

Use R!

Paul S.P. Cowpertwait • Andrew V. Metcalfe **Introductory Time Series with R**

Yearly global mean temperature and ocean levels, daily share prices, and the signals transmitted back to Earth by the Voyager space craft are all examples of sequential observations over time known as time series. This book gives you a step-by-step introduction to analysing time series using the open source software R. Each time series model is motivated with practical applications, and is defined in mathematical notation. Once the model has been introduced it is used to generate synthetic data, using R code, and these generated data are then used to estimate its parameters. This sequence enhances understanding of both the time series model and the R function used to fit the model to data. Finally, the model is used to analyse data taken from a practical application. By using R, the whole procedure can be reproduced by the reader. All the data sets used in the book are available on the website <http://www.massey.ac.nz/~pscowper/ts>.

The book is written for undergraduate students of mathematics, economics, business and finance, geography, engineering and related disciplines, and postgraduate students who may need to analyse time series as part of their taught programme or their research.

Paul S.P. Cowpertwait is a senior lecturer in statistics at Massey University with a substantial research record in both the theory and applications of time series and stochastic models. Andrew V. Metcalfe is an associate professor in the School of Mathematical Sciences at the University of Adelaide, and an author of six statistics textbooks and numerous research papers. Both authors have extensive experience of teaching time series to students at all levels.

STATISTICS

ISBN 978-0-387-88697-8



EAN

9 780387 886978 >

› springer.com



Contents

Preface	vii
1 Time Series Data	1
1.1 Purpose	1
1.2 Time series	2
1.3 R language	3
1.4 Plots, trends, and seasonal variation	4
1.4.1 A flying start: Air passenger bookings	4
1.4.2 Unemployment: Maine	7
1.4.3 Multiple time series: Electricity, beer and chocolate data	10
1.4.4 Quarterly exchange rate: GBP to NZ dollar	14
1.4.5 Global temperature series	16
1.5 Decomposition of series	19
1.5.1 Notation	19
1.5.2 Models	19
1.5.3 Estimating trends and seasonal effects	20
1.5.4 Smoothing	21
1.5.5 Decomposition in R	22
1.6 Summary of commands used in examples	24
1.7 Exercises	24
2 Correlation	27
2.1 Purpose	27
2.2 Expectation and the ensemble	27
2.2.1 Expected value	27
2.2.2 The ensemble and stationarity	30
2.2.3 Ergodic series*	31
2.2.4 Variance function	32
2.2.5 Autocorrelation	33

2.3	The correlogram	35
2.3.1	General discussion	35
2.3.2	Example based on air passenger series	37
2.3.3	Example based on the Font Reservoir series	40
2.4	Covariance of sums of random variables	41
2.5	Summary of commands used in examples	42
2.6	Exercises	42
3	Forecasting Strategies	45
3.1	Purpose	45
3.2	Leading variables and associated variables	45
3.2.1	Marine coatings	45
3.2.2	Building approvals publication	46
3.2.3	Gas supply	49
3.3	Bass model	51
3.3.1	Background	51
3.3.2	Model definition	51
3.3.3	Interpretation of the Bass model*	51
3.3.4	Example	52
3.4	Exponential smoothing and the Holt-Winters method	55
3.4.1	Exponential smoothing	55
3.4.2	Holt-Winters method	59
3.4.3	Four-year-ahead forecasts for the air passenger data	62
3.5	Summary of commands used in examples	64
3.6	Exercises	64
4	Basic Stochastic Models	67
4.1	Purpose	67
4.2	White noise	68
4.2.1	Introduction	68
4.2.2	Definition	68
4.2.3	Simulation in R	68
4.2.4	Second-order properties and the correlogram	69
4.2.5	Fitting a white noise model	70
4.3	Random walks	71
4.3.1	Introduction	71
4.3.2	Definition	71
4.3.3	The backward shift operator	71
4.3.4	Random walk: Second-order properties	72
4.3.5	Derivation of second-order properties*	72
4.3.6	The difference operator	72
4.3.7	Simulation	73
4.4	Fitted models and diagnostic plots	74
4.4.1	Simulated random walk series	74
4.4.2	Exchange rate series	75

4.4.3	Random walk with drift	77
4.5	Autoregressive models	79
4.5.1	Definition	79
4.5.2	Stationary and non-stationary AR processes	79
4.5.3	Second-order properties of an AR(1) model	80
4.5.4	Derivation of second-order properties for an AR(1) process*	80
4.5.5	Correlogram of an AR(1) process	81
4.5.6	Partial autocorrelation	81
4.5.7	Simulation	81
4.6	Fitted models	82
4.6.1	Model fitted to simulated series	82
4.6.2	Exchange rate series: Fitted AR model	84
4.6.3	Global temperature series: Fitted AR model	85
4.7	Summary of R commands	87
4.8	Exercises	87
5	Regression	91
5.1	Purpose	91
5.2	Linear models	92
5.2.1	Definition	92
5.2.2	Stationarity	93
5.2.3	Simulation	93
5.3	Fitted models	94
5.3.1	Model fitted to simulated data	94
5.3.2	Model fitted to the temperature series (1970–2005)	95
5.3.3	Autocorrelation and the estimation of sample statistics*	96
5.4	Generalised least squares	98
5.4.1	GLS fit to simulated series	98
5.4.2	Confidence interval for the trend in the temperature series	99
5.5	Linear models with seasonal variables	99
5.5.1	Introduction	99
5.5.2	Additive seasonal indicator variables	99
5.5.3	Example: Seasonal model for the temperature series	100
5.6	Harmonic seasonal models	101
5.6.1	Simulation	102
5.6.2	Fit to simulated series	103
5.6.3	Harmonic model fitted to temperature series (1970–2005)	105
5.7	Logarithmic transformations	109
5.7.1	Introduction	109
5.7.2	Example using the air passenger series	109
5.8	Non-linear models	113
5.8.1	Introduction	113
5.8.2	Example of a simulated and fitted non-linear series	113

5.9	Forecasting from regression	115
5.9.1	Introduction	115
5.9.2	Prediction in R	115
5.10	Inverse transform and bias correction	115
5.10.1	Log-normal residual errors	115
5.10.2	Empirical correction factor for forecasting means	117
5.10.3	Example using the air passenger data	117
5.11	Summary of R commands	118
5.12	Exercises	118
6	Stationary Models	121
6.1	Purpose	121
6.2	Strictly stationary series	121
6.3	Moving average models	122
6.3.1	MA(q) process: Definition and properties	122
6.3.2	R examples: Correlogram and simulation	123
6.4	Fitted MA models	124
6.4.1	Model fitted to simulated series	124
6.4.2	Exchange rate series: Fitted MA model	126
6.5	Mixed models: The ARMA process	127
6.5.1	Definition	127
6.5.2	Derivation of second-order properties*	128
6.6	ARMA models: Empirical analysis	129
6.6.1	Simulation and fitting	129
6.6.2	Exchange rate series	129
6.6.3	Electricity production series	130
6.6.4	Wave tank data	133
6.7	Summary of R commands	135
6.8	Exercises	135
7	Non-stationary Models	137
7.1	Purpose	137
7.2	Non-seasonal ARIMA models	137
7.2.1	Differencing and the electricity series	137
7.2.2	Integrated model	138
7.2.3	Definition and examples	139
7.2.4	Simulation and fitting	140
7.2.5	IMA(1, 1) model fitted to the beer production series	141
7.3	Seasonal ARIMA models	142
7.3.1	Definition	142
7.3.2	Fitting procedure	143
7.4	ARCH models	145
7.4.1	S&P500 series	145
7.4.2	Modelling volatility: Definition of the ARCH model	147
7.4.3	Extensions and GARCH models	148

7.4.4	Simulation and fitted GARCH model	149
7.4.5	Fit to S&P500 series	150
7.4.6	Volatility in climate series	152
7.4.7	GARCH in forecasts and simulations	155
7.5	Summary of R commands	155
7.6	Exercises	155
8	Long-Memory Processes	159
8.1	Purpose	159
8.2	Fractional differencing	159
8.3	Fitting to simulated data	161
8.4	Assessing evidence of long-term dependence	164
8.4.1	Nile minima	164
8.4.2	Bellcore Ethernet data	165
8.4.3	Bank loan rate	166
8.5	Simulation	167
8.6	Summary of additional commands used	168
8.7	Exercises	168
9	Spectral Analysis	171
9.1	Purpose	171
9.2	Periodic signals	171
9.2.1	Sine waves	171
9.2.2	Unit of measurement of frequency	172
9.3	Spectrum	173
9.3.1	Fitting sine waves	173
9.3.2	Sample spectrum	175
9.4	Spectra of simulated series	175
9.4.1	White noise	175
9.4.2	AR(1): Positive coefficient	177
9.4.3	AR(1): Negative coefficient	178
9.4.4	AR(2)	178
9.5	Sampling interval and record length	179
9.5.1	Nyquist frequency	181
9.5.2	Record length	181
9.6	Applications	183
9.6.1	Wave tank data	183
9.6.2	Fault detection on electric motors	183
9.6.3	Measurement of vibration dose	184
9.6.4	Climatic indices	187
9.6.5	Bank loan rate	189
9.7	Discrete Fourier transform (DFT)*	190
9.8	The spectrum of a random process*	192
9.8.1	Discrete white noise	193
9.8.2	AR	193

9.8.3	Derivation of spectrum	193
9.9	Autoregressive spectrum estimation	194
9.10	Finer details	194
9.10.1	Leakage	194
9.10.2	Confidence intervals	195
9.10.3	Daniell windows	196
9.10.4	Padding	196
9.10.5	Tapering	197
9.10.6	Spectral analysis compared with wavelets	197
9.11	Summary of additional commands used	197
9.12	Exercises	198
10	System Identification	201
10.1	Purpose	201
10.2	Identifying the gain of a linear system	201
10.2.1	Linear system	201
10.2.2	Natural frequencies	202
10.2.3	Estimator of the gain function	202
10.3	Spectrum of an AR(p) process	203
10.4	Simulated single mode of vibration system	203
10.5	Ocean-going tugboat	205
10.6	Non-linearity	207
10.7	Exercises	208
11	Multivariate Models	211
11.1	Purpose	211
11.2	Spurious regression	211
11.3	Tests for unit roots	214
11.4	Cointegration	216
11.4.1	Definition	216
11.4.2	Exchange rate series	218
11.5	Bivariate and multivariate white noise	219
11.6	Vector autoregressive models	220
11.6.1	VAR model fitted to US economic series	222
11.7	Summary of R commands	227
11.8	Exercises	227
12	State Space Models	229
12.1	Purpose	229
12.2	Linear state space models	230
12.2.1	Dynamic linear model	230
12.2.2	Filtering*	231
12.2.3	Prediction*	232
12.2.4	Smoothing*	233
12.3	Fitting to simulated univariate time series	234

12.3.1	Random walk plus noise model	234
12.3.2	Regression model with time-varying coefficients	236
12.4	Fitting to univariate time series	238
12.5	Bivariate time series – river salinity	239
12.6	Estimating the variance matrices	242
12.7	Discussion	243
12.8	Summary of additional commands used	244
12.9	Exercises	244
	References	247
	Index	249