

Functional Data Analysis in Practice

Supervision: Marlène Baumeister and Markus Pauly jointly with Łukasz Smaga and Natalia Stefanska (Adam Mickiewicz University, Poznań)

Research environment: International collaboration within the WEAVE project by Łukasz Smaga and Markus Pauly

Modern technologies frequently generate data that are observed continuously over time, space, or another domain. Examples include measurement curves, sensor signals, biomedical recordings, or environmental trajectories. Such observations are commonly referred to as functional data.

Functional Data Analysis (FDA) provides statistical tools for analysing these data while accounting for their functional nature. In many practical applications, several experimental factors influence the observed curves. For example, researchers may be interested in the effects of treatment, group membership, time period, or their interactions. This leads to functional factorial designs.

We have access to several interesting functional datasets from ongoing or potential collaborations with applied research partners. Depending on the student's interests, possible application areas include, for example, neuroscience, biomedical research, educational science, or other empirical disciplines in which functional observations arise.

The central aim is to conduct a substantive in depth statistical analysis of at least one such dataset in cooperation with the application partner. Typical questions include whether and how experimental factors are associated with differences in the observed functional trajectories, which parts of the domain contribute most strongly to these differences, and how robust the conclusions are with respect to the chosen analysis method.

For this purpose, modern methods from Functional Data Analysis can be used. Depending on the application, these may include

- exploratory visualization and preprocessing of functional observations,
- smoothing or basis representations,
- descriptive functional summaries,
- statistical testing procedures for one-way or factorial designs,
- multiple comparison procedures,
- functional regression models,
- and, where relevant, prediction or classification methods.

Implementations of several relevant methods are already available, for example through the R packages `fdANOVA`, `rmfanova`, `multifanova`, `gmtFD`, and further research software.

The exact application and methodological scope will be selected jointly based on the student's interests and background. This topic is therefore particularly suitable for students who enjoy working at the interface between statistical methodology and substantive scientific questions.

The project can be adapted to both B.Sc. and M.Sc. theses, with the level of methodological depth adjusted accordingly. Experience with R or Python and an interest in applied statistics are desirable.

A particular attraction of the project is the opportunity to work not only with statisticians but also with researchers from the corresponding application field, thereby gaining experience in translating substantive scientific questions into a rigorous statistical analysis.

Contact: Marlène Baumeister (marlene.baumeister@tu-dortmund.de) and Markus Pauly (pauly@statistik.tu-dortmund.de)

Some References

- Morris, J. S. (2015). Functional regression. *Annual Review of Statistics and Its Application*, **2**, 321–359.
- Munko, M., Ditzhaus, M., Pauly, M., & Smaga, Ł. (2025). Multiple comparison procedures for simultaneous inference in functional MANOVA. *Electronic Journal of Statistics*, *19*(1), 2637–2732.
- Ramsay, J. O., & Silverman, B. W. (2005). *Functional Data Analysis*, 2nd ed. Springer.

Further application-specific literature will be selected depending on the chosen dataset and scientific field.