Evaluating Temporal Dynamics using Gaussian Processes in Longitudinal Metabolomics Data

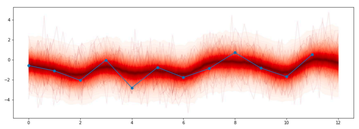
Allen Ross, Ali Rahnavard

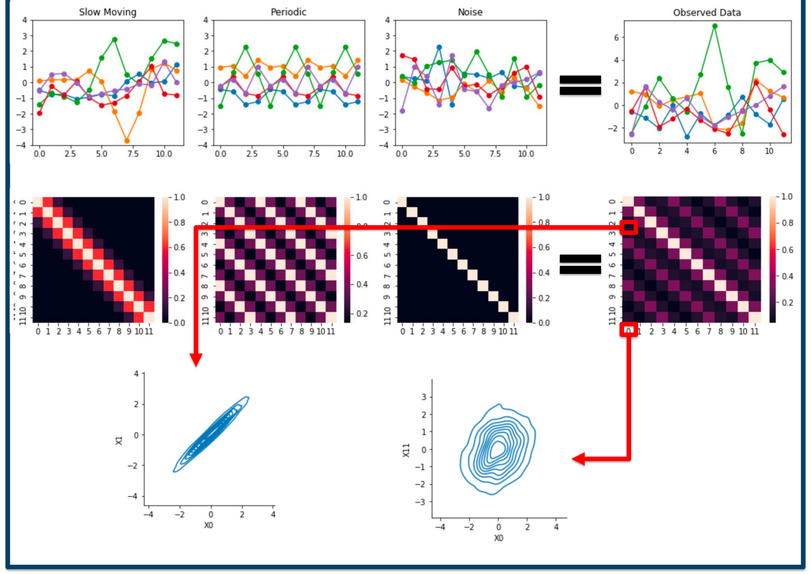
1 Introduction

We are interested in uncovering plausible **functional representations** of temporal dynamics given observed longitudinal omics data. These representations can then be investigated for biologically meaningful **component characteristics** as well as **contributions**. Potential applications include omics normalization and predictive modeling.

2 Gaussian Processes

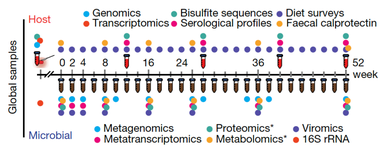
$$y \sim MVN(g(\mu), \Sigma)$$
$$g(\mu) = 0, \Sigma = K(x, x')$$

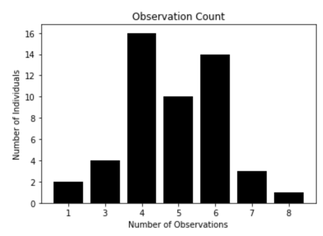


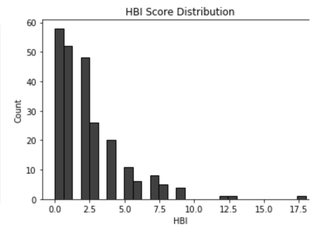


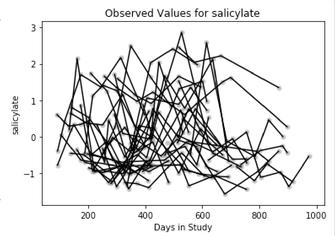
3 iHMP Data Subset Overview

241 samples
50 individuals
100 metabolites



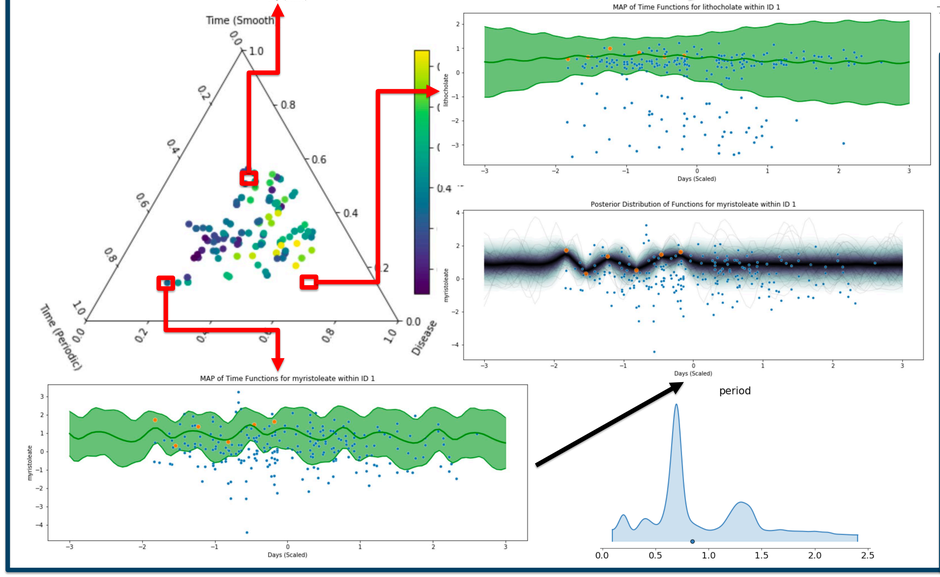


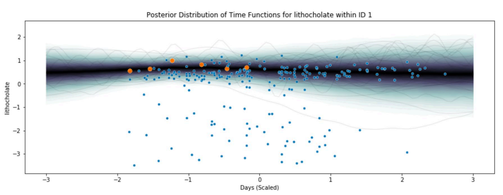


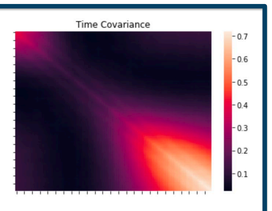


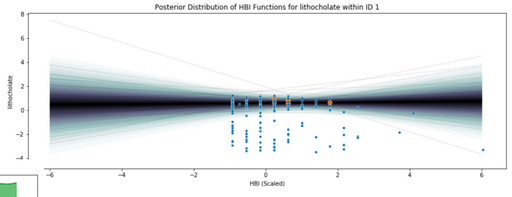
4 Metabolite Gaussian Processes

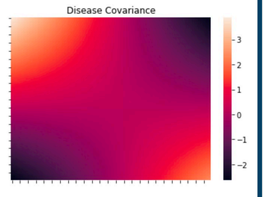
$$K(x, x') = B\left[x_i, x_i'^T\right] \begin{pmatrix} \sigma_s^2 EQ(x_t, x_t') \\ +\sigma_p^2 Per(x_t, x_t') \\ +\sigma_d^2 Poly(x_d, x_d') \end{pmatrix} + \sigma_n^2$$











5 Conclusions

- Nontrivial covariance relationships detected
- Posterior sampling is important but difficult

Next Steps

- Kernel learning
- Clustering
- Multi-omics
- Temporal interrelationship structures
- Approximate inference

6 Software

This work will be incorporated into *Waveome*, a software library that analyzes longitudinal omics data. More information including documentation, manuals, and tutorials will be available at waveome as resources are updated.

Data Source

Inflammatory Bowel Disease Multi'omics Database

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC