

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

The Prairie Naturalist

Great Plains Natural Science Society

12-2015

Remembering the Days of Old: Data Exploration and Simple Statistics

Christopher N. Jacques

Follow this and additional works at: <https://digitalcommons.unl.edu/tpn>



Part of the [Biodiversity Commons](#), [Botany Commons](#), [Ecology and Evolutionary Biology Commons](#), [Natural Resources and Conservation Commons](#), [Systems Biology Commons](#), and the [Weed Science Commons](#)

This Article is brought to you for free and open access by the Great Plains Natural Science Society at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in The Prairie Naturalist by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Remembering the Days of Old: Data Exploration and Simple Statistics

Greetings GPNSS members! By the time you read this editorial, most of you will probably be enjoying spectacular seasonal changes in the outdoors this autumn and a hint of winter in the air (for those of us in the northern Great Plains region). I write this editorial during my favorite time of year, deer hunting in the Midwest. Temperatures are finally starting to resemble the fall and the Christmas Holiday season is just around the corner. Much like my children have images of Santa dancing in their heads on Christmas Eve, each fall I replay images of a trophy whitetail buck as I sit patiently in my tree stand. Though I don't have the patience I did as a younger hunter, I often find myself passing the countless hours reflecting on the real reason for waking up at "o-dark thirty." For me, sneaking to my deer stand with headlamp beaming and through the endless sea of beggar's lice (or stickseed) are experiences I have grown to appreciate, which are a pleasant diversion (all right, maybe I'm stretching it to think picking hundreds of beggar's lice seeds off your sleeves and pant legs is pleasant!) from professional obligations. In short, the enjoyment derived from the outdoor experience are memories I will carry with me for a lifetime. Of course, my personal reflections are in no way related to the topic of this editorial, but more to encourage you to step away from the desk and spend time enjoying your outdoor interests and creating memories during this Holiday Season.

After days of thinking back and forth about a topic of interest for the *TPN's* broad readership, I recently received an annual report from one of my students who was asked to summarize their research activities by reviewing the relevant literature, finalize aspects of study design (e.g., Methods), present preliminary findings, and synthesize the most salient results to date. Once I started reviewing page after page of text, the topic for this editorial hit me like a freight train. It is a topic I have seen sporadically in the past, though with much greater frequency in recent years, not only in graduate research but with papers submitted for publication in scientific journals. I will introduce the topic by asking a simple question. Have we abandoned preliminary data exploration and become too narrowly focused on modern statistical techniques in search of answers to our research questions? I will return to my student's annual report to demonstrate a problem I continue to see and believe is becoming more common in research. As is often the case, we initiate any research project by conducting a thorough review of the literature in search of information needs, followed by creating a rigorous study design to address our study objectives. You then hire a student to spend several years of his or her life working tirelessly to collect field data. At the completion of field work, the student has amassed an impressive data set over multiple field sites with adequate spatial and temporal replication (Chamberlain 2008). The student is now well-positioned to conduct a rigor-

ous analysis to address the objectives of the study. Now the problem arises.

Overflowing with excitement to "crunch some numbers," a favorite statistical package is opened, complex code is written to evaluate a priori candidate models, parameter estimates and 95% confidence intervals are calculated and presented in table after table, and elaborate syntheses of findings to make meaningful inferences and help readers understand their meaning completes the process. Sound familiar? A notable missing part of this process should be obvious. Where in the process did initial data exploration occur? Where are tabular presentations of these efforts? For instance, what about initial data descriptions and tables presenting means and measures of variability (e.g., SD, SE, CV)? What about trying to display data graphically to help evaluate the underlying assumptions of the proposed analysis (Chamberlain 2008). Now don't get me wrong, I rarely (if ever) advocate for the sharing of raw data in a manuscript, though it seems like many of us have lost sight of how to evaluate datasets. I certainly hope I'm not alone in my observations. I can vividly recall a fellow Ph.D. student asking me whether I had graphed my data or sat down to explore the data set to fully understand it prior to model development and evaluation. With my face flushing red with embarrassment, I responded to his questions with a quiet "no." We proceeded to promptly plot the data only to discover patterns that clearly described the less than optimal predictive capability of my regression models. Those two questions all those years ago seem increasingly important to me now, as my concern that researchers are moving further away from the fundamental need for data exploration when it comes to analyzing and interpreting data (Chamberlain 2008).

My student was unable to avoid the temptation to dive into the statistical approaches, of which we had grandiose expectations of being the answer to our questions, rather than first exploring our dataset at hand. The resulting product was an annual report with non-informative results that carried little biological meaning (Chamberlain 2008). How can something like this happen? Simple...the underlying data had been underutilized and not given an opportunity to tell the initial story, which would have greatly facilitated subsequent inference and clarified the novelty of their findings to wildlife management and conservation. I hope the point of this discussion is simple and one that researchers (whether fresh graduate students or seasoned veterans) will give pensive thought to prior to designing their next research project: be diligent to explore your data. That is, establish an intimate relationship with the dataset by learning it in and out (Chamberlain 2008). Initiate your analyses by being mindful to accomplish the simple goal of graphically presenting your data and allowing it to tell a story (Chamberlain 2008).

In other words, avoid the temptation to conduct fancy and cutting edge statistical analyses until you have an intimate understanding of your data. Who knows, perhaps your data will tell you an effective story without the need for complex statistical analyses. If so, why make your job any harder than it needs to be? I offer these final questions solely as potential food for thought.

This is another informative issue that contains a broad range of topics to Great Plains researchers and natural resource managers. Management and Conservation topics include ungulate habitat selection, bird habitat use, walleye food habitat, and bat distribution and diversity. For the botanists among us, you will find survey methods for estimating herbaceous biomass of grassland vegetation and soil conditioning by cheatgrass and western wheatgrass. This issue also includes book reviews ranging from a description of North American bumble bees to the natural history of Minnesota's Superior Coast. Last, this issue contains a special tribute to Dr. Paul Kannowski, founder of the North Dakota Natural Science Society (currently the Great Plains Natural Science Society) and *The Prairie Naturalist*. Dr. Kannowski was the first Editor for TPN (of 27 years) and a lifelong member of GPNSS. He will be missed by all who had the privilege of knowing him.

With the completion of each of my editorials, I find it especially gratifying to acknowledge those that are most responsible for ensuring that the Journal remains an important outlet for disseminating the work conducted throughout the Great Plains Region. I have come to realize that during my tenure as Editor-in-Chief, the work I do is best described as trivial when compared to the contributions of so many others involved in the peer review process and timely preparation of Journal issues. While I don't pretend to know how many

referees take time out of their busy lives to review scientific papers and provide constructive and thoughtful feedback to novel and experienced authors, I imagine the number is in the hundreds. The newly appointed Associate Editors (including Drs. James Lamer and Colter Chitwood) as well as the seasoned veteran AEs continue to do the lion's share of work when it comes to overseeing peer review, critically evaluating manuscripts, and making appropriate decisions and recommendations. The continued efficiency at which the Assistant Editor (Dr. Troy Grovenburg) continues to assemble each issue in a timely manner is nothing short of amazing. I appreciate the efforts of all, who keep the Journal locomotive running. Lastly, I genuinely appreciate you, the reader. Over the past 6 years, many of you have contacted me to offer support for our ongoing efforts to transform the journal, or in other cases to offer ideas for improving *TPN*. It is personally gratifying to interact with natural resource professionals who have such a strong and genuine interest in seeing *TPN* serve their needs. In closing, if you have any questions, comments, or concerns about *TPN*, please feel free to contact me. After all, this is your journal, and I very much appreciate your thoughts about it. Until next time, have a safe winter field season everyone!

—Christopher N. Jacques
Editor-in-Chief

LITERATURE CITED

- Chamberlain, M. J. 2008. Editor's Message – Are the Days of Data Exploration Gone. *73*(1):1–2; DOI: <http://dx.doi.org/10.2193/2008-465>.