

Advances in Geographic Information Science

Robert Gilmore Pontius Jr

Metrics That Make a Difference

How to Analyze Change and Error



Springer

Advances in Geographic Information Science

Series editors

Shivanand Balram, Burnaby, BC, Canada

Suzana Dragicevic, Burnaby, BC, Canada

The series aims to: present current and emerging innovations in GIScience, describe new and robust GIScience methods for use in transdisciplinary problem solving and decision making contexts, illustrate GIScience case studies for use in teaching and training situations, analyze GIScience tools at the forefront of scientific research, and examine the future of GIScience in an expanding knowledge-based economy. The scope of the series is broad and will encompass work from all subject disciplines that develop and use an explicitly spatial perspective for analysis and problem-solving.

Advances in Geographic Information Science.

More information about this series at <https://link.springer.com/bookseries/7712>

Robert Gilmore Pontius Jr

Metrics That Make a Difference

How to Analyze Change and Error

 Springer

Robert Gilmore Pontius Jr
Graduate School of Geography
Clark University
Worcester, MA, USA

ISSN 1867-2434 ISSN 1867-2442 (electronic)
Advances in Geographic Information Science
ISBN 978-3-030-70764-4 ISBN 978-3-030-70765-1 (eBook)
<https://doi.org/10.1007/978-3-030-70765-1>

© Springer Nature Switzerland AG 2022

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

*To my son Nicholas and daughter Olivia who
generate my guiding light*



Photo by Nicholas K Pontius

Foreword

This book has been at least 20 years in the making. Land change science was a reasonably new field in 2000, with comparison of remote sensing observations of land use and cover over time as one of its core tools. While geographers, remote sensing scientists, and statisticians had developed tools for comparisons of maps and other forms of categorical variables by that time, there was a reasonable paucity of tools for fully understanding and quantifying these changes in ways that help unravel the key processes of interest to land change scientists, like urbanization, land degradation, rotational agriculture and forestry, and land-use intensification. The kappa statistic, originally developed for comparison of categorical variables in educational research, had been adapted as the key tool for making quantifying both error in remote sensing classifications and change between two land-cover maps, but it was a reasonably blunt tool for the latter purpose especially. Seeing the need for better tools to understand how land covers are changing and interpreting those changes in terms of processes of interest to domain scientists, Professor Pontius set out to construct a rich mathematical scaffolding that exploits the change matrix, a pairwise accounting of the amounts of land covers in each category at two different times, and provides tools for researchers to better conduct their work.

The results of his work, described in many papers and now summarized smartly in this volume, have informed and enabled hundreds of land change studies, and his generosity in making his algorithms available to others through an easy-to-use Excel spreadsheet and numerous training sessions, have contributed substantially to that impact. His efforts to deconstruct observed changes into mathematical descriptions of component parts of quantity, exchange, and shift, and to build up new summary statistics like total operating characteristic and intensity analysis have brought both fresh thinking to land change science and provided a general framework for map comparisons across a broader range of applications in geographic information science. While I understand that it was never his intention, these efforts at rethinking and recalculating categorical map comparisons led Professor Pontius to declare “Death to Kappa” in his most highly cited publication from 2011. The title along with that iconic paper signaled the birth of a whole new generation of quantitative

map comparison statistics that will continue to serve, thanks also to the introduction of this volume, in advancing land change and other sciences now and on into the future.

Daniel G. Brown
Professor and Director, School of Environmental and Forest Sciences,
Washington University,
Seattle, WA, USA

Preface

Let us begin with a brainteaser so you can get an idea of the concepts you will learn in this book. Take a moment to ponder the riddle before you read the next paragraph. Your government warns that 10% of your neighbors have a deadly contagious virus. The producer of a diagnostic test advertises that 90% of their tests are correct for any population. The test indicates that you have the virus. This book's author claims your test has a 50% chance of being false, given your test is positive. Who do you believe? This book gives you insights necessary to interpret metrics that make a difference in life's decisions.

The solution is that the government, the producer, and the professor are giving consistent information, yet some metrics are more helpful than other metrics. The producer's advertisement that 90% of all tests are correct is partially helpful but potentially misleading for your purpose. You need a metric that makes a difference to your decisions. You need to know the probability of having the virus, given that the test diagnosed the Presence of the virus. If you had read this book, then you would have likely visualized the rectangular Venn diagram below, where the bounding square represents all tests (Fig. 1). The Venn diagram has two sets drawn to scale for our example. The dotted boundary along the bottom outlines the set of true Presence of the virus. The dashed boundary indicates the set of diagnosed Presence of the virus. The label Hit denotes the sets' intersection, which contains the True Positives. Half of the dashed set is in the dotted set. Your test diagnosed the Presence of the virus so you are in the dashed set; therefore, you have a 50% chance of truly having the virus.

This book offers metrics that make an important difference for interpretation, and warns you of metrics that do not. This book's intended audience ranges from undergraduate university students to senior scientists. Most of the mathematics in this book are addition, subtraction, multiplication, and division. Some of the later chapters use high-school-level concepts such as statistical regression. A major concept is a Venn diagram, which you probably have seen since your middle school math class. If you can understand a Venn diagram, then you already have a grasp of a major concept in most of this book. I write intentionally to communicate clearly with readers who might have math anxiety, while I suspect that many of the readers

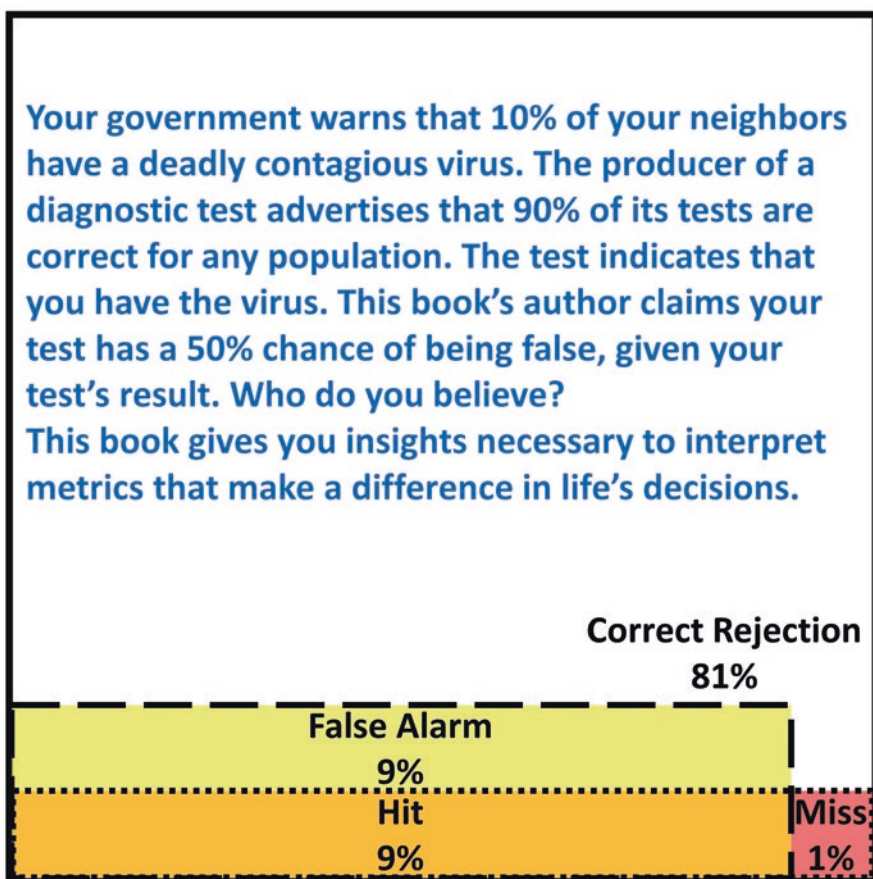


Fig. 1 Sizes of Hits, Misses, False Alarms and Correct Rejections drawn to scale for brainteaser

enjoy math as I do. This book uses math to express concepts that are fundamental to science. If you find science valuable, then this book is for you.

I have generated the ideas in this book by thinking about its concepts for more than two decades. I have developed techniques to communicate the ideas while repeatedly teaching courses that I developed at Clark University in the United States. The ideas concern how to compare variable **X** with variable **Y**, where the observations form pairs (**X**,**Y**), where both **X** and **Y** show the same phenomenon. For example, one possible application is to assess diagnoses, where one variable describes the diagnosis and the other variable describes the truth. A second possible application is to compare two diagnoses, where **X** describes one diagnosis and **Y** describes another diagnosis, while the truth remains unknown. A third possible application is to characterize temporal change of a phenomenon, where **X** describes the start time and **Y** describes the end time. I give cases for several types of variables: binary, rank, categorical, interval, and vector. I am an applied statistician, thus I describe methods so readers can apply them to a variety of scientific subjects:

Biology, Computer Science, Chemistry, Engineering, Environmental Science, Management, Physics, Political Science, Psychology, Sociology, et cetera. My specialty is Geographic Information Science, thus the examples in this book relate to Geography. Scientists who analyze diagnostic errors, temporal changes, or other types of differences will benefit from this book.

I write this book because I offer something constructive to fix many of the problems I see repeatedly in my profession. I have seen the same types of problems in the hundreds of articles that I have reviewed for scientific journals. I see some of the same flaws in published literature and at scientific conferences. Frequently, I see a presentation of an elaborate method to generate a diagnosis or prediction, but then the assessment of the diagnosis or prediction applies methods that are popular, flawed and misleading. The methods frequently either make conceptual blunders or are unnecessarily complicated in ways that render the results uninterpretable. I cannot blame the authors, because authors typically follow methods that universities teach or that have become conventional in the profession due to unfortunate and dysfunctional aspects in the culture of scientists. I write this book to offer help. This book's methods are more straightforward, interpretable and helpful than many of the complicated and misleading methods that I see in the literature. Many metrics exist for the cases that this book considers. I have found that several popular measurements are unnecessarily complicated, frequently misinterpreted, and dangerously distracting. This book recommends the metrics to use and warns of metrics to avoid. I include methods that I have found to be relevant for many types of applications during my decades as a university professor, statistical consultant, and applied scientist.

I write to inspire hope. I hope this book guides others concerning how to present metrics to answer questions in ways that are clear and important for practical applications. Science is a discipline that requires focus, organization, and clarity; science is also an art that requires its practitioners to decide what details to ignore or to demote to lesser importance. This book's methods focus on the most fundamental issues, which one must understand before trying to interpret more subtle details.

A reviewer once described my work this way: *These methods are straightforward, thus any clear-thinking scientists should use them.* I was pleased with that comment because that is my goal. However, the reviewer apparently intended that comment as a negative criticism, which reflects a scientific culture that places value on complicated mathematics. I have found that if I focus on fundamental concepts, then the mathematics are simpler and thus easier to understand. I hope is that you use this book to clarify thoughts, to communicate results, to improve science, and to widen your audience.

Experienced scientists will find in this book several novel ideas that build from familiar fundamental concepts. First, this book's overall approach might be new for some readers because the book focuses on difference, whereas other popular literature focuses on agreement. I have found that metrics of difference are more effective than metrics of agreement at directing attention to the more important information. Differences can indicate errors, which are opportunities for improvement. Differences can indicate change, which is frequently the focus of temporal analysis.

Second, a fundamental concept is the contingency table for a categorical variable, while this book's relatively new concepts include three components of difference: Quantity, Exchange, and Shift. Third, some readers might be familiar with the Relative Operating Characteristic, while this book describes a more informative approach called the Total Operating Characteristic. Fourth, the chapter concerning multiple resolution analysis gives a method to address an issue that many scientists have encountered but have not known how to address, specifically how to distinguish minor Allocation errors from major Allocation errors. Fifth, the chapter that focuses on sampling gives a necessary procedure to convert from sample data to estimated population sizes, which some scientists fail to do. Sixth, the chapters concerning an interval variable gives concepts concerning linear regression, which exist in many software packages. The same chapter defines the difference components of Quantity and Allocation, which are fundamental but appear insufficiently in literature and software. Seventh, the chapter concerning Indices of Agreement describe metrics that are popular across fields or are common only in specialized fields. Eighth, I have rarely seen in practice methods to compare vector variables, which have both magnitude and direction. I include a chapter that offers a method to compare vector variables.

Some of the methods in this book are available in the GIS software TerrSet. Readers can learn more about TerrSet at <https://clarklabs.org/>. The PontiusMatrix42.xlsx file performs the calculations for the methods in Chaps. 1, 3, and 5. The Excel file is available for free at <https://www.clarku.edu/faculty/rpontius/>. My students have written software packages in the language R for some of the methods. Those R packages are available for free at <https://cran.r-project.org/web/packages/>. The diffeR package computes the concepts in Chaps. 1, 3, 4 and 7. The TOC package performs the analysis of Chap. 2. The intensity.analysis R package computes some of the metrics in Chap. 4. Videos concerning the techniques are at <https://www.clarku.edu/faculty/rpontius/videos.html>. I hope you experience as much enjoyment and insight in reading this book as I gained in writing it.

Worcester, MA, USA

Robert Gilmore Pontius Jr

Acknowledgements

Most importantly, I thank my students, who helped me develop the concepts over the decades by working with me in courses, through theses, and via assistantships. They have contributed extremely helpful insights. I thank especially my doctoral advisees: Hao Chen, Safaa Aldwaik, Dan Runfola, Rahul Rakshit, Nicholas Giner, Alf Santacruz, Thomas Bilintoh, and Aiyin Zhang. I thank the students who wrote the 86 Masters and 15 Bachelors theses that I have advised. Some figures in Chap. 10 derive from the Honors thesis of Rebecca Chapman. The progress of science is an inter-generational process. Dr. Charles Hall was my doctoral advisor who inspired several approaches to science that contributed to how I matured as a scientist. One of his best pieces of advice is first to do the obvious straightforward analysis, then see what trouble it gets you into, then decide what to do next. A member of my doctoral committee was Dr. Steve Stehman, who contributed to my development in many ways, including by his lesson that statistical significance means something distinct from practical importance. Interactions with my professional colleagues inspired many of the ideas for this book. My colleagues at Clark University made an enjoyable and fruitful setting for intellectual creativity. The programmers at Clark Labs have incorporated some of the ideas in this book into the GIS software TerrSet. TerrSet's help section is the source of the data concerning sea surface temperature in Chap. 10. The Massachusetts Bureau of Geographic Information is the source of the data concerning land cover in Chaps. 6 and 7. Sam Khallaghi, Benoit Parmentier, Alf Santacruz, Kangping Si, and Amin Tayyebi wrote the `differ` and `TOC` packages in the language R. Priscilla Ahn and Priyanka Verma created the `pontiPy` software with the help of Claire Wang, Xiangyu Wen, and Zexing Zheng. The Edna Bailey Sussman Fund supplied funding to my students. The United States National Science Foundation funded two decades of my research through the Plum Island Ecosystems site of the Long Term Ecological Research network via a sequence of grants: DEB-972682, OCE-0423565, OCE-1026859, OCE-1058747, OCE-1238212, OCE-1637630.