

Contents

1	Why R, and why now?	1
2	An R tutorial for oceanographers	5
2.1	Introduction	6
2.2	Comparison of R and Matlab	7
2.3	First steps with R	8
2.3.1	License and availability	8
2.3.2	R packages	9
2.3.3	Running R	10
2.3.4	R working pattern	10
2.3.5	Getting help	11
2.4	Syntax	12
2.4.1	Expressions	12
2.4.2	Variables	14
2.4.3	Basic storage types	17
2.4.4	Vectors	26
2.4.5	Arrays and matrices	27
2.4.6	Lists	32
2.4.7	Factors	33
2.4.8	Data frames	35
2.4.9	Contingency tables	36
2.4.10	Conditional evaluation	37
2.4.11	Functions	38
2.4.12	Loops	42
2.4.13	Alternative to loops	43
2.5	Graphics	46
2.5.1	Scatter plots	46
2.5.2	Line plots	47

2.5.3	Contour plots	49
2.5.4	Image plots	50
2.5.5	Three-dimensional plots	50
2.5.6	Time-series plots	52
2.5.7	Box plots	53
2.5.8	Lagged autocorrelation plots	53
2.5.9	Histogram plots	54
2.5.10	Spectrum plots	55
2.5.11	Pairs plots	56
2.5.12	Conditioning plots	57
2.5.13	Class-specific plots	58
2.5.14	General hints on R graphics	59
2.5.15	Limitations of R graphics	60
2.6	Probability and Statistics	62
2.6.1	Probability	62
2.6.2	Statistical distributions	62
2.6.3	Summaries and overview functions	64
2.6.4	Hypothesis testing	65
2.6.5	Linear regression	68
2.6.6	Nonlinear regression	75
2.6.7	ANOVA	80
2.6.8	Partitioning decision Trees	82
2.7	Input and output	84
2.7.1	Reading from text files	84
2.7.2	Reading binary data	89
2.7.3	Reading databases	90
2.7.4	Reading NetCDF files	90
2.7.5	Writing files	93
2.8	Using R libraries	93
2.9	Techniques for debugging	94
3	The oce package	97
3.1	Data types and object orientation	98
3.2	Datasets	100
3.3	Functions	100
3.4	Example: world SST chart	104
3.5	Development scheme	105

4	Historical examples	107
4.1	The Redfield Ratio [Redfield 1934]	108
4.2	Ecosystem modelling [Riley 1946]	111
4.3	Plate tectonics [Wilson 1963]	116
4.4	Inferring ocean mixing [Munk, 1966]	119
5	Research operating procedures	125
5.1	Science in the laboratory, science in the field	126
5.2	Hydrographic data	127
5.2.1	CTD profiles	127
5.2.2	Hydrographic sections	132
5.2.3	Calculating water properties	134
5.2.4	Mixed layer detection	136
5.2.5	Density ratio calculation	140
5.2.6	Watermass detection	142
5.3	Acoustic-Doppler data	145
5.3.1	Reading acoustic-Doppler data	145
5.3.2	Trimming acoustic-Doppler data	146
5.3.3	Coordinate transformations	147
5.3.4	Plotting	148
5.4	Sea-level data	149
5.4.1	Time-domain analysis	149
5.4.2	Tidal analysis	149
5.4.3	Tidal atlases	151
5.5	Coastline and topographic data	151
5.6	Drifter data	152
5.7	Other data types	153
5.8	General analysis	154
5.8.1	Pruning faulty data	154
5.8.2	Dealing with log-normally distributed data	158
5.8.3	Time Series Analysis	159
5.8.4	Spatial mapping	169
5.8.5	Differential equations	172
5.8.6	Optimization	178
5.8.7	Eigenanalysis methods	180
6	Solutions	189
	Comparison between Matlab and R	225

Choosing an equation of state	229
B.1 History of modern equations of state	230
B.2 Installing TEOS-10	230
B.3 High-level use of TEOS-10	231
B.4 Low-level use of TEOS-10	232
High-performance calculations with R	233
C.1 Efficiency limitations of high-level languages	234
C.2 Deciding when to optimize	234
C.3 Avoiding loops	235
C.4 Supplementing R with compiled languages	236
C.5 Dealing with memory limitations	238
C.6 The analysis of algorithms	238
C.7 Multiple-processing calculations	240
C.8 Further reading	241
The future of R	243
References	247
Index	257