

Dennis L. Hartmann and Elizabeth A. Barnes

Data Analysis for Atmospheric, Oceanic and Climate Science

July 24, 2025

Springer

Contents

1	Introduction	1
2	Basic Statistical Techniques	3
2.1	Basic statistical quantities: means and other moments	3
2.1.1	The Mean	3
2.1.2	The Variance	3
2.1.3	The Standard Deviation	4
2.1.4	Higher Moments	4
2.2	Probability Concepts and Theorems	5
2.2.1	Unions and Intersections of Probability - Venn Diagram	5
2.2.2	Bayes Theorem	7
2.2.3	Probability philosophies: frequentist vs Bayesian views	8
2.3	Probability Distributions	9
2.3.1	The Normal Distribution	10
2.3.2	Other Common Distributions	12
2.4	Central Limit Theorem	13
2.5	Testing for Significance	14
2.5.1	Small sampling theory: the t-statistic	16
2.5.2	Confidence intervals	19
2.5.3	Chi-Squared Distribution: Tests of Variance	21
2.6	The Binomial Distribution	22
2.6.1	Binomial Distribution	22
2.6.2	Normal Approximation to the Binomial	23
2.7	The Poisson Distribution	25
2.8	Non-parametric Statistical Tests	27
2.8.1	Signs Test	27
2.8.2	Rank Sum Test	28
2.8.3	Runs Test (Wald-Wolfowitz Test)	29
2.8.4	Kolmogorov-Smirnov Test	30
2.9	Hypothesis Testing	31
2.9.1	Terminology and symbology	31
2.9.2	Setting-up the problem	31
2.9.3	Type I and Type II errors in hypothesis testing	32
2.9.4	<i>a priori</i> vs. <i>a posteriori</i>	32
2.9.5	Field significance and False Discovery Rate	33
2.10	Extreme Value Theory	34
2.10.1	Fisher-Tippett Theorem and Generalized Extreme Value Distribution	35
2.11	Monte Carlo and Resampling	36
2.11.1	Monte Carlo Techniques	36
2.11.2	Resampling via Bootstrapping	36

2.11.3	Resampling via Jackknife	38
3	Compositing and Superposed Epoch Analysis	41
3.1	Introduction	41
3.2	Steps in the Compositing Process	41
3.3	Evaluating compositing studies	42
3.4	Example: Daily Precipitation and Temperature	42
4	Regression	45
4.1	Ordinary linear least-squares regression	45
4.1.1	Independent variables are known	45
4.1.2	Independent and dependent variables are uncertain	46
4.1.3	Uncertainty estimates of ordinary least-squares regression	47
4.1.4	Other least-squares fits	48
4.2	Correlation	50
4.2.1	How good is the linear fit?	50
4.2.2	Sampling Theory of Correlation (Pearson's correlation)	53
4.3	Multiple Linear Regression	58
4.3.1	Generalized Normal Equations	58
4.3.2	Derivation of the Normal Equations using Matrix Notation	60
4.3.3	Multiple Regression - How many predictors should I use?	61
5	Seeking Structure in Data	65
5.1	Introduction	65
5.2	Data Sets as Two-Dimensional Matrices	66
5.3	Empirical Orthogonal Functions	67
5.3.1	Two-Dimensional Example	68
5.3.2	EOF/Principal Component Analysis - Introduction	68
5.4	Principal Components and EOFs	69
5.4.1	Orthogonality of the Principle Components	70
5.4.2	EOF Analysis via Singular Vector Decomposition	70
5.4.3	A very simple example	72
5.5	Presenting the Results of EOF and PC Analysis	73
5.5.1	How to scale and plot EOFs and PCs	73
5.6	Significance of EOF Analysis	74
5.6.1	The North Test	75
5.6.2	Assessing Physical Significance	75
5.7	Applications of EOF/PC Analysis	76
5.7.1	Data Compression	76
5.7.2	Determining Degrees of Freedom	77
5.7.3	Prefiltering	78
5.7.4	Statistical Prediction	78
5.7.5	Exploratory Data Analysis	79
5.8	Rotation of Empirical Orthogonal Functions	79
5.8.1	The Eight Physical Variables Example	80
5.8.2	EOF Analysis of Red Noise	81
5.8.3	Wintertime 500hPa Height Example	82
5.9	Maximum Covariance Analysis	84
5.9.1	MCA Mathematics	85
5.9.2	Normalized Root Mean Squared Covariance	86
5.9.3	Heterogeneous and Homogeneous Regression Maps	86
5.9.4	Statistical significance of MCA	87
5.10	Canonical Correlation Analysis	87
5.11	Cluster Analysis	89

5.11.1	k-means Clustering	89
5.11.2	Self-Organizing Maps (SOMs)	92
6	Mapping Data to a Grid and Data Assimilation	95
6.1	Placing data on a regular grid	95
6.1.1	Interpolation with polynomial fits	95
6.1.2	Optimum Interpolation	96
7	Time Series Analysis	101
7.1	Introduction	101
7.2	Autocorrelation and Red Noise	101
7.2.1	The Autocorrelation Function	101
7.2.2	White Noise	102
7.2.3	Red Noise	102
7.3	Statistical Prediction and Red Noise	104
7.4	Degrees of Freedom with Gaussian Red Noise	104
7.5	Harmonic Analysis and the Fourier Transform	106
7.6	The Power Spectrum	107
7.7	Methods of Computing Power Spectra	107
7.7.1	Direct Fourier Transform	107
7.7.2	Lag Correlation Method	107
7.8	The Complex Fourier Transform and Spectral Analysis	108
7.8.1	Parseval's Theorem	108
7.8.2	The Time Shifting Theorem	109
7.8.3	Lagged Covariance and the Power Spectrum	109
7.8.4	Example: The Power Spectrum of Red Noise	110
7.9	Data Windows and Window Carpentry	111
7.9.1	The Hanning Window	112
7.9.2	The Hamming Window	113
7.9.3	Welches Overlapping Segment Analysis: WOSA	113
7.10	Designing a Power Spectral Analysis	114
7.10.1	Bandwidth and Chunk Length	115
7.10.2	Time Step	115
7.10.3	Robustness and Degrees of Freedom	115
7.10.4	Example of 5-Day Wave	116
7.11	Statistical Significance of Spectral Peaks	116
7.11.1	Example: <i>a priori</i> versus <i>a posteriori</i> Spectral Peaks	117
7.11.2	Example: Statistical Design	118
7.11.3	The Red Noise Null Hypothesis	118
7.11.4	Continuous and Discrete Red Noise Spectra	119
7.11.5	Example: Sampling Red Noise	120
7.12	Prewhitening	121
7.12.1	Rossby-Gravity Waves in Reanalysis	121
7.13	Multi-Taper Method of Spectral Analysis	123
7.14	Maximum Entropy Spectral Analysis	123
7.15	Cross Spectrum Analysis	123
7.15.1	Complex Fourier Transform of Cross-Spectrum	124
7.15.2	Example: Rossby-Gravity Wave Cross-Spectral Analysis	125
7.16	Space-Time Spectrum Analysis	126
7.16.1	Standing Waves	127
7.16.2	Example of Space-Time Spectral Analysis	128

8	Filtering	131
8.1	Introduction	131
8.1.1	The Convolution Theorem	131
8.1.2	Parseval's Theorem	131
8.2	Filtering	132
8.2.1	Fourier Method	132
8.2.2	Centered, Non-recursive Filtering Method	133
8.2.3	Obtaining the Response Function	133
8.2.4	Simple Example of Cosine Wave	134
8.2.5	The Running Mean Smoother	135
8.2.6	Construction of Symmetric Non-recursive Filters	136
8.2.7	Frequency Response of Simple Filters	137
8.3	General Symmetric Non-recursive Filter Weights	137
8.3.1	Lanczos Smoothing of Filter Weights	139
8.4	Recursive Filters	140
8.4.1	Response Functions for General Linear Filters	141
8.4.2	A Simple Recursive Filter	141
8.4.3	Impulse Response of a Recursive Filter	142
8.4.4	Construction of Recursive Filters	142
8.4.5	Butterworth Filters	143
8.4.6	Example using Butterworth Filter	144
9	Wavelets	147
9.1	Introduction	147
9.2	Wavelet Types	147
9.3	The Haar Wavelet	148
9.4	Discrete Wavelet Transforms	150
9.5	The Pyramid Scheme of Discrete Wavelet Transforms	151
9.6	Daubechies Wavelet Filter Coefficients	153
9.7	Continuous, Non-orthogonal Wavelets	154
10	Appendix A	157
10.1	Coherence Probability Table	157
11	Appendix B	159
11.1	Matrix Algebra	159
	References	161

Chapter 1

Introduction

This book is intended for graduate students and professionals working in atmospheric and oceanic sciences or closely related fields.

It has been prepared by Professors Dennis L. Hartmann and Elizabeth A. Barnes with the idea of publishing it, but we here make it available for use by all.

It is very helpful if the reader has a background that includes elementary statistics, probability and matrix algebra, but an attempt is made to provide some basic background in these subjects, so that the book is largely self-contained.

The focus of the book is on using some standard methods for practical purposes intelligently, rather than providing a deep and complete theoretical analysis of these methods. Examples are shown of appropriate uses of the techniques in atmospheric and oceanic sciences.

The choice of topics reflects the experience of the authors in doing research in the atmospheric and oceanic sciences over an extended period. Some of the examples presented are of historical interest, while many of the techniques are very applicable today and in the future.

Most of the figures and examples in this book were done using Python and Jupyter Notebooks. The notebooks we used can be found in the GitHub page for this book.

