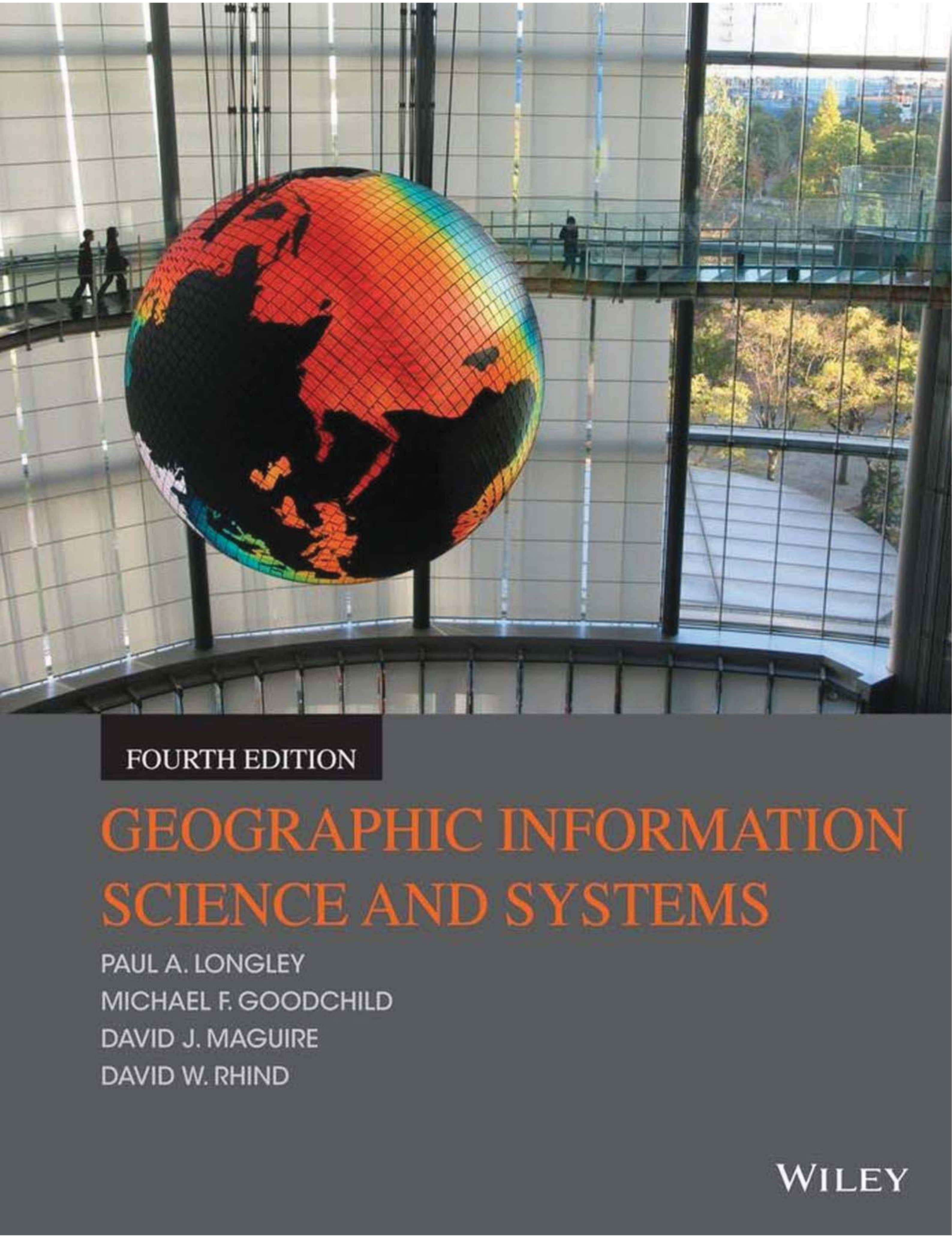


FOURTH EDITION

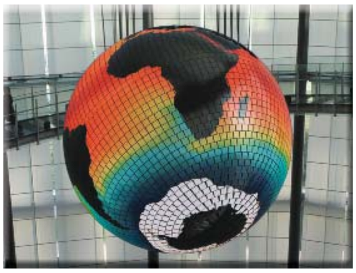
GEOGRAPHIC INFORMATION SCIENCE AND SYSTEMS

PAUL A. LONGLEY
MICHAEL F. GOODCHILD
DAVID J. MAGUIRE
DAVID W. RHIND

WILEY

GEOGRAPHIC INFORMATION SCIENCE & SYSTEMS

FOURTH EDITION


Paul Longley


Alex Singleton


© Eliza Snow/iStockphoto

PAUL A. LONGLEY
UNIVERSITY COLLEGE LONDON, UK

MICHAEL F. GOODCHILD
UNIVERSITY OF CALIFORNIA, SANTA BARBARA, USA

DAVID J. MAGUIRE
UNIVERSITY OF GREENWICH, UK

DAVID W. RHIND
CITY UNIVERSITY LONDON, UK

WILEY

VP & DIRECTOR	Petra Recter
EXECUTIVE EDITOR	Ryan Flahive and Jessica Fiorillo
SPONSORING EDITOR	Marian Provenzano
MARKETING MANAGER	Christine Kushner
ASSOCIATE PRODUCTION MANAGER	Joyce Poh
COVER DESIGN	Kenji Ngieng
FRONT COVER (PHOTO)	Paul Longley
BACK COVER PHOTO	Jack Dangermond

All photographs within this text that do not list a credit were either created by one of the authors or for which the author holds copyright.

This book was set in Avenir LT Std by Aptara®, Inc. and printed and bound by Courier Kendallville.

Founded in 1807, John Wiley & Sons, Inc. has been a valued source of knowledge and understanding for more than 200 years, helping people around the world meet their needs and fulfill their aspirations. Our company is built on a foundation of principles that include responsibility to the communities we serve and where we live and work. In 2008, we launched a Corporate Citizenship Initiative, a global effort to address the environmental, social, economic, and ethical challenges we face in our business. Among the issues we are addressing are carbon impact, paper specifications and procurement, ethical conduct within our business and among our vendors, and community and charitable support. For more information, please visit our website: www.wiley.com/go/citizenship.

Copyright © 2015, 2011 John Wiley & Sons, Inc. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923 (Web site: www.copyright.com). Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030-5774, (201) 748-6011, fax (201) 748-6008, or online at: www.wiley.com/go/permissions.

Evaluation copies are provided to qualified academics and professionals for review purposes only, for use in their courses during the next academic year. These copies are licensed and may not be sold or transferred to a third party. Upon completion of the review period, please return the evaluation copy to Wiley. Return instructions and a free of charge return shipping label are available at: www.wiley.com/go/returnlabel. If you have chosen to adopt this textbook for use in your course, please accept this book as your complimentary desk copy. Outside of the United States, please contact your local sales representative.

Library of Congress Cataloging in Publication Control Number: 2014040579

ISBN: 978-1-118-67695-0 (Paperback)

The inside back cover will contain printing identification and country of origin if omitted from this page. In addition, if the ISBN on the back cover differs from the ISBN on this page, the one on the back cover is correct.

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

CONTENTS

FOREWORD x

DEDICATION xi

PREFACE xii

LIST OF ACRONYMS xiv

Introduction

1 Geographic Information: Science, Systems, and Society 1

1.1 Introduction: What Are GI Science and Systems, and Why Do They Matter? 1

1.1.1 The Importance of Location 2

1.1.2 Spatial Is Special 6

1.2 Data, Information, Evidence, Knowledge, and Wisdom 9

1.3 GI Science and Systems 11

1.4 The Technology of Problem Solving 14

1.5 The Disciplinary Setting of GI Science and Systems (GISS) 16

1.5.1 The Historical Perspective 18

1.5.2 The Business Perspective 19

1.5.3 The Government Perspective 23

1.5.4 Computer-Science and Information-Science Perspectives 24

1.5.5. The Geography Perspective 25

1.5.6 The Societal Perspective 27

1.6 GI Science and Spatial Thinking 30

1.7 GI Systems and Science in Society 31

Questions for Further Study 32

Further Reading 32

1 Principles

2 The Nature of Geographic Data 33

2.1 Introduction 33

2.2 The Fundamental Problem 34

2.3 Spatial Autocorrelation and Scale 37

2.4 Spatial Sampling 39

2.5 Sampling and VGI 42

2.6 Distance Decay 43

2.7 Measuring Distance Effects as Spatial Autocorrelation 48

2.8 Taming Geographic Monsters 51

2.9 Induction and Deduction and How It All Comes Together 53

Questions for Further Study 54

Further Reading 54

3 Representing Geography 55

3.1 Introduction 55

3.2 Digital Representation 57

3.3 Representation of What and for Whom? 58

3.4 The Fundamental Problem 61

3.5 Discrete Objects and Continuous Fields 62

3.5.1 Discrete Objects 62

3.5.2 Continuous Fields 64

3.6 Rasters and Vectors 66

3.6.1 Raster Data 67

3.6.2 Vector Data 68

3.6.3 Representing Continuous Fields 68

3.7 The Paper Map 69

3.8 Generalization 71

3.8.1 Generalization about Places 71

3.8.2 Generalization about Properties 74

3.9 Conclusion 76

Questions for Further Study 76

Further Reading 76

4 Georeferencing 77

4.1 Introduction 77

4.2 Place-Names and Points of Interest 80

4.3 Postal Addresses and Postal Codes 82

4.4 IP Addresses 84

4.5 Linear Referencing Systems 84

4.6 Cadasters and the U.S. Public Land Survey System 85
4.7 Measuring the Earth: Latitude and Longitude 86
4.8 Projections and Coordinates 88
4.8.1 The Plate Carrée or Cylindrical Equidistant Projection 91
4.8.2 The Universal Transverse Mercator (UTM) Projection 92
4.8.3 Web Mercator 93
4.8.4 State Plane Coordinates and Other Local Systems 94
4.9 Measuring Latitude, Longitude, and Elevation: GPS 94
4.10 Converting Georeferences 95
4.11 Geotagging and Mashups 96
4.12 Georegistration 96
4.13 Summary 98
Questions for Further Study 98
Further Reading 98

5 Uncertainty 99
5.1 Introduction 99
5.2 U1: Uncertainty in the Conception of Geographic Phenomena 101
5.2.1 Conceptions of Place: Units of Analysis 101
5.2.2 Conceptions of Attributes: Vagueness and Ambiguity 104
5.2.3 Fuzzy Approaches to Attribute Classification 108
5.3 U2: Further Uncertainty in the Representation of Geographic Phenomena 111
5.3.1 Representation of Place/Location 111
5.3.2 Statistical Models of Uncertainty in Attribute Measures 112
5.3.3 Statistical Models of Uncertainty in Location Measures 117
5.4 U3: Further Uncertainty in the Analysis of Geographic Phenomena 117
5.4.1 Internal and External Validation through Spatial Analysis 117
5.4.2 Validation through Autocorrelation: The Spatial Structure of Errors 118
5.4.3 Validation through Investigating the Effects of Aggregation and Scale 122
5.4.4 Validation with Reference to External Sources: Data Integration and Shared Lineage 123
5.4.5 Internal and External Validation; Induction and Deduction 124
5.5 Consolidation 126
Questions for Further Study 127
Further Reading 127

2 Techniques

6 GI System Software 128
6.1 Introduction 128
6.2 The Evolution of GI System Software 129
6.3 Architecture of GI System Software 131
6.3.1 Project, Departmental, and Enterprise GI Systems 131
6.3.2 The Three-Tier Architecture 131
6.3.3 Software Data Models and Customization 133
6.3.4 GI Systems on the Desktop, on the Web, and in the Cloud 134
6.4 Building GI Software Systems 136
6.5 GI Software Vendors 137
6.5.1 Autodesk 137
6.5.2 Bentley 137
6.5.3 Esri 138
6.5.4 Intergraph 139
6.6 Types of GI Systems 140
6.6.1 Desktop Systems 140
6.6.2 Web Mapping Systems 142
6.6.3 Server GI Systems 142
6.6.4 Virtual Globes 146
6.6.5 Developer GI Systems 147
6.6.6 Mobile GI Systems 148
6.6.7 Other types of GI Software 149
6.7 Conclusion 150
Questions for Further Study 151
Further Reading 151

7 Geographic Data Modeling 152
7.1 Introduction 152
7.1.1 Data Model Overview 152
7.1.2 Levels of Data Model Abstraction 153
7.2 GI Data Models 154
7.2.1 CAD, Graphical, and Image Data Models 155
7.2.2 Raster Data Model 155
7.2.3 Vector Data Model 157
7.2.4 Object Data Model 164
7.3 Example of a Water-Facility Object Data Model 168
7.4 Geographic Data Modeling in Practice 170

Questions for Further Study 172
Further Reading 172

8 Data Collection 173

8.1 Introduction 173
8.1.1 Data Collection Workflow 175
8.2 Primary Geographic Data Capture 175
8.2.1 Raster Data Capture 175
8.2.2 Vector Data Capture 179
8.3 Secondary Geographic Data Capture 181
8.3.1 Raster Data Capture Using Scanners 181
8.3.2 Vector Data Capture 183
8.4 Obtaining Data from External Sources (Data Transfer) 187
8.4.1 Geographic Data Formats 189
8.5 Capturing Attribute Data 190
8.6 Citizen-Centric Web-Based Data Collection 190
8.7 Managing a Data Collection Project 191
Questions for Further Study 193
Further Reading 193

9 Creating and Maintaining Geographic Databases 194

9.1 Introduction 194
9.2 Database Management Systems 195
9.2.1 Types of DBMSs 196
9.2.2 Geographic DBMS Extensions 197
9.3 Storing Data in DBMS Tables 198
9.4 SQL 201
9.5 Geographic Database Types and Functions 202
9.6 Geographic Database Design 205
9.6.1 The Database Design Process 205
9.7 Structuring Geographic Information 206
9.7.1 Topology Creation 206
9.7.2 Indexing 208
9.8 Editing and Data Maintenance 212
9.9 Multiuser Editing of Continuous Databases 213
9.9.1 Transactions 213
9.9.2 Versioning 213
9.10 Conclusion 214
Questions for Further Study 216
Further Reading 216

10 The GeoWeb 217

10.1 Introduction 217
10.2 Distributing the Data 222
10.2.1 Object-Level Metadata 223
10.2.2 Geolibraries and Geoportals 225
10.3 The Mobile User 227
10.3.1 Virtual Reality and Augmented Reality 228
10.3.2 Location-Based Services 230
10.3.3 Issues in Mobile GIS 232
10.4 Distributing the Software: GI Services 233
10.4.1 Service-Oriented Architecture 234
10.5 Prospects 235
Questions for Further Study 236
Further Reading 236

3 Analysis

11 Cartography and Map Production 237

11.1 Introduction 237
11.2 Maps and Cartography 241
11.2.1 Maps and Media 245
11.3 Principles of Map Design 246
11.3.1 Map Composition 247
11.3.2 Map Symbolization 248
11.4 Map Series 257
11.5 Applications 261
11.6 Conclusion 265
Questions for Further Study 265
Further Reading 265

12 Geovisualization 266

12.1 Introduction: Uses, Users, Messages, and Media 266
12.2 Geovisualization, Spatial Query, and User Interaction 268
12.2.1 Overview 268
12.2.2 Spatial Query Online and the Geoweb 271
12.3 Geovisualization and Transformation 274
12.3.1 Overview 274
12.3.2 Cartograms 276
12.3.3 Remodeling Spatial Distributions as Dasymetric Maps 278
12.4 Participation, Interaction, Augmentation, and Dynamic Representation 280

12.4.1 Public Participation and Participatory GI Systems (PPGIS) 280
12.4.2 User Interaction and Representation in 2.5-D and 3-D 282
12.4.3 Handheld Computing and Augmented Reality 284
12.4.4 Visualizing Geotemporal Dynamics 285
12.5 Consolidation 288
Questions for Further Study 289
Further Reading 289

13 Spatial Data Analysis 290
13.1 Introduction: What Is Spatial Analysis? 290
13.1.1 Examples 292
13.2 Analysis Based on Location 295
13.2.1 Analysis of Attribute Tables 296
13.2.2 Spatial Joins 299
13.2.3 The Point-in-Polygon Operation 300
13.2.4 Polygon Overlay 301
13.2.5 Raster Analysis 303
13.3 Analysis Based on Distance 304
13.3.1 Measuring Distance and Length 304
13.3.2 Buffering 306
13.3.3 Cluster Detection 308
13.3.4 Dependence at a Distance 309
13.3.5 Density Estimation 310
13.3.6 Spatial Interpolation 313
13.4 Conclusion 317
Questions for Further Study 318
Further Reading 318

14 Spatial Analysis and Inference 319
14.1 The Purpose of Area-Based Analyses 319
14.1.1 Measurement of Area 319
14.1.2 Measurement of Shape 320
14.2 Centrality 321
14.2.1 Centers 322
14.2.2 Dispersion 324
14.3 Analysis of Surfaces 324
14.3.1 Slope and Aspect 324
14.3.2 Modeling Travel on a Surface 326
14.3.3 Computing Watersheds and Channels 327
14.3.4 Computing Visibility 328
14.4 Design 329
14.4.1 Point Location 330
14.4.2 Routing Problems 332
14.5 Hypothesis Testing 334
14.5.1 Hypothesis Tests on Geographic Data 335
14.6 Conclusion 337
Questions for Further Study 338
Further Reading 338

15 Spatial Modeling with GI Systems 339
15.1 Introduction 339
15.1.1 Why Model? 341
15.1.2 To Analyze or to Model? 342
15.2 Types of Models 343
15.2.1 Static Models and Indicators 343
15.2.2 Individual and Aggregate Models 343
15.2.3 Cellular Models 347
15.2.4 Cartographic Modeling and Map Algebra 349
15.3 Technology for Modeling 351
15.3.1 Operationalizing Models in GI Systems 351
15.3.2 Model Coupling 351
15.3.3 Cataloging and Sharing Models 352
15.4 Multicriteria Methods 352
15.5 Accuracy and Validity: Testing the Model 354
15.6 Conclusion 356
Questions for Further Study 357
Further Reading 357

4 Policy, Management, and Action

16 Managing GI Systems 358
16.1 Introduction 359
16.2 Managing Risk 359
16.3 The Case for the GI System: ROI 360
16.4 The Process of Developing a Sustainable GI System 366
16.4.1 Choosing a GI System: The Classical Acquisition Model 368
16.4.2 Implementing a GI System 373
16.4.3 Managing a Sustainable, Operational GI System 375
16.5 Sustaining a GI System—The People and Their Competences 378
16.5.1 GI System Staff and the Teams Involved 378
16.5.2 Project Managers 379
16.5.3 Coping with Uncertainty 379
16.6 Conclusions 380
Questions for Further Study 380
Further Reading 380

17 Information and Decision Making 381
17.1 Why We Need Information 381
17.1.1 Trade-Offs, Uncertainty, and Risk 383
17.1.2 Organizational Drivers 383
17.2 Information as Infrastructure 386
17.2.1 Information for Management 387
17.3 Different Forms of GI 391
17.3.1 GI about Individuals 394
17.3.2 More Novel Forms of GI 402
17.3.3 The Changing World of GI 402
17.4 Open Data and Open Government 404
17.4.1 The Metadata Issue 405
17.5 Example of an Information Infrastructure: The Military 406
17.5.1 Technological Change and the Military 406
17.5.2 The Military Information Infrastructure 407
17.5.3 Civilian Spin-Offs 408
17.6 Conclusions 409
Questions for Further Study 410
Further Reading 410

18 Navigating the Risks 411
18.1 Clashes Between Scientists and the Judiciary 412
18.2 Business Models for GI-Related Enterprises 412
18.3 Legal and Regulatory Constraints 414
18.3.1 Geography and the Law 414
18.3.2 Three Aspects of the Law and GI 415
18.4 Privacy and GI Systems 421
18.4.1 Preserving Privacy without Losing the Use of Personal Information 422
18.5 Public Trust, Ethics, and Coping with the Media 424
18.5.1 Public Trust 424
18.5.2 Ethics 425
18.5.3 Coping with the Media 426
18.6 Partnerships, Up-Scaling Activities, and Risk Mitigation 426
18.6.1 Spatial Data Infrastructures: The U.S. Experience 427
18.6.2 INSPIRE 429
18.6.3 UN Initiative on Global Geospatial Information Management 430
18.6.4 Have SDIs Worked? 430
18.7 Coping with Spatial Stupidity 432
18.8 Conclusions 433
Questions for Further Study 434
Further Reading 434

19 Epilog: GISS in the Service of Humanity 435
19.1 GISS, the Active Citizen, and Citizen Scientists 435
19.1.1 Who Can Help? 436
19.1.2 Areas Where GISS Contributes 437
19.2 Context: Our Differentiated World 437
19.3 Context: Our Interdependent World 440
19.4 The Process 441
19.4.1 Stage 1: Defining and Describing the Issue 442
19.4.2 Stage 2: Analyzing and Modeling Spatial Interrelationships 442
19.4.3 Stage 3: Devising Possible Solutions 442
19.4.4 Communicating Results and Possible Solutions to Decision Makers 443
19.4.5 Stage 5: Reflect, Learn, and Educate 443
19.5 The Grand Challenges 443
19.6 Grand Challenges Whose Effects We Can Help to Ameliorate 445
19.6.1 Population Growth 445
19.6.2 Poverty and Hunger 446
19.6.3 Human Health 448
19.6.4 Access to Food, Potable Water, and Boundary Disputes 452
19.6.5 Coping with Natural Disasters 453
19.6.6 Coping with Terrorism, Crime, and Warfare 456
19.6.7 Environmental Sustainability 456
19.7 Conclusions 459
Questions For Further Study 460
Further Reading 460

INDEX 461

Supplementary Materials Available Online

SM Supplementary Materials
Supplementary Materials 1
Powerpoint Slides
Instructor Manual

FOREWORD

Joe Lobley here, again.

This is the fourth time the "Gang of Four" authors have asked me to write an introduction to their textbook. I was flattered, so again said yes. Now I know why they keep on doing this: I heard from a Wiley insider that market research shows that my stuff is the most read section of the book! And who can resist having his great thoughts read by 80,000 purchasers (so far) of the book? But next time I'll charge the Gang.

Looking at publisher's blurb for the fourth edition, the first thing my eagle eyes picked up was the "science and systems" goof right there on the front cover—rather than the other way round used previously. I know you won't pulp the print run just because of this, but y'know the new title might just be a blessing in disguise. If you Google "GIS," I've noticed that references to my general infantrymen colleagues keep popping up on the list, so perhaps the term doesn't denote the sunrise industry it once was. Today's bright young wannabes (and old tight-fisted cheapskates like me) are more likely to patch together free and open software than toe the corporate software line. At the end of the day, I buy the line the Gang have spun since I first started helping them write, namely that science is more exciting than this month's favorite software release.

Which brings me to my news. Those who have followed my "most read" contributions will know that life in the GI "system garden" has not always been smooth for me. I've tried all sorts of roles, worked in many countries as a consultant, started businesses, smooched with governments, and got marooned on a desert island for my pains. Despite all my entrepreneurial activities, I'm still not rich. In fact, I'm broke. I'm living in a battered caravan in an alcohol-free Islamic country. Because I worked for the U.S. military for a time, where I am had better be secret 'til I raise enough cash to move on.

So I've been rethinking what's gone wrong, despite my unrivalled experience and scientific skills. Partly it's the structure of our industry. I've noticed that almost all the job adverts are for relatively lowly paid technical roles, and there are not many highly paid employees that are data bashers. I want to be one of the top guys, not a technician—I'm too old to keep up with techie college graduates when the GIS world is changing so rapidly. If putting science before systems presents new market opportunities, count me in, guys.

But at the end of the day, science just isn't where the real action is. When Calvin Coolidge was President, he said that "the business of America is business." So I've retrained: I've used my GIS to acquire a three-month Masters in intellectual property law from a respected online learning provider—my life-experience credits put me on the fast track from the start, and they accepted my successful patent filing for the Lobley Precisional Adjustment to differential GPS instead of a dissertation. The only problem was the huge fee I had to pay an agent to get certified as having passed everything. There have been so many big legal cases of late between Apple, Samsung, Google, and the rest over infringement of patents that I must be able to make it big in the "law and GIS" domain. If I had done it a bit earlier I could have sued one of the street data providers on behalf of users of their error-prone mapping. All I would have needed is for the families of a few people drowned after driving into a river by following these maps to ask me to act for them. OK, timing is everything.

I can only see one problem with GIS and law. It comes, as you might guess, from government. In the United States, government—apart from the military—mostly and until recently hasn't seen data as an asset to be treasured, protected, and exploited (I could help them). Worse, this plague is spreading. Can you believe that 60 or so national governments—including some serious ones (but not yet China or India, and Russia changed its mind)—have signed up to something called the Open Government Partnership? The idea is to flagellate themselves by making public commitments to reform government, foster innovation, and make everything transparent. Making almost all government data free seems to be the way that they will enable armchair auditors to keep watch on their government and politicians. This madness could be a serious barrier to my wealth creation if everything everywhere is free. But hey, maybe I could become a super-auditor, identifying fraud through use of GIS to bring data together. It would need to include lots of personal data, but privacy is an outdated concept anyway. My ex-wife Lolita found that out when I tracked her philandering throughout Lincolnshire some years ago.

All this, of course, is about Big Data—another fashion. We GIS folk have been doing it for years but no one has listened to us. As is normal with new fashions, big consultancies have proclaimed they are experts in it and can change the world. I could try giving them the benefit of my experience. But my best hope is to work for the U.S. National Security Agency or another country's version of it. Those guys—as Snowden pointed out—are focused, with clear aims, limited accountability, and lots of money. My kind of folk in fact. The bad guys have to live somewhere so the good guys need GIS. . . .

Joe Lobley

DEDICATION

We dedicate this fourth edition to Roger Tomlinson (1933–2014). Often called the "Father of GIS," Roger devoted most of his adult life to promoting the systems, technology, and science of geographic information (GI), as an integral part of the discipline of geography. In the 1960s he was the prime instigator behind the Canada Geographic Information System, a federal–provincial project to automate the measurement of Canada's land resource. In the 1970s he argued forcefully for a single, integrated technology for handling geographic information, completed a PhD at University College London, organized groundbreaking conferences through the aegis of the International Geographical Union, and founded a consulting practice to advise government agencies on the adoption of GI systems. His approach is ably detailed in his book, *Thinking about GIS: Geographic Information System Planning for Managers* (Esri Press), which is now in its fifth edition, and in the executive seminars he has led at the Esri International User Conference for many years.

Roger was an unflagging promoter of GI systems, which he saw as an essential part of humanity's interaction with its environment and the key to the solution of many of humanity's problems. He will be remembered for the force of his personality, his wit and charm, and his passionate support of the field, which he did more than perhaps anyone else to establish and support.

PREFACE

It is an old but true adage that everything that happens, happens somewhere. Throughout the history of humankind, geography has played a central role in many types of decision-making, some of which have life or death, or at least major strategic, impacts. In the past 50 years decision-making has benefited enormously, and in very many ways, from access to geographic information (GI), the science that underpins it, and the systems technology that enables it.

The previous edition of this textbook was published in 2011. Since then our world has changed, in some respects dramatically. Many of our interactions with information now occur through mobile devices rather than desktops, laptops, or paper. Location- (i.e., geographic-) based services have been estimated to be worth between $150bn and $270bn annually. Open Data, Open Software, and Open Science have been developing rapidly. The emergence of Big Data—where our community has pioneered many developments—has been hailed by some as obviating many past constraints (such as ensuring that samples are representative of a known population). Virtually all data are now collected in digital form rather than on paper; it is claimed that more data are now collected every two years than in the whole of previous human history. Crowdsourcing has produced many new datasets and changes in the way we tackle some tasks—such as scanning satellite images of a huge area of the South Indian Ocean for wreckage from Malaysian Airlines MH370 flight, a project organized by DigitalGlobe using imagery from its Worldview-2 system. Many governments are at last disgorging the information they hold for general use. And social media data are providing the fuel for real-time analysis of the geotemporal activity patterns of hundreds of millions of citizens. Given all that, this edition attempts to identify, explain, and evaluate the key changes and portray a snapshot of the contemporary world of geographic information, GI science, and GI systems.

In times past we wrote about geographic information systems, or GIS. The world has moved on. Except where we are quoting from others, we no longer use the abbreviation GIS. GI systems continue to evolve rapidly in their functions, ease of use, and number and spread of their users. They continue to provide the tools to describe and analyze the physical or human environments, bringing together data and converting them into information and even evidence (see Section 1.2). But underpinning that use of dazzling new technologies is a rapidly developing GI science. Here we deal with principles, many of which have endured in changing guises ever since the first edition of this book appeared in 2001. Where they exist, we deal with laws akin to those in the physical sciences, but also address the statistical generalizations of the social and environmental sciences. The third driving force of our Gang of Four is geographic information itself: we need to know its many characteristics, including quality, if we are to accommodate the inevitable uncertainty that arises when we admix different data using a variety of algorithms.

The New Vision

Reflecting this emerging GI ecosystem, we have made a subtle change of title in this, the fourth edition. The internal structure and content of the book reflects the change. After an introductory chapter, we develop a section on principles. This encompasses the nature of geographic data and information, representing geography, georeferencing, and uncertainty. We follow this with the "how"—a section on techniques, dealing with GI system software, data modeling, data collection, creating and maintaining geographic databases, and the Geoweb. The fourth section on analysis covers cartography, geovisualization, spatial data analysis, inferential spatial analysis, and spatial modeling. The fifth section covers human factors in relation to what we now term geographic information science and systems (GISS). It deals with information and decision-making, and with navigating the legal, ethical, and many other risks that GISS practitioners face. The concluding chapter—the Epilog—looks ahead. But it does this not by seeking to assess technological change, important as that is. Rather, it seeks to identify where we can use our GISS understanding, knowledge, skills, and tools to tackle major problems.

Throughout the book we emphasize the commonalities and the differences between groups of GI system users. Thus those in business, in governments at a variety of levels, in academia, and in not-for-profit organizations have overlapping concerns but some different drivers. This extends to differences between national and subnational cultures (and even between individuals), where our value systems and preferred modes of operating vary greatly. We have tried to give due credence to these similarities and differences.

Throughout the book we use examples and descriptions of luminaries whom we judge to have made a substantial contribution. We have tried throughout the text to provide detail because "the devil is in the detail" while also trying to highlight key points (such as through use of short tweet-like "factoids" that appear in bold), further reading, and a set of questions at the end of each chapter to test how much the student has gained from it and whether the student can develop new ideas or practice.

Online Supplementary Materials

This fourth edition is available both in print and online. In addition to the full content of the print edition, the online Web site includes significant supplementary material:

- A detailed discussion of four examples of GI system application, chosen to illustrate both the breadth of applications of GI technology, and the importance of the scientific principles elaborated throughout the book.
- Powerpoint slides for each of the chapters of the book, designed to be used as the basis for a course of lectures on the book's contents.
- An Instructor's Manual, giving pointers to the most effective ways to use the book in courses.

The Best of Times

In short, we are in the most exciting of times. Human ingenuity is transforming the way we can describe, analyze, and communicate what is occurring on the face of the Earth (and beyond). We have good enough science, information, and tools to make a real impact in improving societies, business performance, and much else—at all levels from the very local to the global. Central to all this is geographic variation and the awareness and skills to cope with it or even to reshape it. We authors are excited by what GISS practitioners have already achieved and by the prospects for the future. This book seeks to tell you why and convince you to join us.

Acknowledgments

We take complete responsibility for all the material contained herein. But we have drawn on contributions made by many others from across the world, which have helped us create this edition. We especially thank Muhammad Adnan, Stuart Ashfield, Brian Baker, Lawrence Ball, Sir John Beddington, Ann Blake, James Borrell, Peter Benton, Budhu Bhaduri, Paul Boyle, Alisdair Calder, Folke Carlsson, Tao Cheng, James Cheshire, Keith Clarke, Steve Coast, David Cowen, Max Craglia, Jack Dangermond, Keith Dugmore, Sarah Elwood, Ryan Flahive, Chris Gale, Vishal Gaudhar, Paul Hancock, David Hand, Stuart Houghton, Liz Hughes, Indy Hurt, Pete Jones, Jens Kandt, Milan Konecny, Kira Kowalska, Guy Lansley, Vanessa Lawrence, Alistair Leak, Dan Lewis, Antonio Lima, Yu Liu, Ross Maciejewski, Sara McLafferty, Mirco Musolesi, Tomoki Nakaya, Ollie O'Brien, Giles Pavey, Joyce Poh, Marian Provenzano, Denise Pumain, Muttukrishnan "Raj" Rajarajan, Jonathan Rhind, Doug Richardson, James Russiello, Matt Sims, Alex Singleton, Karen Slaght, Carl Steinitz, Lynn Usery, Ruth Wembridge, Roy Wood, Michael Worboys, Bradley Williams, Qingling Wu, and Keiji Yano. We sincerely apologize to anyone else whose contribution we may have overlooked.

Many others have contributed to the content of previous editions; we reiterate our thanks to them.

Paul Longley, University College London
Michael Goodchild, Emeritus Professor, University of California, Santa Barbara
David Maguire, University of Greenwich
David Rhind, Emeritus Professor, City University London

List of Acronyms and Abbreviations

AAG	Association of American Geographers
ABM	agent-based model
AEGIS	Advanced Emergency Geographic Information System
AHP	Analytical Hierarchy Process
AJAX	Asynchronous Javascript and XML
ALSM	airborne laser swath mapping
AM/FM	automated mapping/facilities management
AOL	America On-Line
API	application programming interface
AR	augmented reality
ARPANET	Advanced Research Projects Agency Network
ASCII	American Standard Code for Information Interchange
AVHRR	Advanced Very High Resolution Radiometer
AVIRIS	Airborne Visible Infrared Imaging Spectrometer
BLOB	binary large object
BS	Bachelor of Science
CA	cellular automaton
CAD	computer-assisted design
CAMS	Capacity Area Management System (Sears)
CARS	Computer-Aided Routing System (Sears)
CAS	Chinese Academy of Sciences
CASE	computer-aided software engineering
CCTV	closed-circuit television
CD	compact disk
CDR	carbon dioxide removal
CEN	Comité Européen de Normalisation
CERCO	Comité Européen de Responsibles de la Cartographie Officielle
CERN	European Organization for Nuclear Research
CGIS	Canada Geographic Information System
CGS	Czech Geological Survey
CIA	Central Intelligence Agency (US)
CODATA	Committee on Data for Science and Technology (International Council for Science)
COGO	coordinate geometry
COM	component object model
COTS	commercial off-the-shelf
CPI	consumer price index
CPU	central processing unit
CSAIL	Laboratory for Computer Science and Artificial Intelligence (MIT)
CSDGM	Content Standards for Digital Geospatial Metadata
CSV	comma-separated values
DBA	database administrator
DBMS	database management system
DCL	Data Control Language (SQL)
DCM	digital cartographic model
DDL	Data Definition Language (SQL)
DEM	digital elevation model
DGPS	Differential GPS
DIG	Decentralized Information Group (MIT)
DIME	dual independent map encoding
DLG	digital line graph
DLM	digital landscape model
DMI	distance measuring instrument
DML	Data Manipulation Language (SQL)
DNA	deoxyribonucleic acid
DoD	Department of Defense (US)
dpi	dots per inch
DRASTIC	model named for its inputs: depth, recharge, aquifer, soils, topography, impact, conductivity
DRG	digital raster graphic
DVD	digital video disk
DWD	National Meteorological Service (Germany)
EC	European Commission
ECU	Experimental Cartography Unit
EDA	exploratory data analysis
EL.STAT	Hellenic Statistical Agency (Greece)
EOS	Earth Observing System
EOSDIS	Earth Observing System Data and Information System
EPA	Environmental Protection Agency (US)
EPSG	European Petroleum Study Group
ERDAS	Earth Resource Data Analysis System
ESDA	exploratory spatial data analysis
Esri	Environmental Systems Research Institute
Esri BIS	Esri Business Information Solutions
EU	European Union
ExCiteS	Extreme Citizen Science (University College London)
FEMA	Federal Emergency Management Agency (US)
FGDC	Federal Geographic Data Committee (US)
FOIA	freedom of information act
FOSS4G	Free and Open-Source Software for Geospatial
FSA	Forward Sortation Area
GA	genetic algorithm
GAO	Government Accountability Office (US)
GDP	gross domestic product
GDT	Geographic Data Technology Inc.
GEOINT	geospatial intelligence
GFIS	Geographic Facilities Information System (IBM)
GGIM	(Initiative on) Global Geospatial Information Management (UN)
GI	geographic information
GIF	Graphics Interchange Format
GIS	geographic information system
GISS	geographic information science and systems
GIST	Geographic Information Science and Technology group (ORNL)

GIS-T	geographic information systems for transportation
GITA	Geospatial Information and Technology Association
GLONASS	Global Orbiting Navigation Satellite System
GML	Geography Markup Language
GPS	Global Positioning System
GRASS	Geographic Resources Analysis Support System
GSDI	global spatial data infrastructure
GSN	Global Spatial Network
GUI	graphical user interface
GWR	geographically weighted regression
HIV-AIDS	Human Immunodeficiency Virus - Acquired Immune Deficiency Syndrome
HLS	hue, lightness, saturation
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
HUMINT	human intelligence
IARPA	Intelligence Advanced Research Projects Activity
IBRU	International Boundaries Research Unit
ICSU	International Council for Science
ICT	information and communication technology
ID	identifier
IDE	integrated developer environment
IDW	inverse-distance weighting
IGN	Institut Géographique National
IJDE	International Journal of Digital Earth
IM	Instant Messenger
INPE	Instituto Nacional de Pesquisas Espaciais (Brazil)
INSPIRE	Infrastructure for Spatial Information in the European Community (Europe)
IP	Internet Protocol
IPCC	Intergovernmental Panel on Climate Change
IPR	intellectual property rights
ISDE	International Society for Digital Earth
ISO	International Organization for Standardization
IT	information technology
ITS	intelligent transportation systems
ITT	invitation to tender
JPEG	Joint Photographic Experts Group
JPL	Jet Propulsion Laboratory
KML	Keyhole Markup Language
LAN	local-area network
LBS	location-based service
LDO	Local Delivery Office
LiDAR	light detection and ranging
LIESMARS	State Key Laboratory for Information Engineering in Surveying, Mapping, and Remote Sensing (China)
LMIS	Land Management Information System (South Korea)
MAT	(point of) minimum aggregate travel
MAUP	Modifiable Areal Unit Problem
MBR	minimum bounding rectangle
MCDM	multi-criteria decision making
MDGs	Millennium Development Goals
MER	minimum enclosing rectangle
MGCP	Multinational Geospatial Co-Production Program
MIDI	Musical Instrument Digital Interface
MIT	Massachusetts Institute of Technology
MOCT	Ministry of Construction and Transportation (South Korea)
MODIS	Moderate Resolution Imaging Spectroradiometer
MOOC	massive open online course
MP3	MPEG Audio Layer III
MPEG	Motion Picture Experts Group
MrSID	Multiresolution Seamless Image Database
MSC	Mapping Science Committee (US National Research Council)
NAD27	North American Datum of 1927
NAD83	North American Datum of 1983
NASA	National Aeronautics and Space Administration
NATO	North Atlantic Treaty Organization
NCGIA	National Center for Geographic Information and Analysis (US)
NGA	National Geospatial-Intelligence Agency (US)
NGO	non-governmental organization
NII	national information infrastructure
NIMA	National Imagery and Mapping Agency (US)
NIMBY	not in my backyard
NLS	National Land Survey
NMCA	national mapping and charting agency
NMO	national mapping organization
NMP	National Mapping Program
NOAA	National Oceanic and Atmospheric Administration (US)
NPWS	National Parks and Wildlife Service (Australia)
NSDI	National Spatial Data Infrastructure (US)
NSF	National Science Foundation (US)
OAC	Output Area Classification (UK Office of National Statistics)
OAS	Organization of American States
OCR	optical character recognition
OD	Open Data
ODBMS	object database management system
OGC	Open Geospatial Consortium
OGL	Open Government License (UK)
OGP	Open Government Partnership
OLM	object-level metadata
OLS	ordinary least-squares
OMB	Office of Management and Budget (US)
ORDBMS	object-relational database management system
ORNL	Oak Ridge National Laboratory
OS	Ordnance Survey (Great Britain, or Northern Ireland)
OSINT	open-source intelligence

OSM	Open Street Map
PAIGH	PanAmerican Institute of Geography and History
PAF	Postal Address File
PARC	Palo Alto Research Center (Xerox)
PB	petabyte
PC	personal computer
PCC	percent correctly classified
PCRaster	Personal Computer Raster (GIS)
PDA	personal digital assistant
PDF	Portable Document Format
PERT	Program Evaluation and Review Technique
PGIS	participatory geographic information systems
PLSS	Public Land Survey System
PPGIS	public-participation geographic information systems
PROTECT	Port Resilience for Operational/Tactical Enforcement to Combat Terrorism (US Coast Guard)
PSI	public-sector information
QA	quality assurance
QR	quick response (code)
RADI	Institute of Remote Sensing and Digital Earth (Chinese Academy of Sciences)
R&D	research and development
RDBMS	relational database management system
RDFa	Resource Description Framework in Attributes
REST	Representation State Transfer Protocol
RFI	request for information
RFID	radio frequency identification
RFP	request for proposals
RGB	red, green, blue
RGS	Royal Geographical Society (UK)
RMSE	root mean squared error
ROI	return on investment
RS	remote sensing
RSS	Rich Site Summary or Really Simple Syndication
SDE	Spatial Database Engine
SDI	spatial data infrastructure
SDSS	spatial decision support system
SDTS	Spatial Data Transfer Standard
SETI	search for extra-terrestrial intelligence
SIGINT	signals intelligence
SOA	service-oriented architecture
SOAP	Simple Object Access Protocol
SOHO	small office/home office
SPC	State Plane Coordinates
SPOT	Système Probatoire d'Observation de la Terre
SQL	Structured (or Standard) Query Language
SQL/MM	Structured (or Standard) Query Language/ Multimedia
SRM	solar radiation management
SWMM	Storm Water Management Model
SWOT	strengths, weaknesses, opportunities, threats
TB	terabyte
TIFF	Tagged Image File Format
TIGER	Topologically Integrated Geographic Encoding and Referencing
TIN	triangulated irregular network
TOID	topographic identifier
TSP	traveling-salesperson problem
TV	television
UAM	Metropolitan Autonomous University (Mexico)
UAV	unmanned aerial vehicle
UCAS	University of the Chinese Academy of Sciences
UCGIS	University Consortium for Geographic Information Science (US)
UK	United Kingdom (of Great Britain and Northern Ireland)
UML	Unified Modeling Language
UN	United Nations
UNAM	National Autonomous University of Mexico
UNIGIS	University GIS Consortium
URI	uniform resource identifier
URL	uniform resource locator
US	United States (of America)
USA	United States of America
USGS	United States Geological Survey
USLE	Universal Soil Loss Equation
UTM	Universal Transverse Mercator projection
VACCINE	Visual Analytics for Command, Control, and Interoperability Environments (Purdue University)
VBA	Visual Basic for Applications
VfM	value for money
VGA	video graphics array
VGI	volunteered geographic information
VR	virtual reality
W3C	World Wide Web Consortium
WAN	wide-area network
WCS	Web Coverage Service
WFS	Web Feature Service
WGS84	World Geodetic System of 1984
WHO	World Health Organization
WIMP	windows, icons, menus, pointers
WMS	Web Map Service
WTO	World Trade Organization
WWF	World Wide Fund for Nature
WWW	World Wide Web
WYSIWYG	what you see is what you get
XML	Extensible Markup Language
XSEDE	Extreme Science and Engineering Discovery Environment

1

Geographic Information: Science, Systems, and Society

This chapter sets the conceptual framework for and summarizes the content of the book by addressing several major questions:

- What exactly is geographic information (GI), and why is it important? What is special about it?
- What new technological developments are changing the world of GI?
- How do GI systems affect the lives of average citizens?
- What kinds of decisions make use of geographic information?
- What is a geographic information system (GI system), and how would you recognize one?
- What is geographic information science (GI science), and why is it important to GI systems?
- How do scientists and governments use GI systems, and why do they find them helpful?
- How do companies make money from GI systems?

LEARNING OBJECTIVES

After studying this chapter you will:

- Know definitions of many of the terms used throughout the book.
- Be familiar with a brief history of GI science and GI systems.
- Recognize the sometimes invisible roles of GI systems in everyday life, business, and government.
- Understand the significance of GI science and how it relates to GI systems.
- Understand the many impacts that GI systems and its underpinning science are having on society and the need to study those impacts.

1.1 Introduction: What Are GI Science and Systems, and Why Do They Matter?

Almost everything that happens, happens somewhere. We humans confine our activities largely to the surface and near-surface of the Earth. We travel over it and through the lower levels of its atmosphere, and we go through tunnels dug just below the surface. We dig ditches and bury pipelines and cables, construct mines to get at mineral deposits, and drill wells to access oil and gas. We reside on the Earth and interact with others through work, leisure, and family pursuits. Keeping track of all this activity is important, and knowing where it occurs can be the most convenient basis for tracking. Knowing where something happens is of critical importance if we want to go there ourselves or send someone there, to find more information about the same place, or to inform people who live nearby. In addition, geography shapes the range of options that we have to address things that happen, and once they are made, decisions have geographic consequences. For example, deciding the route of a new high-speed railroad may be shaped by topographic and environmental considerations, and the chosen route will create geographic winners and losers in terms of access. Therefore geographic location is an important component of activities, policies, strategies, and plans.

> **Almost everything that happens, happens somewhere. Knowing where something happens can be critically important.**

The focus of this book is on geographic information, that is, information that records *where* as well as *what* and perhaps also *when*. We use the abbreviation *GI* throughout the book. GI systems were originally conceived as something separate from the world they represent—a special kind of information system, often located on a user's desk, dedicated to performing special kinds of operations related to location. But today such information pervades the Internet, can be accessed by our smartphones and other personal devices, and is fundamental to the services provided by governments, corporations, and even individuals. Locations are routinely attached to health records, to Twitter feeds and photographs uploaded to Flickr, and to the movements of mobile phone users and vehicles. In a sense, then, the whole digital world has become one vast, interconnected GI system. This book builds on what users of this system already know—that use of GI services is integral to many of our interactions through the Internet. Later chapters will describe, for example, how storage and management of more and more data entail use of the Cloud, how Big Data and Open Data have become ubiquitous (but not necessarily useful), and how Web-based GI systems have become a fact of life.

Underlying these changes are certain fundamentals, however, and these have a way of persisting despite advances in technology. We describe them with the term *GI science*, which we define as the general knowledge and important discoveries that have made GI systems possible. GI science provides the structure for this book because as educators we believe that knowledge of principles and fundamentals—knowledge that will still be valid many years from now—is more important than knowledge of the technical details of today's versions of GI technology. We use the acronym *GISS*—geographic information science and systems—at various points in this book to acknowledge the interdependence between the underpinning science and the technology of problem solving.

At the outset, we also observe that GI science is also fundamentally concerned with solving applied problems in a world where business practices, or the realpolitik of government decision making, are important considerations. We also discuss the practices of science and social science that, although governed by clearly defined scientific principles, are imperfectly coupled in some fast-developing areas of citizen science.

1.1.1 The Importance of Location

Because location is so important, it is an issue in many of the problems society must solve. Some of these problems are so routine that we almost fail to notice them—the daily question of which route to take to and from work, for example. Others are quite extraordinary and require rapid, concerted, and coordinated responses by a wide range of individuals and organizations—such as responding to the major emergencies created by hurricanes or earthquakes (see Box 1.1). Virtually all aspects of human life involve location. Environmental and social scientists recognize the importance of recording location when collecting data; major information companies such as Google recognize the importance of providing mapping and driving directions and prioritizing searches based on the user's location; and citizens are increasingly familiar with services that map the current positions of their friends. Here are some examples of major decisions that have a strong geographic element and require GI:

- Health-care managers decide where to locate new clinics and hospitals.
- Online shopping companies decide the routes and schedules of their vehicles, often on a daily basis.
- Transportation authorities select routes for new highways and anticipate their impacts.
- Retailers assess the performance of their outlets and recommend how to expand or rationalize store networks.
- Forestry companies determine how best to manage forests, where to cut trees, where to locate roads, and where to plant new trees.
- National park authorities schedule recreational path creation, maintenance, and improvement (Figure 1.1).
- Governments decide how to allocate funds for building sea defenses.
- Travelers and tourists give and receive driving directions, select hotels in unfamiliar cities, and find their way around theme parks (Figure 1.2).
- Farmers employ new GI technology to make better decisions about the amounts of fertilizer and pesticides to apply to different parts of their fields.

If location and GI are important to the solution of so many problems, what distinguishes those problems from each other? Here are three bases for classifying problems. First, there is the question of *scale*, or level of geographic detail. The architectural design of a building involves GI, but only at a very detailed or local scale. The information needed to service the building is also local—the size and shape of the

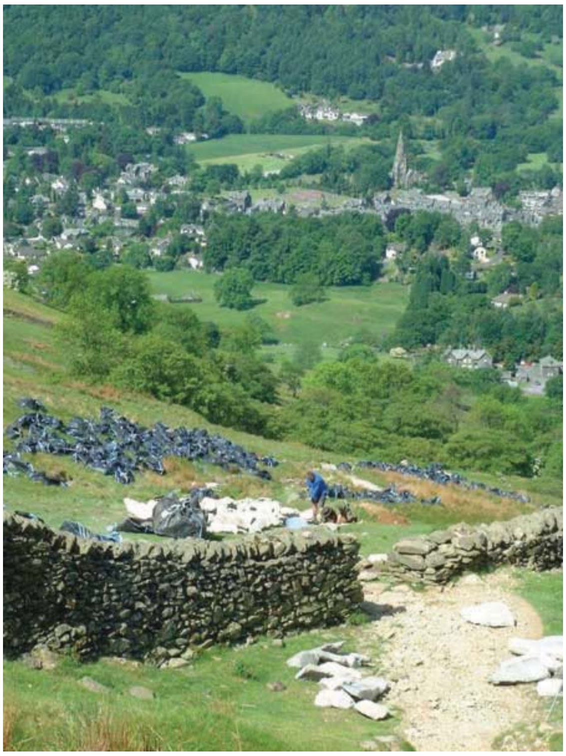


Figure 1.1 Maintaining and improving footpaths in national parks is a geographic problem.

parcel, the vertical and subterranean extent of the building, the slope of the land, and its accessibility using normal and emergency infrastructure. At the other end of the scale range, the global diffusion of epidemics and the propagation of tsunamis across the Pacific Ocean (Box 1.1) are phenomena at a much broader and coarser scale.

> **Scale or level of geographic detail is an essential property of any project.**

Second, problems can be distinguished on the basis of *intent*, or *purpose*. Some problems are strictly practical in nature—they must often be solved as quickly as possible and at minimum cost to achieve such practical objectives as saving lives in an emergency, avoiding fines by regulators, or responding to civil disorder. Others are better characterized as driven by human curiosity. When GI is used to verify the theory of continental drift, to map distributions of glacial deposits, or to analyze the historic movements of people in anthropological or biosocial research (see Box 1.2 and Figure 1.5), there is no sense of an immediate problem that needs to be solved. Rather, the intent is to advance human understanding of the world, which we often recognize as the intent of science.

Although science and practical problem solving can be thought of as distinct human activities, it is often argued that there is no longer any effective distinction between their methods. Many of the tools and methods used by a retail analyst seeking a site for a new store are essentially the same as those used by a scientist in a government agency to ensure the protection of an endangered species, or a transport planner trying to ameliorate peak-hour traffic congestion in a city. Each requires the most accurate measurement devices, employs terms whose meanings have been widely shared and agreed on, produces results that are replicable by others, and in general follows all the principles of science that have evolved over the past centuries. The knowledge-exchange activities carried out between research organizations and the government and business sectors can be used to apply many of the results of curiosity-driven science to the practical world of problem solving.

Figure 1.2 Navigating tourist destinations is a geographic problem.

The use of GI systems in support of science, routine application, and knowledge exchange reinforces the idea that science and practical problem solving are no longer distinct in their methods, as we will discuss later. As a consequence, GI systems are used widely in all kinds of organizations, from academic institutions to government agencies, not-for-profit organizations, and corporations. The use of similar tools and methods across so much of science and problem solving is part of a shift from the pursuit of curiosity within traditional academic disciplines to solution-centered, interdisciplinary teamwork.

Nevertheless, in this book we distinguish between uses of GI systems that focus on applications such as inventory or resource management, or so-called normative uses, and uses that advance science, or so-called positive uses (a rather confusing meaning of that term, unfortunately, but the one commonly used by philosophers of science—its use implies that science confirms theories by finding positive evidence in support of them and rejects theories when negative

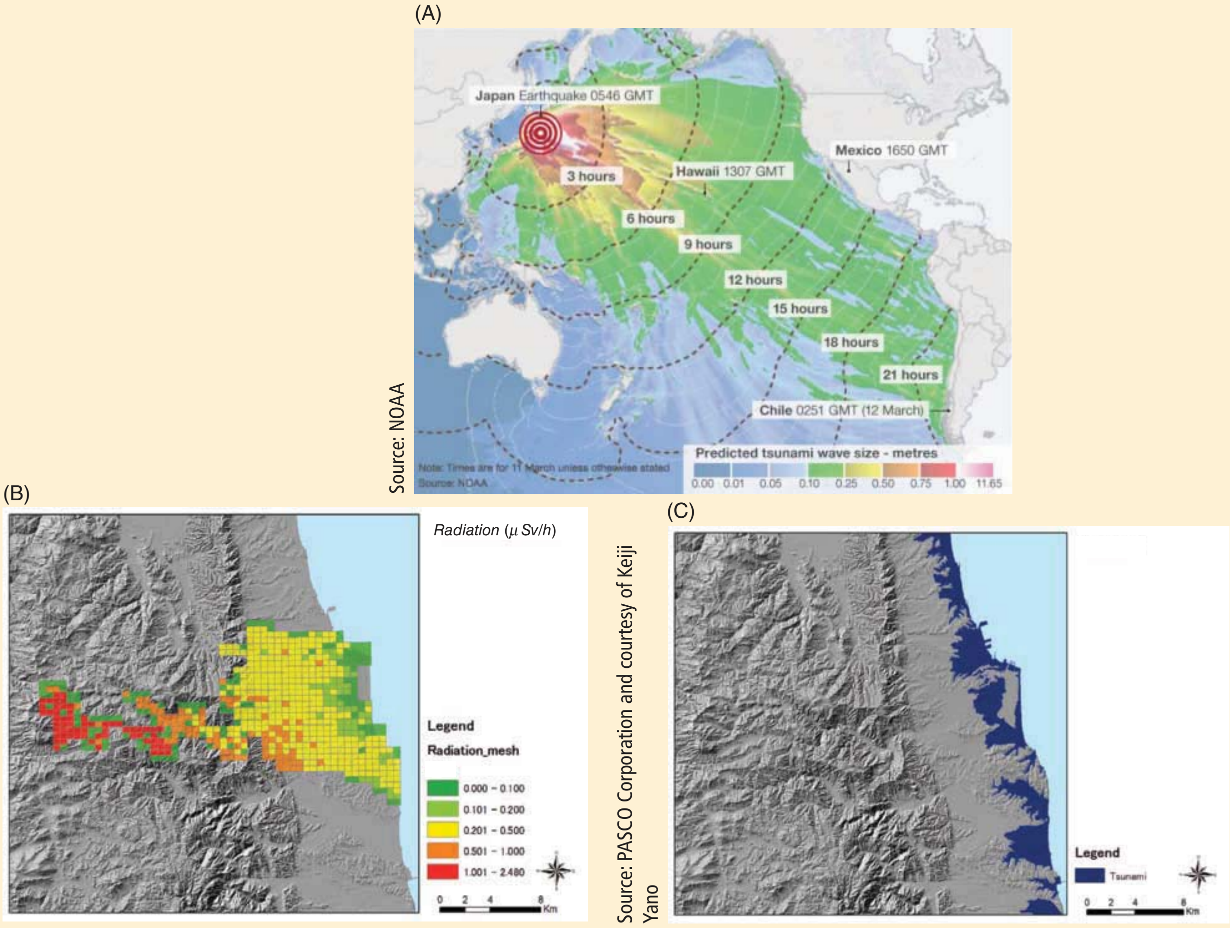
Applications Box 1.1

The 2011 Tōhoku Earthquake and Tsunami

At 14.46 local time (05.56 GMT) on March 11, 2011, an undersea earthquake measuring 9.0 on the Richter scale occurred approximately 43 miles (70 kilometers) east of the Japanese coast of Tōhoku. This was the most powerful earthquake ever to have been scientifically documented in Japan, and the fifth most powerful earthquake in the world since modern record-keeping began in c. 1900. The earthquake moved Honshu (the main island of Japan) 2.4 m (8 ft) east and shifted the Earth on its axis by estimates of between 10 cm (4 in) and 25 cm (10 in). Of more immediate significance, the earthquake caused severe earth tremors on the main islands of Japan and triggered powerful tsunami waves that reached heights of up to 40.5 meters (133 ft) in Tōhoku Prefecture and traveled up to 10 km (6 mi) inland in Sendai.

Directly or indirectly, the earthquake led to at least 15,883 deaths and the partial or total collapse of over 380,000 buildings. It also caused extensive and severe structural damage in northeastern Japan (Figure 1.3B), including heavy damage to roads and railways, as well as fires in many areas and a dam collapse. In its immediate aftermath, 4.4 million households in northeastern Japan were left without electricity and 1.5 million without water. In the following days, the tsunami set in action events that led to cooling system failures, explosions, and major meltdowns at three reactors of the Fukushima Daiichi Nuclear Power Plant and the associated evacuation of hundreds of thousands of residents. The World Bank estimated the economic cost at US$235 billion, making it the costliest natural disaster in world history.

All of this happened to a very advanced economy in an earthquake-prone region, which was almost certainly the best prepared in the world for a natural disaster of this kind. GI systems had been used to assemble information on a full range of spatially distributed



Source: NOAA

Source: Soma City and courtesy of Keiji Yano

Source: PASCO Corporation and courtesy of Keiji Yano

Figure 1.3 (A) The passage of the tsunami arising out of the Great East Japan (Tōhoku) earthquake of March 11, 2011. It had subsequent effects on Soma City in terms of (B) radiation (measured in µSv/h (micro Sievert per hour) and (C) tsunami inundation.

▶

phenomena—including the human population, the built environment, and transportation infrastructure—in preparation for a major earthquake disaster and protection against many of its foreseeable consequences.

Yet the science of predicting the location, timing, and intensity of earthquakes has made little progress over the past century. A magnitude-9.0 earthquake is a very rare event and so did not fall within any disaster-management scenario prior to the event. For example, the Fukushima reactors had been built to withstand a magnitude-8.6 earthquake on the basis of historic occurrences plus a safety margin: but not an event of magnitude 9.0. However, even when major events are unforeseen, GI science and systems are integral to response and recovery in the short term (e.g., alerting populations to the imminent arrival of a tsunami, coordinating citizen reports of how localities have been affected, and organizing evacuation), the medium term (e.g., managing the disruption to industrial supply chains), and the long term (e.g., prioritizing repair and replacement of damaged transport infrastructure). All these actions take place in an organizational context. Early warning systems are very much an international effort. In terms of addressing effects after the event, the Tōhoku earthquake raised issues that were best addressed at the national level, whereas much of the implementation was best effected at local levels.

The three Ps of disaster management are prevention, preparedness, and protection. GI science and systems are integral to each of them.

evidence is found). Finding new locations for retailers, with its focus on design, is an example of a normative application of GI systems. But to predict how consumers will respond to new locations, it is necessary for retailers to analyze and model the actual patterns of behavior they exhibit. Therefore, the models they use will be grounded in observations of messy reality that have been tested in a positive manner.

Design is concerned with improving the world—with decisions that when implemented achieve certain desired objectives, such as constructing new housing subdivisions, developing conservation plans, or defining sales territories. In recent years the term *geodesign* has become a popular way of referring to design decisions at geographic scales, supported by GI systems. All of us would like to design improvements to the world, and GI systems are valuable tools for doing so. Although most work with GI systems is considerably more mundane, it is always good to bear its grander potential in mind. As we show in Section 14.4, geodesign combines two important functions of GI systems—the ability to capture new ideas through sketching (creating/editing new features) and the ability to evaluate them and assess their impacts. A user might sketch a design for a new development, for example, and ask the GI system to predict its impacts on transportation, groundwater, and air pollution.

With a single collection of tools, GI systems are able to bridge the gap between curiosity-driven science and practical problem solving

The third way in which problems can be distinguished is on the basis of their *time scale*, ranging in human terms from the dynastic (perhaps thousands of years; see Box 1.2) to the diurnal, but very much longer with respect to understanding geological or geomorphological change. At one end of the human time spectrum, some decisions are operational and are required for the smooth day-to-day functioning of an organization, such as how to control electricity inputs into grids that experience daily surges and troughs in usage. At slightly longer timescales, tactical decisions might include where to cut trees in next year's forest harvesting plan. Still other decisions are more infrequent and strategic in nature, such as those required to give an organization long-term direction, as when a retailer decides to expand or rationalize its store network (Figure 1.4). At the far end of the human time spectrum, Box 1.2 describes how the geographic

Figure 1.4 Many store location principles are generic across different retail markets, as with Tesco's investment in Ostrava, Czech Republic.

distributions of family names, past and present, can be used to indicate how settled (or otherwise) is the population of different places, and even the geography of the DNA of long-settled residents consequent on population movements in early human history (see Box 1.4).

Although humans like to classify time frames into hours, days, years, centuries, and epochs, the real world is somewhat more complex than this, and these distinctions may blur—what is theoretically and statistically a 1000-year flood in a river system influences strategic and tactical considerations, but may arrive a year after the previous one! Other problems that interest geophysicists, geologists, or evolutionary biologists may occur on timescales that are much longer than a human lifetime, but are still geographic in nature, such as predictions about the future physical environment of Japan or about the animal populations of Africa. GI databases are often *transactional* (see Section 9.9.1), meaning that they are constantly being updated as new information arrives, unlike paper maps, which stay the same once printed.

Applications are discussed to illustrate particular principles, techniques, analytic methods, and management practices (such as risk minimization) as these arise throughout the book.

1.1.2 Spatial Is Special

The adjective *geographic* refers to the Earth's surface and near surface, at scales from the architectural to the global. This defines the subject matter of this book, but other terms have similar meaning. *Spatial* refers to any space, not only the space of the Earth's surface; this term is used frequently in the book, almost always with the same meaning as *geographic*. But many of the methods used in GI systems are also applicable to other nongeographic spaces, including the surfaces of other planets, the space of the cosmos, and the space of the human body that is captured by medical images. Techniques that are integral to GI systems have even been applied to the analysis of genome sequences on DNA. So the discussion of analysis

Applications Box 1.2

Researching Family Histories and Geo-Genealogy

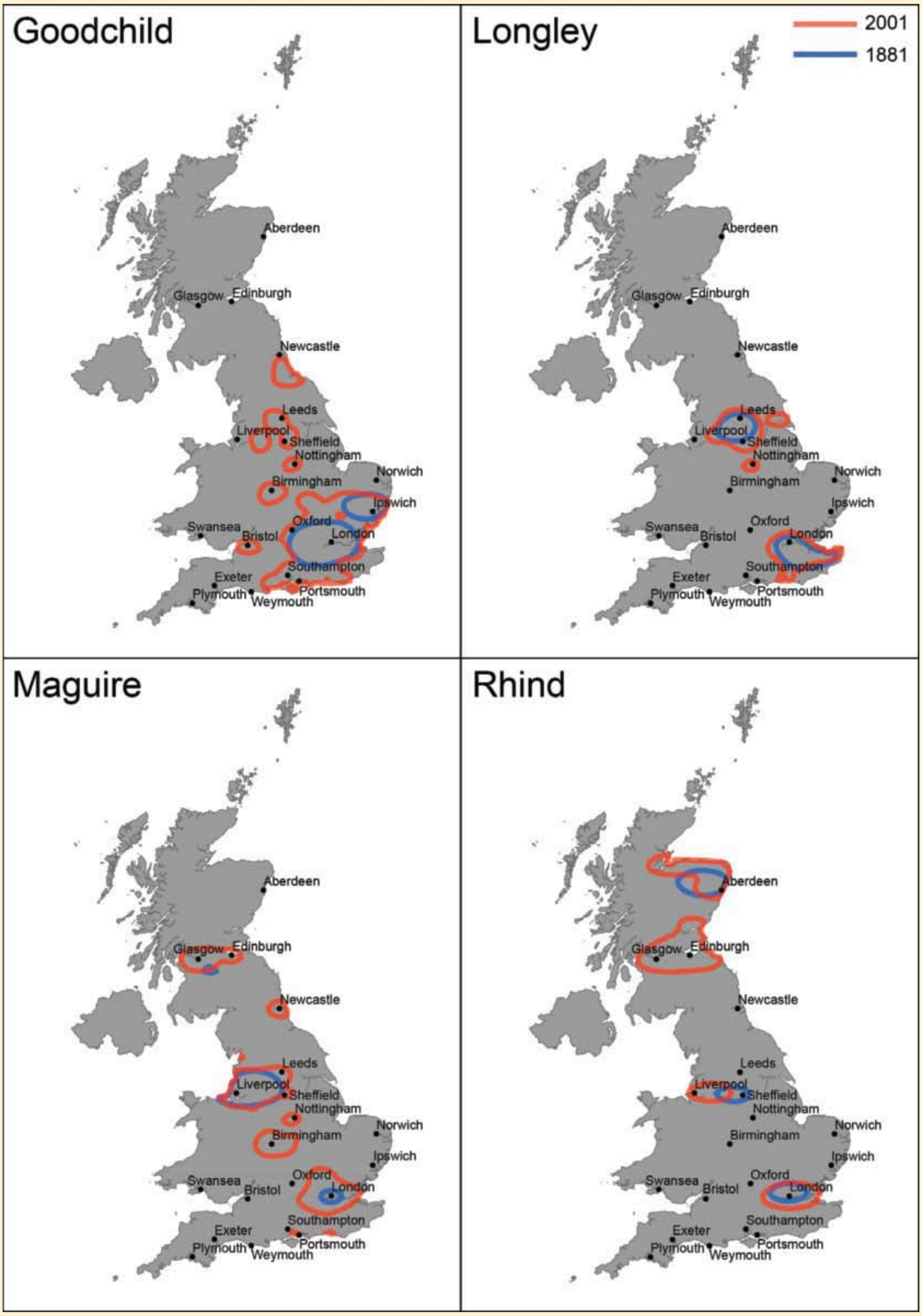
As individuals, many of us are interested in *where*, in general terms, we came from at different points in recorded human history—for example, whether we are of Irish, Spanish, or Italian descent. More specific locational information can provide clues about the work and other lifestyle characteristics of our ancestors. Some of the best clues to our ancestry may come from our surnames (family names) because many surnames indicate geographic origins to greater or lesser degrees of precision (such clues are less important in some Eastern societies, where family histories are generally much better documented). Research at University College London uses GI systems to analyze historic and present-day lists of names to investigate the changing local and regional geographies of surnames across the world. Figure 1.5 illustrates how the bearers of four selected Anglo-Saxon names in Great Britain (the ancestors of the authors of this book) have mostly stayed put in those parts of the island where the names first came into common parlance at some point between the 12th and 14th centuries—although some have evidently migrated to urban centers.

It also turns out that the mix of names with similar geographic origins in any given area can provide a good indication of regional identity. Figure 1.6, derived from the PhD thesis of Jens Kandt, presents a regionalization of Great Britain on the basis of the present-day residences of bearers of different surnames. (This is essentially a geography of rural Britain. Note that the major urban areas have been excluded because they are characterized by mixes of names arising from urban–rural, interregional, and international migration over the last 200 or so years).

All of this is most obviously evident for Great Britain and many of the countries of Europe, where populations have remained settled close to the locations at which their names were first coined. But there is also evidence to suggest that the spatial patterning of names in former colonies, such as North America, Australia, and New Zealand, is far from random. Figure 1.7 illustrates this for the surname Singleton, which can be used to build evidence about the migration patterns of bearers of this name from their documented origins in northwest England.

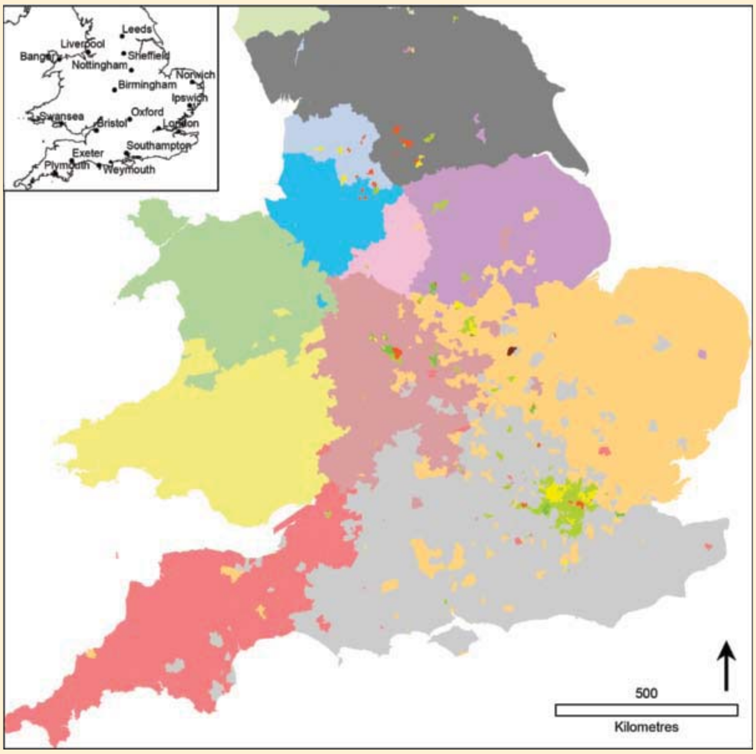
Fundamentally, this is curiosity-driven research, driven by the desire among amateur genealogists to discover their roots. But the same techniques can be used to represent the nature and depth of affiliation that people feel toward the places in which they live. Moreover, the work of Sir Walter Bodmer and colleagues (Box 1.4) is highlighting probable links between surnames and genetics, rendering this curiosity-driven research relevant to the development of drug and lifestyle interventions.

▶



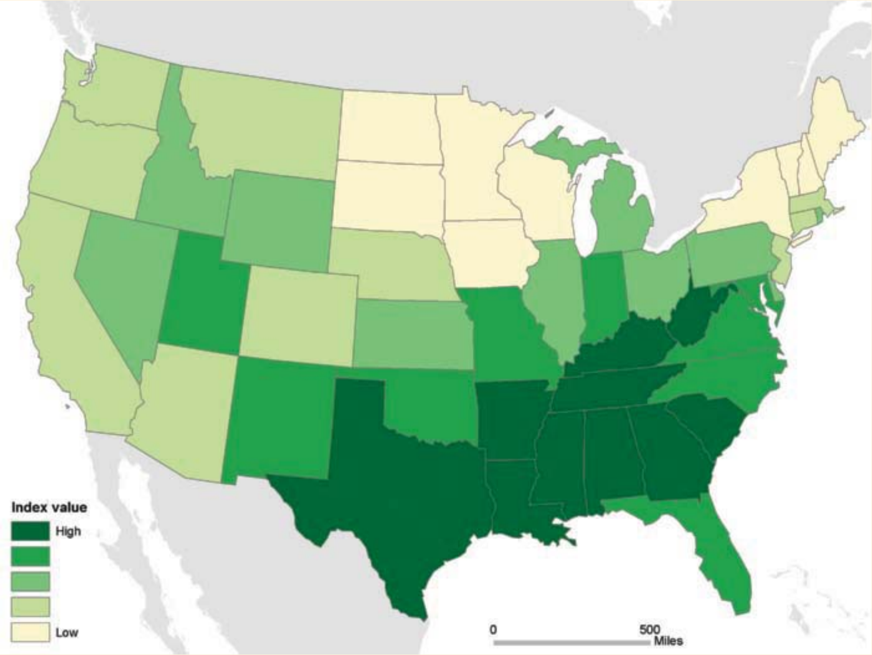
Courtesy: James Cheshire

Figure 1.5 The Great Britain Geography of the Longleys, Goodchilds, Maguires, and Rhinds. In each case the shorter (blue) line delineates the smallest possible area within which 95% of name bearers reside, based on 1881 Census of Population figures, and the outer (red) line encloses the smallest area that accommodates the same proportion of adult name bearers according to a recent address register.



Courtesy: Jens Kandt

Figure 1.6 A regionalization based on the coincidence of distinctive patterns of surnames, showing the southern part of Great Britain. Major urban areas do not fit into this regional pattern because their residents are drawn from a wide range of national and international origins.



Courtesy: Alex Singleton

Figure 1.7 The Singleton family name derives from a place in north-west England, and understandably the greatest concentration of this name today still occurs in this region. But why should the name be disproportionately concentrated in the south and west of the United States? Geographical analysis of the global pattern of family names can help us to hypothesize about the historic migrations of families, communities, and cultural groups.

Technical Box 1.3

Some Technical Reasons Why Geographic Information Is Special and Why GI Science and Systems Have Developed

- It is multidimensional, because *two* coordinates must be specified to define a location, whether they be *x* and *y* or latitude and longitude; and a third coordinate is needed when elevation is important.
- It is voluminous because a geographic database can easily reach a terabyte in size (see Table 1.2).
- It may be collected by citizens, governments, or other organizations, and it may prove useful to pool information from these diverse sources.
- It may be represented at different levels of spatial resolution, for example, by using a representation equivalent to a 1:1 million-scale map or a 1:24,000-scale one (see Section 3.7).
- It may be represented in different ways inside a computer (see Chapter 3), and how this is done can strongly influence the ease of analysis and the end results.
- It must often be projected onto a flat surface, for reasons identified in Section 4.8.
- It requires many special methods for its analysis (see Chapters 13 and 14).
- It may be transformed to present different views of the world, for example, to aid interpretation.

in this book is of *spatial* analysis (see Chapters 13 and 14), not geographic analysis, to emphasize this versatility.

Another term that has been growing in usage in recent years is *geospatial*—implying a subset of spatial applied specifically to the Earth's surface and near surface. In this book we have tended to avoid *geospatial*, preferring *geographic*, and we use *spatial* where we need to emphasize generality.

Although there are subtle distinctions between the terms *geographic(al), spatial,* and *geospatial,* for many practical purposes they can be used interchangeably.

People who encounter GI for the first time are sometimes driven to ask why geography is so important; why, they ask, is spatial special? After all, there is plenty of information around about geriatrics, for example, and in principle one could create a geriatric information system. So why has GI spawned an entire industry, if geriatric information has not done so to anything like the same extent? Why are there unlikely to be courses in universities specifically in geriatric information science and systems? Part of the answer should be clear already: almost all human activities and decisions involve a location component, and the location component is important. Another reason will become apparent in Chapter 2, where we will see that working with GI involves complex and difficult choices that are also largely unique. Other, more technical reasons will become clear in later chapters and are briefly summarized in Box 1.3.

1.2 Data, Information, Evidence, Knowledge, and Wisdom

Information systems help us to manage *what we know*, by making it easy to organize and store, access and retrieve, manipulate and synthesize, and apply to the solution of problems. We use a variety of terms to describe what we know, including the five that head this section and that are shown in Table 1.1. There are no universally agreed-on definitions of these terms. Nevertheless it is worth trying to come to grips with their various meanings because the differences between them can often be significant, and what follows draws on many sources and thus provides the basis for the use of these terms throughout the book. Data clearly refers to the most mundane kind of information and wisdom to the most substantive. *Data* consist of numbers, text, or symbols, which are in some sense neutral and almost context-free. Raw geographic facts, such as sensor measurements of temperature at a specific time and location, are examples of data. When data are transmitted, they are treated as a stream of bits; a crucial requirement is to preserve the integrity of the data set. The internal meaning of the data is irrelevant in such considerations. Data (the noun is the plural of datum) are assembled together in a