

SILAS

Statistical Inference on Latent Structures from Sequential Data

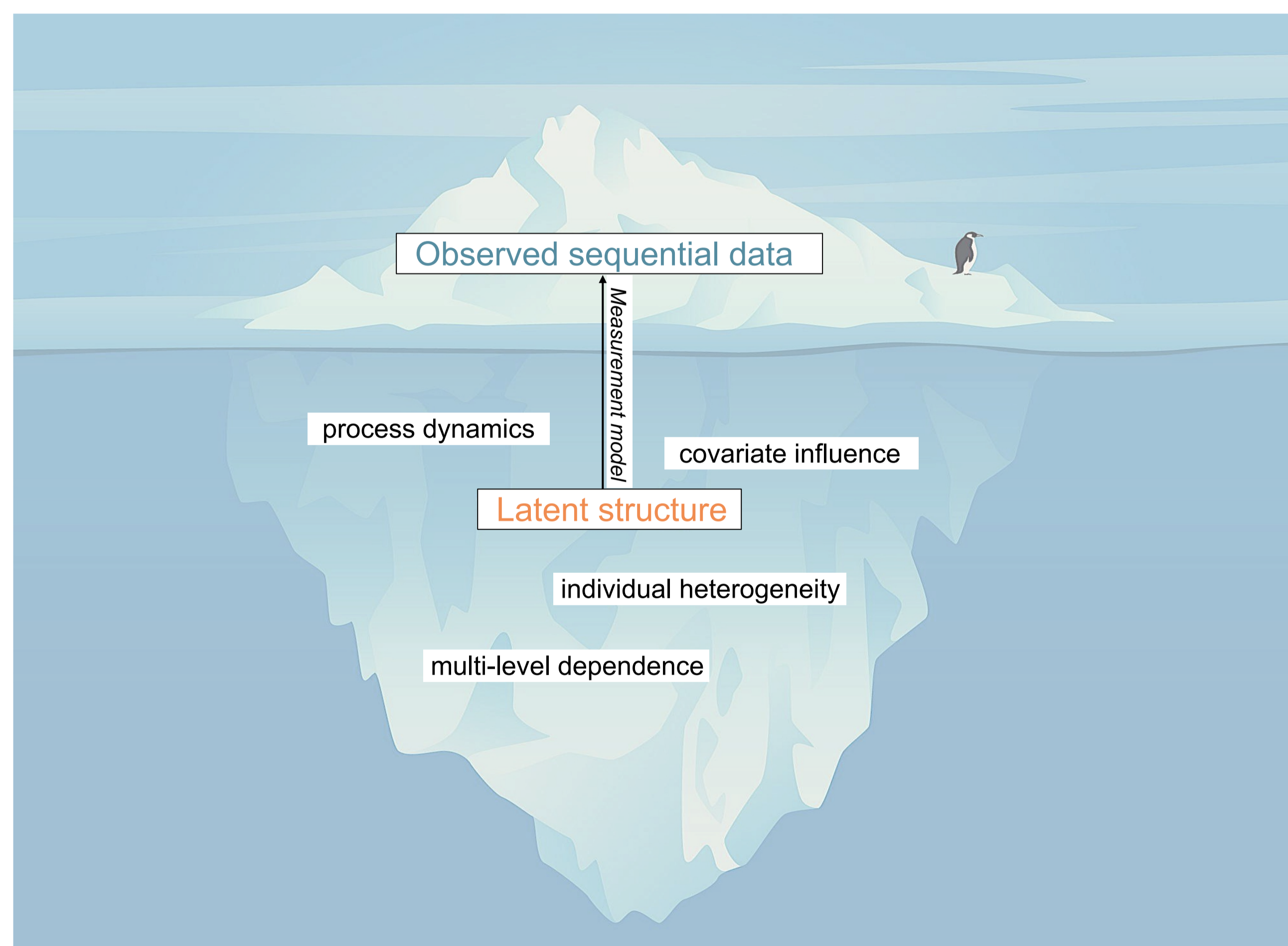
Summary

Sequential data enable valuable inference on underlying latent structures:

- animal behaviour can be decoded from movement measurements
- electronic health records allow the monitoring of disease progression
- customers' sentiments can be inferred from their interactions with companies
- social media activities may reveal political attitudes
- traits, emotions, and cognitions can be uncovered using psychological questionnaires

Novel types of data driven by digitalisation enable us to use sophisticated statistical modelling to obtain insights into very comprehensive latent structures (cf. Figure below).

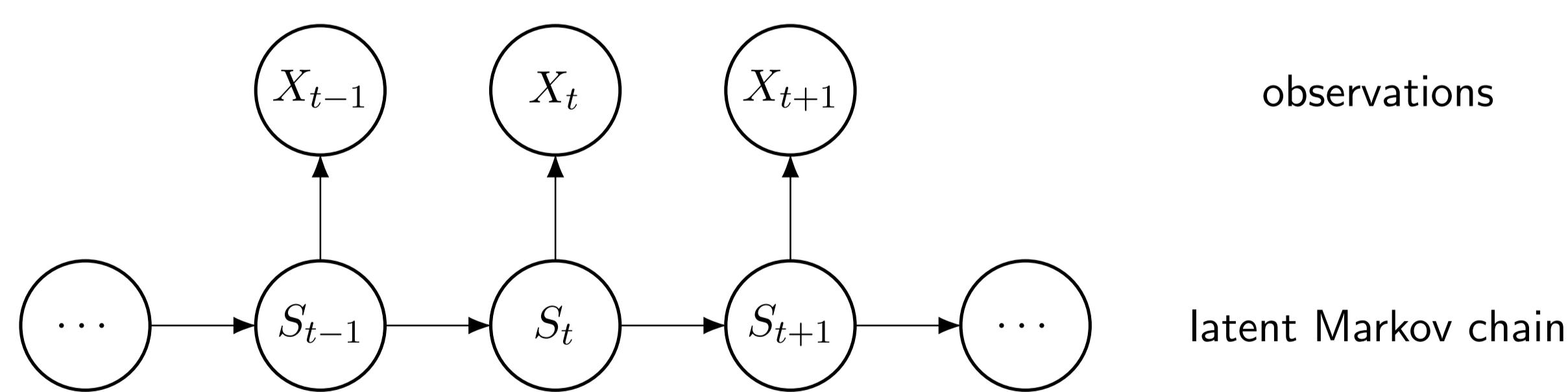
Method development for inference on latent structures from sequential data has lagged behind the very rapid increase in the volume and the complexity of new types of data.



Methods & models

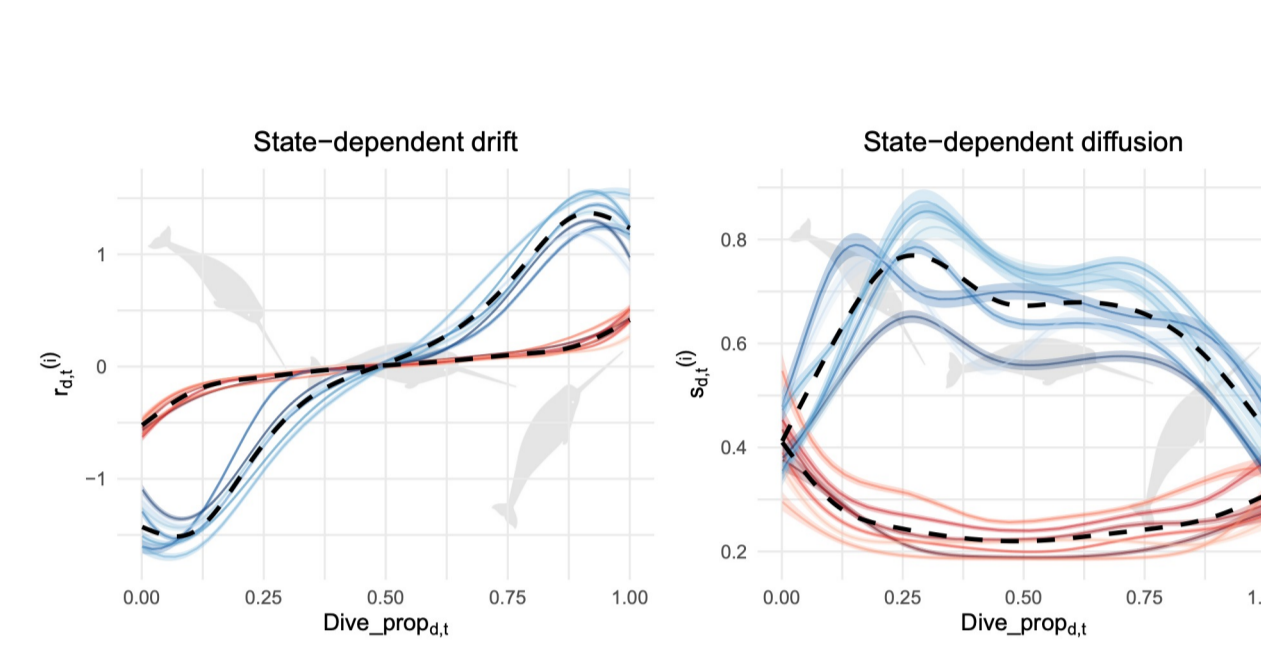
Within SILAS, there is methodological expertise on the following models:

- hidden Markov models (HMMs)
- Markov-modulated Poisson processes (MMPPs)
- state-space models (SSMs)
- stochastic differential equations (SDEs)
- structural equation models (SEMs)

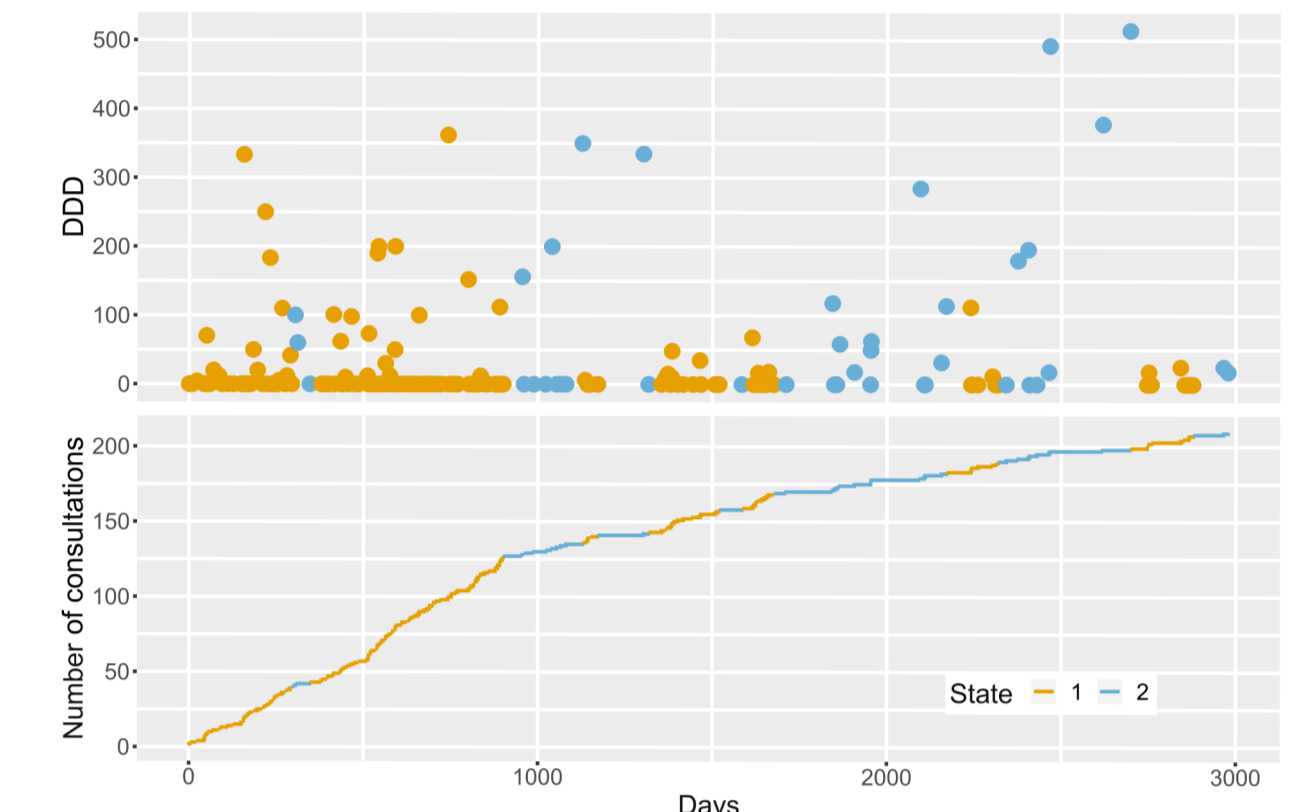


Example statistical model (HMM) for learning latent structures from sequential data.

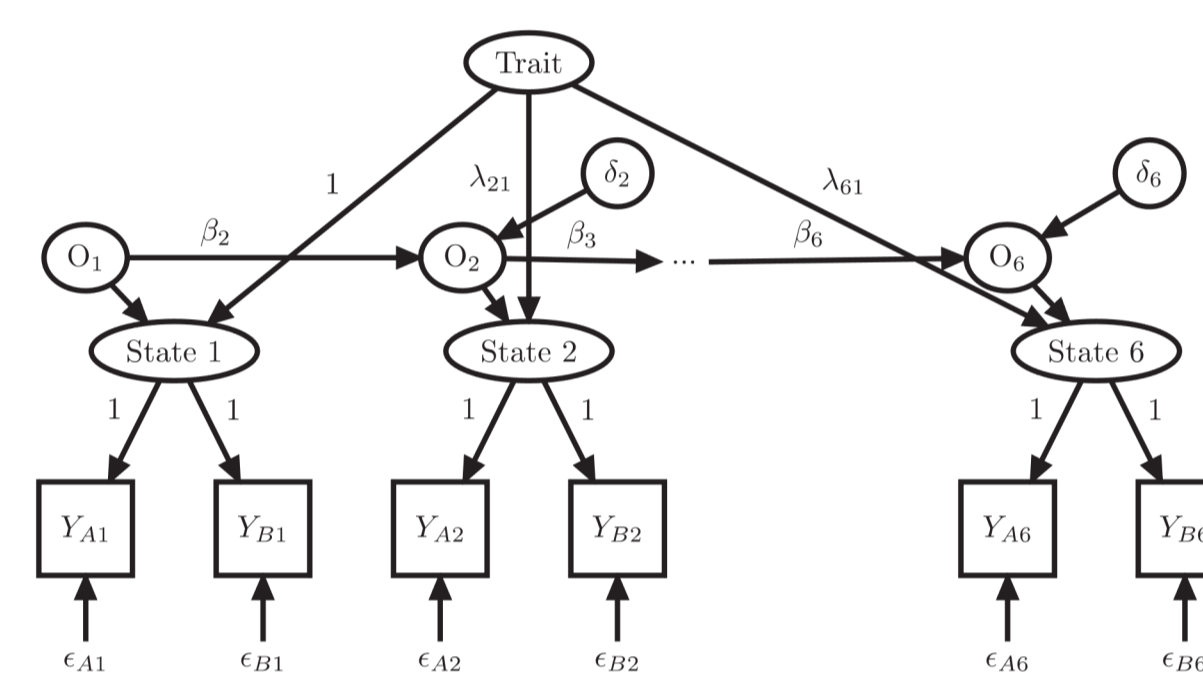
Some example projects



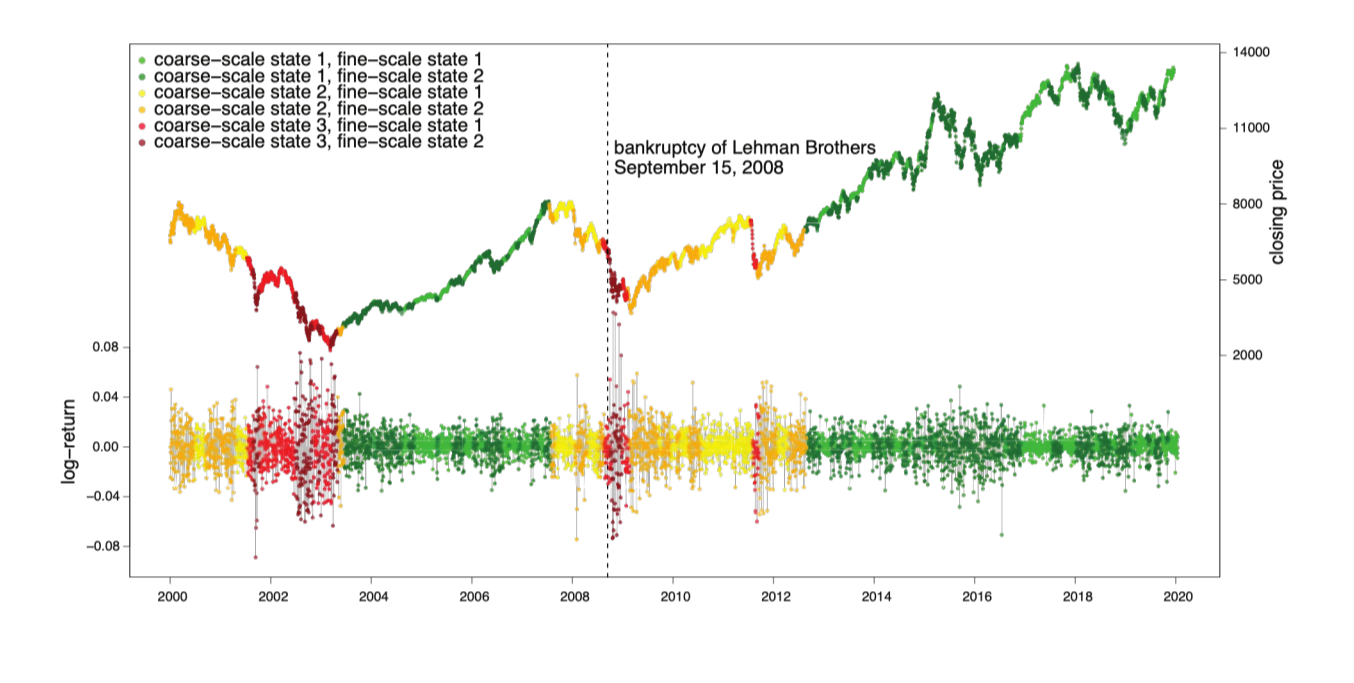
State switching SDEs: estimated state-dependent drift and diffusion for dive profiles of multiple narwhals.



MMPPs: Daily defined dose sequence (upper plot) and step function for number of consultations (lower plot).



The trait-state occasion SEM includes a latent trait variable and an autoregressive component.



Decoded time series of the daily DAX prices and the corresponding log-returns decoded according to a 3-state HMM.

Objectives

SILAS combines

- the expertise on several key methodological approaches and...
- ...the multi-disciplinary view on empirical applications

to consolidate the wild-west like situation concerning inference on latent structures from increasingly complex sequential data.

Specifically, we aim to identify synergies between the models and methods used in the distinct disciplines, bridging gaps where necessary.

Concrete objective: application for a DFG Research Training Group (under way).

References for the example projects

- Adam, T., Heide-Jørgensen, M. P., & Ditlevsen, S. (2024, July). Modelling narwhal diving behaviour and responses to sound exposure using stochastic differential equations with state-switching coefficients. In *38th International Workshop on Statistical Modelling*: 35-38.
- Mews, S., Surmann, B., Hasemann, L., & Elkenkamp, S. (2023). Markov-modulated marked Poisson processes for modeling disease dynamics based on medical claims data. *Statistics in Medicine*, 42(21): 3804-3815.
- Loncke, J., Mayer, A., Eichelsheim, V. I., Branje, S. J. T., Meeus, W. H. J., Koot, H. M., Buysse, A., & Loey, T. (2017). Latent state-trait models for longitudinal family data: Investigating consistency in perceived support. *European Journal of Psychological Assessment*, 33(4): 256-270.
- Oelschläger, L., Adam, T., & Michels, R. (2024). fHMM: Hidden Markov Models for Financial Time Series in R. *Journal of Statistical Software*, 109: 1-25.

