update

Stemming "ignorance creep" in paleoecology and biogeography

The continued success and relevance of any scientific field depends on critical examination of its foundational knowledge. A recent perspective in *Quaternary Science Reviews* by Steve Jackson (U. of Wyoming) highlights the different ways that knowledge is formed, recognized, and then lost in paleoecology (Jackson 2012). Although he was writing for a specific audience—paleoecologists—his call for greater efforts to acknowledge and combat "ignorance creep" applies to all researchers, including biogeographers.

Jackson (2012) uses an unlikely pairing of quotes by Donald Rumsfeld and Henry David Thoreau to highlight the scope of our knowledge about any given discipline. He presents an epistemological framework for scientific understanding: knowledge versus uncertainty on the one hand, and cognizance versus ignorance on the other. These contrasts generate four categories: known knowns and known unknowns, i.e., knowledge and uncertainty of which we are cognizant; and unknown knowns and unknown unknowns, knowledge and uncertainty of which we are ignorant.

The focus of Jackson (2012) is on the often-ignored category that arises from this classification scheme—unknown knowns. Unknown knowns include "the hidden and unquestioned assumptions that underlie a discipline, the things so seemingly obvious that they are beyond question or reflection, all the things that are routinely taken for granted..." (p. 3). Jackson is particularly concerned about "ignorance creep", the process that converts "knowns" into "unknowns" and argues for increasing attention to forward models to identify and combat ignorance creep within paleoecology. Forward models trace the mechanistic processes along the path from the target variable to the proxy used to measure it. Jackson argues that the construction of forward models serves two purposes: 1) to make processes and assumptions explicit, minimizing ignorance creep, and 2) to allow quantification of uncertainty, thereby assessing the strength of resulting inferences.

Jackson uses woodrat middens as a study system to provide a worked example of forward model construction for paleoecological inference. The forward model for this system starts with a variable of interest, regional vegetation, and outlines the path through data collection and analysis to the resulting inferences a researcher might make about regional vegetation. In between those two points lie a set of assumptions, observations, and models, dealing with the ultimate source of the data (the regional vegetation), the animal vector bringing information about the source into the middens (decisions and assumptions made by or about the woodrats), issues with diagenesis (physical, chemical, and biological processes that transform the samples over time), and finally the set of analytical and inferential issues that arise in field, lab, and computer work. This forward model (and most forward models that paleoecological studies would generate) essentially recapitulates the field of taphonomy. Taphonomy is a component of paleoecology that studies how the processes of preservation affect the information found in the fossil record (Behrensmeyer et al. 2000). However, when defined more broadly, every field grapples with "taphonomic" processes, whether in the field or lab, that affect inferences made from the raw data. Every field has, at the very least, source issues and analytical and inferential issues; they may or may not have vector or diagenesis issues.

Jackson argues that forward models are important when data, originally collected for one purpose, are used for another. Distribution modeling is a good example of this. Often, the original data (e.g., occurrences of individuals at particular localities) were collected for a very specific purpose: to determine some aspect of the ecology and evolution of a particular group at a particular location. Today, however, data collected for disparate purposes, with different collection methodologies, are aggregated to understand a new set of questions, at much broader spatial and temporal scales (Boakes et al. 2010): What was the past distribution of a species? How did the distribution change through time? What sets of variables influence distribution changes? The assumptions and traditions that were appropriate for the original inferences may no longer be appropriate for these new uses. The result is a set of unknown knowns, assumptions about data that persist in the literature, but that may no longer be valid since the target variable, and thus the forward model, has changed. As Jackson notes, taphonomic studies have waned in recent decades, but evolving use of data requires revisiting and updating the taphonomic pathways.

Ignorance creep may be particularly important for biogeographers. The inherent multidisciplinarity of biogeography requires transparency about the limitations of each constituent discipline. Each community has norms that have evolved over the decades, a set of known knowns and known unknowns, either explicit or implicit. Interdisciplinary collaborations can highlight areas where norms are not well supported and thus represent unknown knowns. At other times, outsiders are unaware of known knowns or, more importantly, the known unknowns. This may result in inappropriate data use and conclusions not supported by the data or the process by which it was generated. For example, in paleo-distribution modeling, the forward model outlined by Jackson (2012) would need to be modified to fit the particular paleo-occurrence data types and then merged with, e.g., the set of processes outlined by Guisan and Thuiller (2005) for distribution modeling.

Distribution modeling and biogeography in general have done a good job of trying to understand the components of its forward models (e.g., Guisan and Thuiller 2005, Araújo and Guisan 2006, Heikkinen et al. 2006, Elith and Leathwick 2009). However, Jackson (2012) is a timely reminder that we should continue to be cognizant of these issues, particularly in biogeography. Every field has its own forward models, sets of inferences, best practices, and rules of thumb that help understand how a source variable is translated into data. Attention to the 'taphonomic' processes underlying those forward models, and critical stages of information loss, will help combat ignorance creep and strengthen the conceptual foundations of our field.

Jessica L. Blois

Center for Climatic Research, University of Wisconsin
Madison, WI 53706, USA. blois@wisc.edu;
https://mywebspace.wisc.edu/blois/web/

References

Araújo, M. & Guisan, A. (2006) Five (or so) challenges for species distribution modelling. Journal of Biogeography, 33, 1677–88.

Behrensmeyer, A.K., Kidwell, S.M. & Gastaldo, R.A. (2000) Taphonomy and paleobiology. Paleobiology, 26, 103–147.

Boakes, E.H., McGowan, P.J.K., Fuller, R.A., Chang-qing, D., Clark, N.E., O'Connor, K. & Mace, G.M. (2010) Distorted views of biodiversity: spatial and temporal bias in species occurrence data. PLoS Biology, 8, e1000385.

Elith, J. & Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics, 40, 677–697.

Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. Ecology Letters, 8, 993–1009.

Heikkinen, R.K., Luoto, M., Araújo, M.B., Virkkala, R., Thuiller, W. & Sykes, M.T. (2006) Methods and uncertainties in bioclimatic envelope modelling under climate change. Progress in Physical Geography, 30, 751–777.

Jackson, S.T. (2012) Representation of flora and vegetation in quaternary fossil assemblages: known and unknown knowns and unknowns. Quaternary Science Reviews, 49, 1–15.

Edited by Daniel Gavin