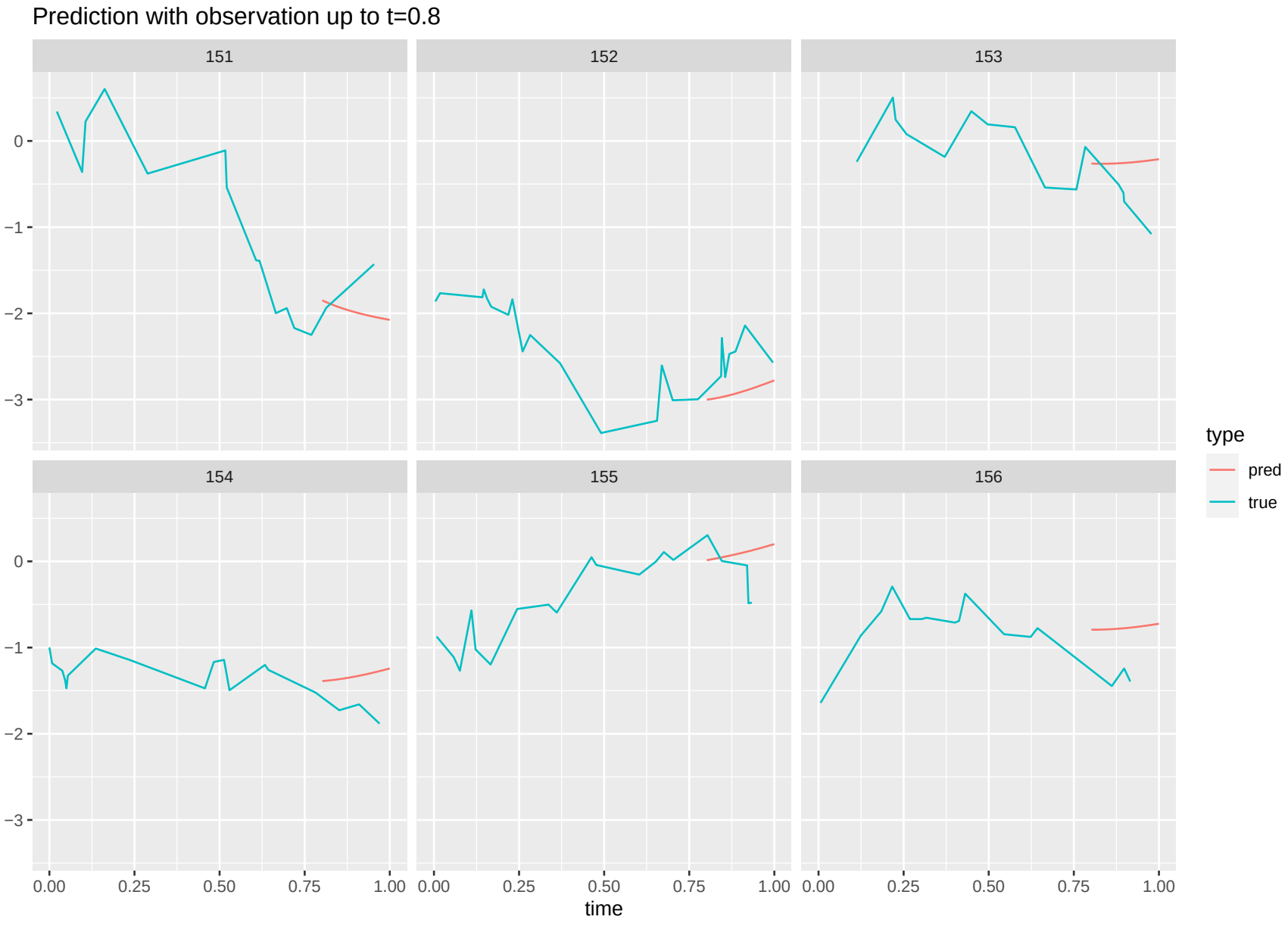
Prediction of new partially observed sample



Prediction of new partially observed sample



Prediction of new partially observed sample



Next steps

- Establish interval prediction
- ? fPCA on pooled prediction, instead of original functions observed with noise
- Extension to exponential family data