Christine Müller
Mirko Jakubzik

Technische Universität Dortmund
Summersemester 2025
Case Studies / Fallstudien II

Project 1
Forecasting electrical power consumption using individual load profiles

The file „*Haushalte mit Waermepumpe OHNE PV.zip*“ contains 33 data sets of the power consumption of 33 households with heat pump but without photovoltaic system. The column „*PUMPE_TOT*“ contains the power consumption of the heat pump in watts and the column „*HAUSHALT_TOT*“ the household's further electricity consumption in watts. The column „*TEMPERATURE:TOTAL*“ contains the temperature at the measurement times. All measurements were carried out in 2019. In this project, only the column „*HAUSHALT_TOT*“ should be considered.

Problems:

1) Create a data set with smoothed daily power consumptions as dependent multivariate variable and with day, weekday, and household as explanatory variables. Perform the smoothing as follows:
    a) Smooth only the values of the day and some neighbour hours.
    b) Create the smoothed day curve by using not only the power consumption of the day but also of some past and future days of the same weekday.
2) Choose randomly 30 households and study the influence of the household, the weekday and the season on the smoothed daily power curves. Can the models be simplified? Is it in particular necessary to assume interactions between the explanatory variables?
3) Study the influence of the weekday and the season on the smoothed daily power curves for each household separately. Can now the models be simplified? Is it here in particular necessary to assume interactions between the explanatory variables?
4) Cluster the in 2) randomly chosen households with respect to the estimated effects in the model obtained in 3). Check whether you should do the clustering for each weekday separately.
5) Classify the remaining 3 households to the clusters obtained in 4) by using their estimated effects in the model obtained in 3). Possibly, you may do the classification for each weekday separately.
6) For each cluster obtained in 4), calculate load profiles depending only on the weekday and the season.
7) Forecast the power consumption in the next hour for the remaining 3 households by

i) using no individual load profiles,
ii) using the classification result in 5) and the load profiles obtained in 6).

Use for the forecast

a) the Kalman filter,
b) a particle filter.

8) Compare the forecast errors for the different smoothing methods in 1) and the different forecast methods in 6) with and without individual load profiles.

Literature

- Brockwell, P. J. and Davis, R. A. (2002). *Introduction to Time Series and Forecasting*. Springer, New York.
- Chatfield, C. (2019). *The Analysis of Time Series: An Introduction with R*. Chapman & Hall/CRC Texts in Statistical Science, Boca Raton.
- Christensen, R. (2001). *Advanced Linear Modeling*. Springer, New York.
- Hastie, T., Tibshirani, R., and Friedman, J. (2002). *The Elements of Statistical Learning. Data Mining, Inference and Prediction*. Springer, New York.
- James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An Introduction to Statistical Learning*. Springer, New York.
- Petris, G., Petrone, S., and Campagnoli, P. (2009). *Dynamic Models with R*. Springer, New York.
- Rencher, A.C. (1998). *Multivariate Statistical Inference and Applications*. Wiley, New York.
- Schlemminger, M., Ohrdes, T., Schneider, E., and Knoop, M. (2022). Dataset on electrical single-family house and heat pump load profiles in Germany. *Sci Data **9**, 56.* https://doi.org/10.1038/s41597-022-01156-1

Submission deadline: Tuesday, June 17, 2025.