

Statistical Methods for Trend Detection and Analysis in the Environmental Sciences

Richard E. Chandler

Department of Statistical Science, University College London, UK

E. Marian Scott

School of Mathematics and Statistics, University of Glasgow, UK



A John Wiley & Sons, Ltd., Publication

Contents

Preface	xi
Contributing authors	xiii
Part I METHODOLOGY	1
1 Introduction	3
1.1 What is a trend?	4
1.2 Why analyse trends?	5
1.3 Some simple examples	6
1.3.1 Dutch wind speeds	6
1.3.2 North Sea haddock stocks	8
1.3.3 Alkalinity in the Round Loch of Glenhead	10
1.3.4 Atmospheric ozone in eastern England	12
1.4 Considerations and difficulties	14
1.4.1 Autocorrelation	14
1.4.2 Effect of other variables	15
1.4.3 Lack of designed experiments	15
1.4.4 Consideration of auxiliary information	16
1.4.5 The necessity of extrapolation	17
1.5 Scope of the book	19
1.6 Further reading	20
References	21
2 Exploratory analysis	25
2.1 Data visualisation	25
2.1.1 Time series plots	26
2.1.2 Boxplots	28
2.1.3 The autocorrelation function	30
2.1.4 Irregularly spaced data – the variogram	35
2.1.5 Relationships between variables	38
2.2 Simple smoothing	41
2.2.1 Moving averages	41
2.2.2 Local polynomial fitting	42
2.2.3 Further considerations	43

2.3	Linear filters	45
2.3.1	Frequency considerations	46
2.3.2	The convolution theorem and filter design	48
2.3.3	Dealing with end effects	50
2.3.4	Other applications	51
2.4	Classical test procedures	54
2.5	Concluding comments	57
	References	58
3	Parametric modelling – deterministic trends	61
3.1	The linear trend	63
3.1.1	Checking the model assumptions	65
3.1.2	Choosing the time index	69
3.1.3	An overview of least squares	70
3.1.4	Extrapolation	74
3.1.5	Influential observations	76
3.1.6	Other methods of model fitting	79
3.2	Multiple regression techniques	82
3.2.1	Representing seasonality in regression models	83
3.2.2	Interactions	86
3.2.3	Model building and selection	87
3.3	Violations of assumptions	94
3.3.1	Dealing with heteroscedasticity	94
3.3.2	Dealing with non-normality	96
3.3.3	Dealing with autocorrelation	97
3.4	Nonlinear trends	105
3.4.1	Nonlinear least squares	105
3.4.2	Cycles	106
3.4.3	Changepoints and interventions	108
3.5	Generalised linear models	111
3.5.1	Parameter estimation	113
3.5.2	Model comparison	115
3.5.3	Model checking	116
3.5.4	Prediction with GLMs	117
3.5.5	Extensions and refinements	118
3.6	Inference with small samples	120
	References	122
4	Nonparametric trend estimation	127
4.1	An introduction to nonparametric regression	127
4.1.1	Linear smoothing	128
4.1.2	Local linear regression	129
4.1.3	Spline smoothing	130
4.1.4	Choice of smoothing parameter	131
4.1.5	Variance estimators	135
4.1.6	Standard errors for the regression function	135
4.1.7	Testing for consistency with parametric models	136
4.2	Multiple covariates	140
4.2.1	Additive, semiparametric and bivariate models	140

4.2.2	The backfitting algorithm	142
4.2.3	Inference for additive models	144
4.2.4	Handling autocorrelation	148
4.3	Other nonparametric estimation techniques	151
4.3.1	Lowess smoothing	151
4.3.2	Wavelets	154
4.3.3	Varying coefficient models	160
4.3.4	Discontinuity detection	161
4.3.5	Quantile regression	162
4.4	Parametric or nonparametric?	166
	References	167
5	Stochastic trends	171
5.1	Stationary time series models and their properties	171
5.1.1	Autoregressive processes	171
5.1.2	Moving average processes	174
5.1.3	Mixed ARMA processes	174
5.1.4	Model identification	175
5.1.5	Parameter estimation	177
5.1.6	Model checking	182
5.1.7	Forecasting	186
5.1.8	The backshift operator	190
5.2	Trend removal via differencing	193
5.2.1	ARIMA models	194
5.2.2	Spurious regressions	199
5.3	Long memory models	201
5.4	Models for irregularly spaced series	205
5.5	State space and structural models	207
5.5.1	Simple structural time series models	207
5.5.2	The state space representation	209
5.5.3	The Kalman filter	213
5.5.4	Parameter estimation	217
5.5.5	Connection with nonparametric smoothing	219
5.6	Nonlinear models	228
	References	231
6	Other issues	235
6.1	Multisite data	235
6.1.1	Visualisation	236
6.1.2	Modelling	239
6.2	Multivariate series	241
6.2.1	Dimension reduction	241
6.2.2	Multivariate models	244
6.3	Point process data	245
6.3.1	Poisson processes	246
6.3.2	Other point process models	249
6.3.3	Marked point processes	250
6.4	Trends in extremes	250
6.4.1	Approaches based on block maxima	251

6.4.2	Approaches based on threshold exceedances	253
6.4.3	Modern developments	256
6.5	Censored data	257
	References	260

Part II CASE STUDIES 265

7 Additive models for sulphur dioxide pollution in Europe 267

Marco Giannitrapani, Adrian Bowman, E. Marian Scott and Ron Smith

7.1	Introduction	267
7.2	Additive models with correlated errors	269
7.2.1	An introduction to additive models	269
7.2.2	Smoothing techniques	270
7.2.3	Smoothing correlated data	272
7.2.4	Fitting additive models	273
7.2.5	Comparing nonparametric models	276
7.3	Models for the SO ₂ data	277
7.4	Conclusions	281
	Acknowledgement	282
	References	282

8 Rainfall trends in southwest Western Australia 283

Richard E. Chandler, Bryson C. Bates and Stephen P. Charles

8.1	Motivation	283
8.2	The study region	285
8.3	Data used in the study	285
8.4	Modelling methodology	289
8.4.1	Generalised linear models for daily rainfall	289
8.4.2	Temporal and spatial dependence	290
8.4.3	Covariates considered	291
8.4.4	Modelling strategy	292
8.5	Results	293
8.5.1	Diagnostics	294
8.5.2	Trends in wet-day precipitation amounts	295
8.5.3	Trends in precipitation occurrence	296
8.5.4	Combined trends in occurrence and amounts	302
8.6	Summary and conclusions	303
	References	304

9 Estimation of common trends for trophic index series 307

Alain F. Zuur, Elena N. Ieno, Cristina Mazziotti, Giuseppe Montanari, Attilio Rinaldi and Carla Rita Ferrari

9.1	Introduction	307
9.2	Data exploration	311
9.3	Common trends and additive modelling	314
9.3.1	Adding autocorrelation to the additive model	316
9.3.2	Combining the data from the eight stations	319
9.3.3	Doing it all within a parametric model	323

9.4	Dynamic factor analysis to estimate common trends	324
9.4.1	The underlying model	324
9.5	Discussion	328
	Acknowledgement	329
	References	329
10	A space–time study on forest health	333
	<i>Thomas Kneib and Ludwig Fahrmeir</i>	
10.1	Forest health: survey and data	333
10.2	Regression models for longitudinal data with ordinal responses	336
10.3	Spatiotemporal models	340
10.3.1	Penalised splines	341
10.3.2	Interaction surfaces	343
10.3.3	Spatial trends	344
10.3.4	Inference in spatiotemporal models	346
10.4	Spatiotemporal modelling and analysis of forest health data	348
	Acknowledgements	357
	References	357
	Index	359