Preface

Abstract The topic of this contribution is the statistical analysis of climatological time series. The data sets consist of monthly (and daily) temperature means and precipitation amounts gained at German weather stations. Emphasis lies on the methods of time series analysis, comprising plotting, modeling and predicting climate values in the near future. Further, correlation analysis (including principal components), spectral and wavelet analysis in the frequency domain and categorical data analysis are applied.

Introduction Within the context of the general climate discussion, the evaluation of climate time series gains growing importance. Here we mainly use the monthly data of temperature (mean) and precipitation (amount) from German weather stations, raised over many years. We analyze the series by applying statistical methods and describe the possible relevance of the results. First the climate series (annual and seasonal data) will be considered in their own right, by employing descriptive methods. Long-term trends—especially the opposed trends of temperature in the nineteenth and twentieth century—are statistically tested. The auto-correlations, that are the correlations of (yearly, seasonally, monthly, daily) data, following each other in time, are calculated before and after a trend or a seasonal component is removed. In the framework of correlation analysis, we use principal components to structure climate variables from different stations. We also formulate well-known folk (or country) sayings about weather in a statistical language and check their legitimacy.

The notion of auto-correlation leads us to the problem, how to model the evolution of the underlying data process. For annual data, we use ARMA-type time series models, applied to the differenced series, with a subsequent residual analysis to assess their adequacy. For the latter task, GARCH-type models can be employed. In the present text, predictions of the next outcomes are understood as forecasts: The prediction for time point $t + 1$ is strictly based on information up to time t only (thus parameters must be estimated for each t anew). The ARMA-type modeling is compared with (left-sided) moving averages by using a goodness-of-fit criterion calculated from the squared residuals.

Guided by the modeling of annual data, we similarly proceed with monthly data. Here, it is the detrended series, to which we fit an ARMA-model. With this method, the yearly seasonality can correctly be reproduced.

Daily records on temperature reveal a seasonal component—as known from monthly data, such that the adjusting of the series is advisable. We study a spatial effect, namely the cross-correlation between five German stations. Half of the daily precipitation data consists of zeros; here we are led to logistic regression approaches, to categorical data analysis and—with repect to heavy precipitation—to event-time analysis.

We continue with analyses in the frequency domain. Periodograms, spectral density estimations, and wavelet analyses are applied to find and trace periodical phenomena in the series.

Then, we present two approaches for predicting annual and monthly data, which are quite different from those based on ARMA-type models, namely growing polynomials and sin-/cos-approximations, respectively. Further, the one-step predictions of the preceding sections are extended to l-step (i. e. l-years) forecasts. This is done by the Box and Jenkins and by the Monte Carlo method. Finally, specific features of temperature and of precipitation data are investigated by means of multiple correlation coefficients.

The numerical analysis is performed by using the open-source package

`R [cran.r-project.org].`

An introductory manual as for instance the book of Dalgaard (2002) is useful. The `R` codes are presented within complete programs. We have two kinds of comments. If comments should appear in the output, they are standing between "…" signs. If they are only directed to the reader of the program and should be ignored by the program, they begin with the ♯ sign. Together with the `read.table(...)` command in program `R 1.1`, the programs are ready to run. Optionally the `sink(...)` command in program `R 2.1` can be employed (to divert the output to an external file). The index lists the `R` commands with the page of their first occurrence.

This book addresses

- Students and lecturers in statistics and mathematics, who like to get knowledge about statistical methods for time series (in a wide sense) on one side and about an interesting and relevant field of application on the other
- Meteorologists and other scientists, who look for statistical tools to analyze climate series and who need program codes to realize the work in `R`.

Programs, which are ready to run, and data sets on climatological series (both provided on the author's homepage) enable the reader to perform own exercises and allow own applications.

www.math.lmu.de/∼pruscha/ Helmut Pruscha

http://www.springer.com/978-3-642-32083-5

Statistical Analysis of Climate Series
Analyzing, Plotting, Modeling, and Predicting with R
Pruscha, H.
2013, VIII, 176 p., Hardcover
ISBN: 978-3-642-32083-5