

---

# **Analyzing Precision Ag Data**

**A Hands-On Case Study  
in Spatial Analysis and Data Mining**

*Updated November, 2002*

**Joseph K. Berry**

*W.M. Keck Scholar in Geosciences  
University of Denver  
Denver, Colorado*

---

# Table of Contents

<b>Forward</b> by Grant Mangold .....	v
<b>Preface</b> .....	ix
<b>1. Overview of the Case Study</b> .....	1
1.1 Introduction	
1.2 The Big Picture	
1.3 Case Study	
1.4 Some Assembly Required	
1.5 Exercises	
Example of Analysis Capabilities	
Viewing Full Color Figures	
Installing MapCalc Software	
<b>2. Mapped Data Visualization and Summary</b> .....	7
2.1 Visceral Visions	
2.2 Preprocessing and Normalization	
2.3 Exercises	
Interacting with Map Displays	
Thematic Mapping	
Map Summary Statistics and Charts	
Map Normalization	
<b>3. Comparing Mapped Data</b> .....	15
3.1 Comparing Yield Maps (Discrete)	
3.2 Comparing Yield Surfaces (Continuous)	
3.3 Exercises	
Simultaneously Viewing Yield Maps	
Comparing Yield Maps	
Comparing Yield Surfaces	
Other Map Comparisons	
<b>4. Spatial Interpolation</b> .....	23
4.1 From Point Samples to Map Surfaces	
4.2 The Keystone Concept	
4.3 Benchmarking Interpolation Results	
4.4 Assessing Interpolation Results	
4.5 Exercises	
Interpolating Point data (IDW)	
Interpolating Point Data (Kriging)	
Comparing Interpolation Results	
Evaluating Interpolation Performance	
<b>5. Characterizing Data Groups</b> .....	33
5.1 Calculating Field Similarity	
5.2 Identifying Data Zones	
5.3 Mapping Data Clusters	
5.4 Exporting Mapped Data	
5.5 Exercises	
Characterizing Map Similarity	

---

Identifying Data Zones  
Mapping Data Clusters  
Exporting Data for Further Analysis

<b>6. Developing Predictive Models</b> .....	45
6.1 Predicting Yield Maps	
6.2 Assessing Prediction Model Results	
6.3 Stratifying for Better Predictions	
6.4 Exercises	
Predictive Modeling	
Calculating Error	
Deriving a Stratified Model	
<b>7. Analyzing Spatial Context</b> .....	55
7.1 Spatial Analysis Operations	
7.2 Applying Spatial Analysis Tools	
7.3 GIS Modeling	
7.4 Exercises	
Delineating Field Edges	
Deriving Erosion Potential	
Localized Yield Variability	
<b>8. Generating Prescription Maps</b> .....	63
8.1 The Big Picture	
8.2 Technical Issues	
8.3 Decision Rules	
8.4 Economic Maps	
8.5 On-Farm Testing	
8.6 Conclusion	
8.7 Exercises	
Deriving a Prescription Map	
Deriving Economic Maps	
<b>Appendix A. Quick Set of Exercises</b> .....	71
<b>Appendix B. Creating Your Own Ag Database</b> .....	79

---

## Preface

“Make everything as simple as possible, but not simpler” (*Einstein*)

Many farmers see *precision agriculture* as an oxymoron. With mud up to the axles and a hundred acres left to plow, precision seems worlds away. Yet *site-specific management* makes sense to a rapidly growing number of producers as years of experience confirm the spatial variability in field conditions and yield. Mapping and analyzing this variability and linking the relationships to management action places production agriculture at the cutting edge of *geographic information systems (GIS)* applications. Its use down on the farm is both down to earth and downright ambitious.

Until the 1990s, maps played a minor role in production agriculture. Soil maps and topographic sheets, for the most part, were far too generalized for application at the farm level. Acquisition of spatial data with the detail and information farmers needed for spatially responsive operations were beyond reach. Management actions were dominated by the principle of *whole-field management* based on broad averages of field data. Weigh-wagon, or grain elevator measurements, established a field’s average yield performance. Soil sampling determined the typical average nutrient levels within a field. From these and other data the best overall seed variety was chosen and a constant rate of fertilizer applied, as well as a bushel of other decisions—all treating the entire field as uniform within its boundaries.

*Site-specific management*, on the other hand, recognizes the variability within a field and is about doing the right thing, in the right way, at the right place and time. It involves assessing and reacting to field variability by tailoring management actions, such as fertilization levels to match changing field conditions. The site-specific approach assumes that managing field variability leads to cost savings and/or production increases, as well as improved stewardship and environmental benefits.

Site-specific farming isn’t just a bunch of pretty maps, but a set of new technologies and procedures linking mapped data to appropriate management actions. The *global positioning system (GPS)* locates equipment within a few feet anywhere in a field. On-the-fly data collection devices provide continuous data logging of crop yield and variable rate control units alter the amount of farm inputs as needed. Working in concert, these *intelligent devices and implements (IDI)* effectively apply seeding rates, fertilizers and herbicides, among other inputs, precisely where they are needed for economic and environmental gains. *Geographic information systems (GIS)* technology is used to store, display and analyze these data. It provides the link between crop productivity and field conditions used in constructing “prescription maps” for effective planning and management of farm activities.

Of the three underlying technologies (GPS/IDI/GIS), the extension of GIS from simply pretty maps to map analysis is the least understood yet holds the greatest promise to revolutionize farming practices. *Analyzing Precision Ag Data* is designed to introduce the underlying concepts used in spatial analysis and data mining that are needed to “cross the chasm” from simply mapping to map analysis.

### Book Organization

*Analyzing Precision Ag Data* is organized into eight topics and two appendices that lead the reader from an understanding of the fundamental nature of mapped data through a series of basic procedures used in deriving, analyzing and applying spatial information. A case study approach is used with each topic area describing the application of a set of related analysis techniques. The discussion is followed by a series of hands-on exercises providing practical

---

experience in applying the techniques. The exercises include step-by-step instructions that are thoroughly annotated.

### **Companion Software**

*MapCalc* software by Red Hen Systems, Inc. is used for the hands-on exercises. A free evaluation version of MapCalc is included with this book. The software includes the basic set of data and operations needed to complete the hands-on exercises; **the evaluation version expires after a two-week period.**

The full *MapCalc Learner* software for individual use contains additional materials and provides for extended experience in map analysis procedures and applications. It contains a basic set of functions for import/export of your own data that is constrained to a field-level (100row x 100col) analysis grid configuration. *MapCalc Academic* is a full, multi-seat licensed educational version for classroom/lab use and contains an additional instructor CD with supporting teaching materials including PowerPoint slides, exercises and quizzes for a several workshop and college-level course offerings. *MapCalc Professional* is licensed for commercial use and includes a full range of features including image data routines, coordinate/datum transformation and an extended set of import/export formats for data exchange.

For more information on the Learner, Academic and Professional versions of MapCalc, visit [www.redhensystems.com/mapcalc/](http://www.redhensystems.com/mapcalc/). Pricing for the MapCalc versions is US\$21.95 for Learner, \$495 for Academic and \$695.00 for Professional plus shipping and handling. (September, 2002; prices subject to change)

### **Case Study Data Set**

The data used in this book was provided through joint cooperation among USDA-ARS, Colorado State University, Red Hen Systems and the producer. The geographic coordinates for field have been altered to provide anonymity. The field is a 189 acre center pivot irrigated cornfield on the Colorado high plains. A 50x50 foot analysis grid resolution is used to represent field conditions over four years. Base map layers include:

#### Tabular Data:

- Field boundary file (.tab file)
- Soil sample data (1997; .tab and .xls files)

#### Map Data:

- Feature data including field edge and access road
- High resolution digital elevation surface
- Order III soil survey
- Soil properties including percent sand, clay, silt, organic matter (1996)
- Soil nutrients/chemistry including phosphorous, potassium, nitrogen and pH (1996 and 1997)
- Veris conductivity deep and shallow (1997)
- Yield volume, mass and moisture (1997, 1998 and 2000)
- Image data including green, red and derived NDVI (2000)

### **Intended Audience**

This book is ideal for farmers, crop consultants, farm input advisors, technology developers and scientists who are interested in a basic understanding of the concepts, procedures and considerations in analyzing precision agriculture data. The material is presented in an informal manner designed so the reader can grasp the broad issues and then delve into hands-on exercises for practical experience in applying the techniques.

---

## Internet Extended Environment

There are several benefits for readers with a self-published book. Traditionally published texts have inordinately slow publishing cycles compared to the speed of software development and application innovation. This book is printed in small batches and frequently revised to keep the material current. The Internet is a great mechanism for feedback so the book can evolve *with your help*. Please help in identifying sections that need correction or clarification, as well as providing suggestions for extending the current topics or adding entirely new ones—email the author at [jberry@innovativegis.com](mailto:jberry@innovativegis.com).

In addition, the author's website at <http://www.innovativegis.com/basis/> serves as a mechanism for updates, enhancements, extensions and related articles. It is recommended that you periodically check this website for up-to-the-minute information pertaining to the book. Notices and updates for the companion MapCalc software are posted at [www.redhensystems.com/mapcalc/](http://www.redhensystems.com/mapcalc/).

## Acknowledgements

Preparation of a book as complete as this one requires input from several organizations and individuals. *Colorado State University* and the *USDA-ARS* unit in Fort Collins, Colorado graciously shared their field data—numerous students and researchers were involved in the design and implementation of a comprehensive field data collection program far beyond the basic data set used in this book. *Meredith Publishing* supported the “Inside GIS Toolbox” column in the @gInnovator newsletter from 1993 through 2000 that provides the book's framework—Grant Mangold served as editor but provided far more guidance and inspiration than normally is asked of an editor. *Red Hen Systems* provided a critical ingredient—powerful yet inexpensive educational software that is very easy to use. In addition, Neil Havermale, David Wright, Carol Snyder and Elaine McCallum of Red Hen Systems contributed valuable insight over several years into agricultural practices and practical applications—a group of exceptionally bright individuals setting the early course of precision agriculture.

Many Thanks!

*Joseph K. Berry  
University of Denver  
Denver, Colorado  
September 2002*