Hitchhiker’s guide to analyzing neural and behavioral time-series data

Rich Pang

2025

Don’t panic!

Make predictions

This guide

Comprehensive guide to analyzing time-series data in neuroscience

For

Anyone interested in working with modern neuroscience data

Whether you want to

Analyze data collected to address a specific question

Extract new scientific results from existing data

Note

Designed to be comprehensive yet concise.
To be used as both course and guidebook/reference.

Why is analyzing time-series data important?

Default output of most modern measurement devices

<img>
EEG-like time-series

<img>
Population raster plots

<img>
Tracked animal video still

Crucial to understand how to work rigorously with such data, but not easy.

Analysis is much more than classical signal processing.

Generally requires both technical and domain knowledge.

Challenges of analyzing time-series data in neuroscience

- Many small decisions to make along the way
- Easy to make mistakes
- Takes time to be rigorous
- *Many* methods to choose from
- Big datasets
- Multi-modal datasets
- Weird statistics/ lack of trials
- Missing data/variable trial lengths
- High dimensionality
- Violates many assumptions of classic signal processing
- Model-fitting can be highly complex
- Difficult to interpret analysis results

Goals of this course

<u>Fast-track your way to expertise in time-series analysis</u>

- Fundamental concepts and philosophy
- Canonical and state-of-the-art methods
- Common challenges and solutions
- Learn to efficiently transform data into trustworthy scientific results
- Learn to design new analyses to ask new questions

The complete outline: everything you need to hit the ground running

Fundamental concepts

- Common data types in neuroscience
  - Intracellular voltage
  - Spike trains
  - LFP/ECoG
  - Calcium imaging
  - fMRI/EEG/MEG
  - Fiber photometry?
  - Video/tracked behavior
  - Stimuli/other sensors
  - Simulation data
- Modeling as analysis
- Model fitting
  - Parameters
  - Loss functions
  - Training/test data
  - Overfitting/bias-variance tradeoff
  - Model comparison
- Random processes perspective
- Dynamical systems perspective

Philosophy

- What you see depends on how you look
- Frameworks/theories/models
- Descriptive/mechanistic/normative levels
- Importance of making predictions
- Transforming data into scientific results

Methods survey I: Canonical methods

- Common pre-processing steps
  - Spike sorting
  - Image processing/ROI extraction
  - Estimating firing rates
  - Tracking
  - Smoothing
  - Detrending
  - Creating artificial trials
- Classical signal processing
  - Correlation functions
  - Fourier transforms/power spectra
  - Linear filters/impulse response
- Canonical neural data analyses
  - Raster plots and ISIs
  - PSTHs and tuning curves
  - LNP/spike-triggered average

Methods survey II: Modern methods

- Dimensionality reduction
- Clustering/segmentation
- Statistical models
- Latent variable models
- Dynamics/state-space models
- Mechanistic models
- Spatiotemporal analyses
- Neural networks/VAEs/ELBO
- Learning/inference techniques
- Information theoretic methods

Common issues

- Multicollinearity
- Controlled vs naturalistic data
- Nonstationarity and long timescales
- Trials/Sessions/Animals/Conditions
- Binning
- Missing data
- Too much data
- Gotchas: normalization, correlated training/test data, what is N? etc.
- Interpreting data-driven analyses

Practical tips

- Approaching/vetting a new dataset
- Data pre-processing
- Statistics
- Data munging/storage
- Null/control datasets
- Deconstructing fit models
- How to not make mistakes
- Leveraging AI
- Designing custom analyses
- Coding strategies
- Reproducibility/data sharing

Miscellaneous

- Other methods
- Common software
- Other resources