Contents

Preface vii
Acknowledgements xi
List of Algorithms xxi
List of Figures xxv
List of Tables xxxi

Part I Fundamental Concepts

1. Introduction 3
1.1 Climate archives, variables and dating 5
1.2 Noise and statistical distribution 6
1.3 Persistence 11
1.4 Spacing 14
1.5 Aim and structure of this book 24
1.6 Background material 26

2. Persistence Models 33
2.1 First-order autoregressive model 33
2.1.1 Even spacing 34
2.1.1.1 Effective data size 36
2.1.2 Uneven spacing 37
2.1.2.1 Embedding in continuous time 38
2.2 Second-order autoregressive model 39
2.3 Mixed autoregressive moving average model 41
2.4 Other models 42
2.4.1 Long-memory processes 42

2.4.2 Nonlinear and non-Gaussian models 43
2.5 Climate theory 44
2.5.1 Stochastic climate models 45
2.5.2 Long memory of temperature fluctuations? 47
2.5.3 Long memory of river runoff 51
2.6 Background material 54
2.7 Technical issues 63

3. Bootstrap Confidence Intervals 65
3.1 Error bars and confidence intervals 66
3.1.1 Theoretical example: mean estimation of Gaussian white noise 68
3.1.2 Theoretical example: standard deviation estimation of Gaussian white noise 69
3.1.3 Real world 71
3.2 Bootstrap principle 74
3.3 Bootstrap resampling 76
3.3.1 Nonparametric: moving block bootstrap 78
3.3.1.1 Block length selection 78
3.3.1.2 Uneven spacing 80
3.3.1.3 Systematic model parts and nonstationarity 81
3.3.2 Parametric: autoregressive bootstrap 83
3.3.2.1 Even spacing 83
3.3.2.2 Uneven spacing 83
3.3.3 Parametric: surrogate data 86
3.4 Bootstrap confidence intervals 86
3.4.1 Normal confidence interval 87
3.4.2 Student's t confidence interval 88
3.4.3 Percentile confidence interval 88
3.4.4 BCa confidence interval 88
3.5 Examples 89
3.6 Bootstrap hypothesis tests 91
3.7 Notation 94
3.8 Background material 99
3.9 Technical issues 106

Part II Univariate Time Series

4. Regression I 113

4.1 Linear regression 114

4.1.1 Weighted least-squares and ordinary least-squares estimation 114

4.1.1.1 Example: Arctic river runoff 115

4.1.2 Generalized least-squares estimation 116

4.1.3 Other estimation types 118

4.1.4 Classical confidence intervals 119

4.1.4.1 Prais–Winsten procedure 121

4.1.4.2 Cochrane–Orcutt transformation 121

4.1.4.3 Approach via effective data size 123

4.1.5 Bootstrap confidence intervals 124

4.1.6 Monte Carlo experiments: ordinary least-squares estimation 124

4.1.7 Timescale errors 129

4.1.7.1 Nonparametric: pairwise-moving block bootstrap 131

4.1.7.2 Parametric: timescale-autoregressive bootstrap 131

4.1.7.3 Hybrid: timescale-moving block bootstrap 136

4.1.7.4 Monte Carlo experiments 136

4.2 Nonlinear regression 141

4.2.1 Climate transition model: ramp 142

4.2.1.1 Estimation 143

4.2.1.2 Example: Northern Hemisphere Glaciation 145

4.2.1.3 Bootstrap confidence intervals 146

4.2.1.4 Example: onset of Dansgaard–Oeschger event 5 146

4.2.2 Trend-change model: break 150

4.2.2.1 Estimation 150

4.2.2.2 Example: Arctic river runoff (continued) 152

4.2.2.3 Bootstrap confidence intervals 152

4.3 Nonparametric regression or smoothing 153

4.3.1 Kernel estimation 153

4.3.2 Bootstrap confidence intervals and bands 156

4.3.3 Extremes or outlier detection 157
4.3.3.1 Example: volcanic peaks in the NGRIP sulfate record 159
4.3.3.2 Example: hurricane peaks in the Lower Mystic Lake varve thickness record 159
4.4 Background material 161
4.5 Technical issues 173

5. Spectral Analysis 177
5.1 Spectrum 177
5.1.1 Example: AR(1) process, discrete time 180
5.1.2 Example: AR(2) process, discrete time 180
5.1.3 Physical meaning 181
5.2 Spectral estimation 183
5.2.1 Periodogram 183
5.2.2 Welch's Overlapped Segment Averaging 186
5.2.3 Multitaper estimation 188
5.2.3.1 F test 190
5.2.3.2 Weighted eigenspectra 191
5.2.3.3 Zero padding 192
5.2.3.4 Jackknife 192
5.2.3.5 Advanced topics: CI coverage accuracy and uneven spacing 194
5.2.3.6 Example: radiocarbon spectrum 195
5.2.4 Lomb–Scargle estimation 196
5.2.4.1 Bias correction 197
5.2.4.2 Covariance 199
5.2.4.3 Harmonic filter 199
5.2.4.4 Advanced topics: degrees of freedom, bandwidth, oversampling and highest frequency 201
5.2.5 Peak detection: red-noise hypothesis 202
5.2.5.1 Multiple tests 202
5.2.6 Example: peaks in monsoon spectrum 205
5.2.7 Aliasing 205
5.2.8 Timescale errors 208
5.2.9 Example: peaks in monsoon spectrum (continued) 209
5.3 Background material 215
5.4 Technical issues 225

6. Extreme Value Time Series **229**

6.1 Data types 229

6.1.1 Event times 230

6.1.1.1 Example: Elbe winter floods 230

6.1.2 Peaks over threshold 230

6.1.2.1 Example: volcanic peaks in the NGRIP sulfate record (continued) 231

6.1.3 Block extremes 231

6.1.4 Remarks on data selection 232

6.2 Stationary models 232

6.2.1 Generalized Extreme Value distribution 232

6.2.1.1 Model 233

6.2.1.2 Maximum likelihood estimation 233

6.2.2 Generalized Pareto distribution 235

6.2.2.1 Model 235

6.2.2.2 Maximum likelihood estimation 235

6.2.2.3 Model suitability 237

6.2.2.4 Return period 238

6.2.2.5 Probability weighted moment estimation 239

6.2.3 Bootstrap confidence intervals 240

6.2.4 Example: Elbe summer floods, 1852 to 2002 241

6.2.5 Persistence 243

6.2.5.1 Condition $D(u_n)$ 243

6.2.5.2 Extremal index 243

6.2.5.3 Long memory 244

6.2.6 Remark: tail estimation 244

6.2.7 Remark: optimal estimation 246

6.3 Nonstationary models 246

6.3.1 Time-dependent Generalized Extreme Value distribution 247

6.3.2 Inhomogeneous Poisson process 248

6.3.2.1 Model 248

6.3.2.2 Nonparametric occurrence rate estimation 249

6.3.2.3 Boundary bias reduction 250

6.3.2.4 Bandwidth selection 251

6.3.2.5 Example: Elbe winter floods (continued) 252

6.3.2.6 Bootstrap confidence band 253
6.3.2.7 Example: Elbe winter floods (continued) 256
6.3.2.8 Example: volcanic peaks in the NGRIP sulfate record (continued) 257
6.3.2.9 Example: hurricane peaks in the Lower Mystic Lake varve thickness record (continued) 257
6.3.2.10 Parametric Poisson models and hypothesis tests 258
6.3.2.11 Monte Carlo experiment: Cox–Lewis test versus Mann–Kendall test 260
6.3.3 Hybrid: Poisson–extreme value distribution 264
6.4 Sampling and time spacing 266
6.5 Background material 269
6.6 Technical issues 279

Part III Bivariate Time Series

7. Correlation 285
7.1 Pearson's correlation coefficient 286
7.1.1 Remark: alternative correlation measures 287
7.1.2 Classical confidence intervals, non-persistent processes 287
7.1.3 Bivariate time series models 289
7.1.3.1 Bivariate white noise 289
7.1.3.2 Bivariate first-order autoregressive process 290
7.1.4 Classical confidence intervals, persistent processes 291
7.1.5 Bootstrap confidence intervals 293
7.1.5.1 Pairwise-moving block bootstrap 293
7.1.5.2 Pairwise-autoregressive bootstrap 295
7.2 Spearman's rank correlation coefficient 295
7.2.1 Classical confidence intervals, non-persistent processes 298
7.2.2 Classical confidence intervals, persistent processes 300
7.2.3 Bootstrap confidence intervals 301
7.2.3.1 Pairwise-moving block bootstrap 301
7.2.3.2 Pairwise-autoregressive bootstrap 302
7.3 Monte Carlo experiments 302

7.4 Example: Elbe runoff variations 309
7.5 Unequal timescales 311
7.5.1 Binned correlation 312
7.5.2 Synchrony correlation 314
7.5.3 Monte Carlo experiments 316
7.5.3.1 Optimal estimation 321
7.5.4 Example: Vostok ice core records 322
7.6 Background material 323
7.7 Technical issues 338

8. Regression II 339

8.1 Linear regression 340
8.1.1 Ordinary least-squares estimation 340
8.1.1.1 Bias correction 341
8.1.1.2 Prior knowledge about standard deviations 341
8.1.2 Weighted least-squares for both variables estimation 343
8.1.2.1 Prior knowledge about standard deviation ratio 343
8.1.2.2 Geometric interpretation 344
8.1.3 Wald–Bartlett procedure 345
8.2 Bootstrap confidence intervals 346
8.2.1 Simulating incomplete prior knowledge 348
8.3 Monte Carlo experiments 350
8.3.1 Easy setting 350
8.3.2 Realistic setting: incomplete prior knowledge 353
8.3.3 Dependence on accuracy of prior knowledge 355
8.3.4 Mis-specified prior knowledge 357
8.4 Example: climate sensitivity 359
8.5 Prediction 362
8.5.1 Example: calibration of a proxy variable 364
8.6 Lagged regression 367
8.6.1 Example: CO_2 and temperature variations in the Pleistocene 368
8.7 Background material 373
8.8 Technical issues 379

Part IV Outlook

9. Future Directions **383**

9.1 Timescale modelling 383

9.2 Novel estimation problems 384

9.3 Higher dimensions 385

9.4 Climate models 385

9.4.1 Fitting climate models to observations 387

9.4.2 Forecasting with climate models 388

9.4.3 Design of the cost function 389

9.4.4 Climate model bias 390

9.5 Optimal estimation 391

References **397**

Subject Index **455**

Author Index **467**