## EcoHealth

Book Review

You've Got the Silver!

THE SIGNAL AND THE NOISE: WHY SO MANY PREDICTIONS FAIL—BUT SOME DON'T, NATE SILVER, 2012, NEW YORK: THE PENGUIN PRESS, 544 PP, ISBN: 978-1-59-420411-1

I love data! Locating old and new sources of data is my personal constant quest for the Holy Grail. Plotting data out, reinterpreting it, and watching it emerge from an experiment, or long-term study, is to me as edifying as deconstructing a Shakespearean sonnet, or composing an Adagio for the piano. Several of my mates have similar addictions. One of them fondly recalls a conference when he was sitting between myself and another data-addict during a post-lunch Plenary speech; repeated slides of squabbling and copulating sea birds sent us off to a happier place, until about 30 min into the talk a data-slide appeared. Instantly, we both woke up, copied down the data, and promptly fell back to sleep. As soon as the talk ended the first question came from my fellow data-addict. He had somnolently re-analyzed the one slide that had kept his mind in the room and asked a question that changed the audience's perception of the rest of the talk. I suspect that Silver is a lot like my colleague. His wonderful and already much vaunted book *The Signal and the Noise: Why So Many Predictions Fail—But Some Don't* (Silver 2012) provides a whole spectrum of wonderful insights on how to interpret, collect, and analyze data, and an equally perceptive set of examples of the ways in which data has been, and is, misinterpreted, misappropriated, and generally abused, either by accident or political design. It is a book that everybody working in science, politics, policy, and sports broadcasting should read as it has insights and criticisms from all these fields.
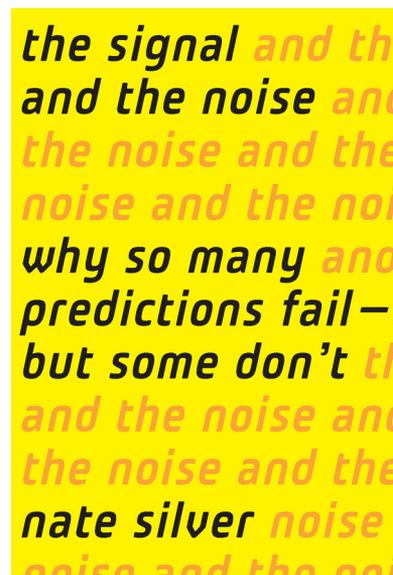
The book is carefully and constructively organized; first introducing you to the central core of concepts you should keep in mind when thinking about interpreting any observation or series of observations. Silver kicks off the book by explaining basic probability theory using the savings and loan crisis as an example of the misuse of probability theory. This might sound intimidating, but as written we instantly learn that the fundamentals of probability theory can quickly be grasped, while simultaneously realizing that most people working in savings and loans in recent decades have either failed to grasp some very basic mathematics, or willfully ignored it. Silver then goes on to use examples from baseball statistics and political pollsters to illustrate other key aspects of data analysis, while concomitantly illustrating how adept politicians are at misusing and misinterpreting happenstance as determinism. At the end of the first section of the book, most readers will realize that the old adage of "Lies, dam lies, and statistics"

needs to be revised to "Lies, dam lies and political polling". Amusingly, the most vehement critics of the book in the Amazon online reviews appear to be the staffers and research assistants of the political pundits so assiduously unveiled in the chapters on political pollsters.

Like many subscribers to EcoHealth, I was most interested in the sections of the book that deal with epidemiological data. Silver does a super job of describing the logic and pitfalls of epidemiological models and their use for understanding disease dynamics and misunderstanding disease forecasts. Moreover, he quite seems to like epidemiologists "- in a refreshing contrast to their counterparts in some other fields, epidemiologists were strongly aware of the limitations of their models" (Chap. 7, p. 229). Phew! The insights he provides from different influenza outbreaks would provide the ground work for a graduate discussion of the uses and abuses of epidemic data.

The book provides one of the best introductions to the underlying logic of Bayesian statistics that I have ever read, Silver does this by initially using poker as an example, he confesses to dabbling in professional poker at an earlier stage of his career. He then expands this into a discussion of chess grand masters and chess computers and the way that additional knowledge is used to determine the odds of making a series of winning, or losing moves. There is much that is insightful and eloquent in the way that these (and all) chapters are organized. From here forward, I am going to recommend this book and Mick Crawley's The R Book (Crawley 2007), to all incoming graduate students and undergraduate juniors in our ecology and epidemiology programs; armed with these two volumes they will have a central underlying conceptual and mechanistic understanding of 99% of the analyses they will ever have to undertake as scientists.

Is there anything missing from the book? Each morning when I cycle into work I pass the final resting place of John Tukey, the statistician who invented "Exploratory Data Analysis" and someone to whom Silver pays both an acknowledged and unacknowledged debt in this book. On some days, I swear I can hear a rumbling or spinning noise coming from just beneath the ground, and it seems to coincide with the publication in "Nature" or "Science" of yet another paper mapping the global distribution of malaria or potential emerging disease hotspots: yet another spectacular application of "minimum likelihood". One looks forward to seeing Silver's opinion of this

type of low data, high graphic imagery in a second addition of "The Signal and the Noise."

Ultimately, this is a book that I would recommend to everyone I know from undergraduates taking their first glance at scientific papers, graduate students interpreting their first field season, right through to the editorial boards of distinguished scientific journals. If more people read this book we would all have less scientific dross to sift through while trying to keep up with analyses that really allow science to make progress. So buy multiple copies and distribute them to your friends, students, and colleagues.

I have a final caveat or warning to make. My wife listens to audio books, but she expressed extreme angst when, at my recommendation she downloaded and began listening to a copy of the book from the local library. At best, the audio version of "The Signal and the Noise" is deeply horrible. The pompous and pedestrian style of the narrator crucifies Silver's pithy prose in a manner that reminds me of Mel Gibson death's by hanging, drawing and quartering in "Braveheart"; indeed it may be the nearest that many audio-listeners get to such a painful experience. In general, I have this difficulty with audio books on academic subjects; it makes me very cautious about believing that the drive for "on-line courses" in the US Academic System is the way of the future. Alternatively, this might be wonderfully subtle justice from the author and publisher if they assume that those so assiduously criticized in the book will be prominent among those who listen to it; for the rest of us, the book is best enjoyed as the written word.

Andrew P. Dobson

Department of Ecology and Evolutionary Biology, Princeton University, 117 Eno Hall, Princeton, NJ 08544, USA
e-mail: dobber@princeton.edu

REFERENCES

Crawley MJ (2007) The R Book, Chichester: John Wiley & Sons Ltd, 942 pp
Silver N (2012) The Signal and the Noise: Why So Many Predictions Fail—But Some Don't, New York: The Penguin Press