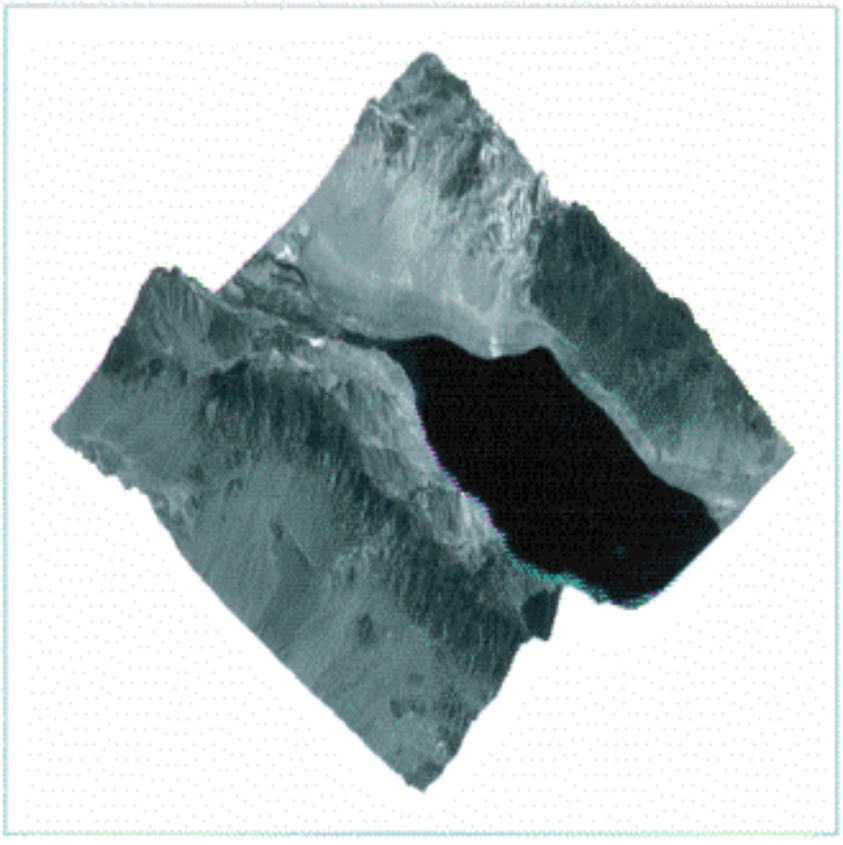
CD-ROM Included



DATA VISUALIZATION IN THE GEOSCIENCES

JAMES R. CARR

CONTENTS

Preface xiii

CHAPTER 1 1

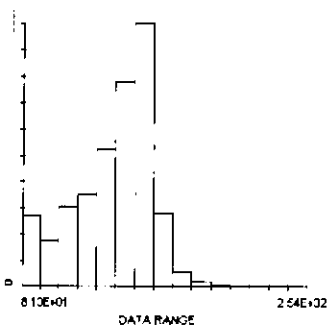
Introduction

1.1 **Purpose** 1
1.2 **Scope** 2
1.3 ***Nevada_Landsat* Data Set** 4
1.4 **Literature** 8

Exercises 8

CHAPTER 2 9

Univariate Data Analysis

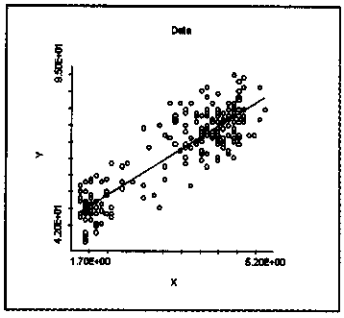


2.1 **Histograms** 9
2.1.1 Visualizing Univariate Data Through Statistical Parameters: A Summary 14
2.2 **Data Distributions: Probability Density Functions** 15
2.2.1 The Normal Distribution 16
2.2.2 Selected Comments About the Normal Distribution 17
2.2.3 Testing the Normal Probability Density for Goodness of Fit 17
2.3 **Statistical Hypothesis Testing** 19
2.3.1 The Chi-Square Test 19
2.3.1.1 Discussion 22
2.3.2 Hypothesis Test for the Probability Plot 23
2.3.3 So, Which Hypothesis Test Should Be Used? 23
2.4 **Other Distribution Functions and Suggested Applications** 24
2.4.1 The Poisson Distribution 24
2.4.2 The Exponential Distribution 24
2.4.3 The Log-Normal Distribution 26
2.5 **How Do I Reproduce Results in This Chapter, or Analyze My Own Data...** 26
2.5.1 ... Using *Visual_Data*? 26
2.5.2 ... Using *Microsoft Excel*? 27
2.5.3 ... Using MATLAB Student Version 5.3? 28
2.6 **Literature** 29

Exercises 29

CHAPTER 3 32

Bivariate Data Analysis



3.1 **Correlation 32**
- 3.1.1 The Method of Least Squares 32
  - 3.1.1.1 Linear Least Squares Regression 34
  - 3.1.1.2 Linear Regression, Type II 36
- 3.1.2 On the Adequacy of the Linear Regression Model 38
- 3.1.3 Hypothesis Testing of the Slope, M 40
  - 3.1.3.1 The F-Test 40
- 3.1.4 Correlation Coefficient 42
- 3.1.5 Covariance 43
- 3.1.6 On the Statistical Significance of the Correlation Coefficient 43

3.2 **Visual Regression 44**
- 3.2.1 Linear Models 44
  - 3.2.1.1 Least Absolute Deviation (LAD) Regression 45
  - 3.2.1.2 Comment: Least Squares or Least Absolute Deviation, Which Should be Used? 46
- 3.2.2 Analysis of Residuals to Infer the Need for Nonlinear Regression 47

3.3 **Application of Correlation Analysis to the *Nevada_Landsat* Data Set 49**

3.4 **How Do I Reproduce Results in This Chapter, or Analyze My Own Data... 54**
- 3.4.1 ... Using *Visual_Data?* 54
- 3.4.2 ... Using MATLAB 5.3? 55
- 3.4.3 ... Using *Microsoft Excel?* 56

3.5 **Weighted Regression 58**

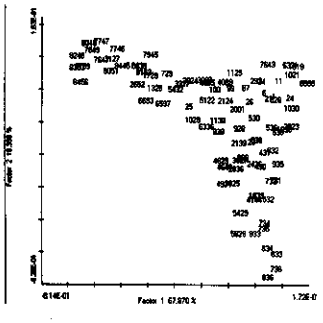
3.6 **A Review of the *Nevada_Landsat* Data Set 59**

3.7 **Literature 59**

Exercises 59

CHAPTER 4 61

Multivariate Data Analysis



4.1 **Analysis of Variance (ANOVA) 61**
- 4.1.1 Application to the *Nevada_Landsat_6x_Data* 63

4.2 **Statistical Hypothesis Test for Two Data Sets 63**

4.3 **Principal Components Analysis 65**
- 4.3.1 Principal Components 65
- 4.3.2 Standardized Principal Components Analysis 65
  - 4.3.2.1 Principal Components are Eigenvectors 66
  - 4.3.2.2 The Matrix Algebra 67
  - 4.3.2.3 Eigendecomposing the Correlation Matrix for the *Nevada_Landsat_6x_Data* 69
- 4.3.3 Correspondence Analysis 71
  - 4.3.3.1 Application to the *Nevada_Landsat* Data Set 74

4.3.3.2 Accommodating Missing Data in Correspondence Analysis 76

4.4 Multivariate, Linear Regression (Multiple Regression) 77

4.4.1 Logistic Regression: An Application of Multivariate, Linear Regression to the *Nevada_Landsat* Data Set 77

4.5 How Do I Reproduce Results in This Chapter, or Analyze My Own Data... 78

4.5.1 ... Using *Visual_Data?* 78

4.5.2 ... Using *Microsoft Excel?* 79

4.5.3 ... Using MATLAB 5.3? 80

4.6 A Summary of Analyses Thus Far Obtained of the *Nevada_Landsat* Data Set 83

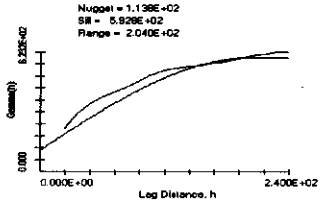
4.7 Literature 83

4.8 Final Thoughts on the Robustness of Principal Components Methods 83

Exercises 84

CHAPTER 5 85

Univariate Spatial Analysis



5.1 Autocorrelation 85

5.1.1 Introduction to Time Series Analysis 85

5.2 Spatial Autocorrelation 88

5.2.1 The Variogram 89

5.2.1.1 Application to the *Nevada_Landsat* Data Set 93

5.2.1.2 Directional Variogram Analysis 96

5.3 Fractal Geometry 96

5.4 Kriging: Spatial Interpolation as a Function of Spatial Autocorrelation 99

5.4.1 Covariance is Obtained from the Variogram 102

5.5 The Practice of Kriging 103

5.5.1 On the Normality of the Spatial Data 104

5.5.2 On Second Order Stationarity 104

5.5.3 On the Variogram Model 104

5.5.4 On the Design of the Grid 105

5.5.5 More on the Number of Nearest Neighbors used for Estimation 105

5.5.6 On the Size of the Search Window (And the Type of Search Strategy) 105

5.5.7 On the Concept of Sample Support: Punctual or Block 106

5.5.8 On the Need for a Data Transform 106

5.5.9 On Anisotropic Spatial Autocorrelation Modeling 108

5.6 Visualizations 108

5.7 Application to the *Nevada_Landsat_Data* 113

5.8 *M*-Kriging 117

5.8.1 Cross-Validation 119

5.9 How Do I Reproduce Results in This Chapter, or Analyze My Own Data 119

5.9.1 . . . Using *Visual_Data?* 119
5.9.1.1 Time Series Analysis 119
5.9.1.2 Variogram Analysis 121
5.9.1.3 Kriging 122
5.9.2 . . . MATLAB 125
5.9.2.1 Time Series Analysis 125
5.9.2.2 Variogram Analysis 126
5.9.2.3 Kriging 127
5.10 Literature 130
Exercises 130

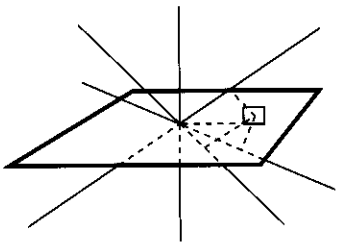
CHAPTER 6 — 132

Multivariate Spatial Data Analysis

6.1 Theory 132
6.1.1 Solving for the Cokriging Weights 133
6.1.1.1 The Matrix System: Kriging: 133
6.1.1.2 The Matrix System: Cokriging: 134
6.2 On the Practice of Cokriging 139
6.3 Autokrigeability 146
6.3.1 Application to Principal Components Images 146
6.3.2 Extension to Indicator Cokriging 147
6.4 The Undersampled Case 148
6.4.1 Accommodating Undersampling in Cokriging 149
6.4.2 Application to the *Nevada_Landsat* Data Set: Improving the Resolution of Thermal Images 149
6.5 How Do I Reproduce Results in This Chapter Using . . . 151
6.5.1 *Visual_Data?* 151
6.5.2 Using MATLAB? 155
6.5.3 Using *Microsoft Excel?* 156
6.6 Literature 157
6.7 Final Thoughts on Cokriging 157
Exercises 158

CHAPTER 7 — 159

Spatial Simulation



7.1 Random Numbers and Their Generation 159
7.2 One-Dimensional Spatial Simulation 162
7.3 Extension to Three-Dimensional Space: The Method of Random Lines 164
7.4 Simulation Using Fractals 165
7.5 Nonconditional Simulation: The Need for Data Transformation 167
7.6 Nonconditional Simulation of *Nevada_Landsat_Data* 168
7.6.1 Random Lines Algorithm 168
7.6.1.1 Visible Blue Reflectance 169
7.6.1.2 Thermal Emission 170
7.6.2 The Fractal Algorithm 172

7.7 **Conditioning the Simulation 173**
7.7.1. Application to the *Nevada_Landsat_Data* 173
7.7.1.1 Visible Blue Reflectance 174
7.7.1.2 Thermal Emissivity 174
7.8 **Why Is Spatial Simulation Useful? 175**
7.9 **How Do I Reproduce Results in This Chapter, or Experiment With My Own Data . . . 176**
7.9.1 . . . Using *Visual_Data?* 176
7.10 **Literature 178**
Exercises 179

CHAPTER 8 **180**

Digital Image Processing

8.1 **Pixels 180**
8.2 **Adjusting Pixel Contrast 181**
8.2.1 Algorithms for Contrast Adjustment 182
8.2.1.1 Simple, Linear 183
8.2.1.2 Multiple, Linear 185
8.2.1.3 Histogram Equalization 185
8.2.1.4 Binary Stretch 186
8.3 **Filtering Digital Images 188**
8.3.1 In a Digital Image, What Is a Low Frequency Feature? What Is a High Frequency Feature? 189
8.3.2 Spatial Convolution Filtering: Low-Pass, High-Pass, High-Boost, and Custom Strategies 190
8.3.2.1 Boosting Factor of 4 193
8.3.2.2 Boosting Factor of 8 193
8.3.2.3 East-West Custom High-Pass Filter (3x3) 193
8.3.2.4 North-South Custom High-Pass Filter (3x3) 193
8.4 **Principal Components Analysis of Multispectral Digital Images 194**
8.5 **Dust Devils on Mars 198**
8.5.1 True Color Compositing: The 24-Bit Bitmap 199
8.6 **How Do I Reproduce Results in This Chapter, or Analyze My Own Images . . . 203**
8.6.1 . . . Using *Visual_Data?* 203
8.6.1.1 File Format Conversion 206
8.6.1.2 Contrast Adjustment 207
8.6.1.3 Digital Filtering 208
8.6.1.4 Image Addition and Subtraction 210
8.6.2 . . . Using MATLAB 5.3 (Student Version)? 210
8.6.3 . . . With the Help of *Microsoft Excel?* 215
8.7 **Literature 217**
Exercises 217

CHAPTER 9 219

Composite Visualizations

9.1 **Multispectral Digital Image Compositing for Classification** 219
9.1.1 Classification of Texture 224
9.2 **Compositing a Contour Map With a Digital Image** 225
9.3 **Three-Dimensional Perspectives** 227
9.4 **Animation** 233
9.5 **How Do I Reproduce Results in This Chapter, or Pursue My Own Ideas . . .** 234
9.5.1 . . . Using *Visual_Data?* 234
9.5.2 . . . Using MATLAB? 237
9.6 **Literature** 239

Exercises 240

Epilogue 241
Why Be Normal? 241
Human Prerogative 241
Is Kriging the Best Spatial Interpolation Method? 243
The Robustness of Kriging 244
Appendices 247
A Critical Values of the Chi-Square Distribution 248
B Critical Values of Squared Correlation Coefficient, *p*-plot 251
C Critical Values of *F* Distribution 253
D Critical Values of *t* Distribution 257
Bibliography 259
Index 263