Journal of Archaeological Science 63 (2015) 160–163

Contents lists available at ScienceDirect

Journal of Archaeological Science

journal homepage: http://www.elsevier.com/locate/jas

Focus article

“All models are wrong but some are useful”: A response to Campbell's comment on estimating *Mytilus californianus* shell size

Gerald G. Singh [a], Iain McKechnie [b, c, *], Todd J. Braje [d], Breana Campbell [d]

[a] Institute for Resources, Environment and Sustainability, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada
[b] Department of Anthropology, 1218, University of Oregon, Eugene, OR, 97403, USA
[c] Department of Archaeology, Hakai Institute, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada
[d] Department of Anthropology, San Diego State University, 5500 Campanile Dr., San Diego, CA, 92182-6040, USA

ARTICLE INFO

Article history:
Received 25 June 2015
Received in revised form
26 August 2015
Accepted 31 August 2015
Available online 3 September 2015

Keywords:
Marine ecology
Morphometric regression
Historical ecology
Zooarchaeology
Size estimation
Allometry
Shellfish

ABSTRACT

Developing accurate methods for estimating animal body size from fragmentary remains is a key focus of zooarchaeological research. Here, we respond to Greg Campbell's critique regarding methods we recently developed to predict *Mytilus californianus* shell size from archaeological contexts using linear regression. We show that Campbell's assertion that our regressions are “inaccurate” is incorrect and mischaracterizes the premise and results of our study. We appreciate that Campbell draws attention to the importance of allometry but do not agree that archaeologists must first describe ontogenetic size relationships before developing a practical method for size prediction in zooarchaeology. We further argue that pooling data from broad geographic scales incorporates diverse growing conditions into a predictive model to account for the uncertainties across archaeological time scales. We conclude by highlighting the difference between zoological and zooarchaeological research goals and emphasize that the precision required for a particular analysis can create a mismatch between analytical expectations and archaeologically applicable research questions.

1. Introduction

It is an accepted truism that “all models are wrong but some models are useful” (Box, 1978:2). In a series of papers, we developed a method that provides a useful estimate of total mussel shell size from archaeological fragments (Campbell and Braje, 2015; McKechnie et al., 2015; Singh and McKechnie, 2015). Within the range of sizes considered, our models work well as shown by the fits of the data to the regression lines. Gregory Campbell's – henceforth GC's – comment (2015) focuses on what he considers to be faulty methods for estimating *Mytilus californianus* shell size. Unfortunately, GC's criticisms reflect mischaracterizations about the premise and results of our study. We find his narrowly focused comment worrisome, not just because he performs statistical procedures inappropriately but because of a mistaken belief that archaeologists must first describe the precise morphological growth relationships before developing a method for size prediction in zooarchaeology. Here, we respond to the “assumptions” he presents in his comment and reflect more generally on differences between zoological and zooarchaeological approaches to size estimation.

1.1. GC's assumption #1: organisms' dimensions have straight-line relationships

We do not dispute the importance of allometry (the relationship of body size and shape) as a foundational method for understanding how organisms develop. While GC may be correct in his claim that the relationships between particular umbo measurements (and other measurements) and shell length are allometric, these specific relationships have not been found in the *M. californianus* literature and our studies were not seeking to investigate whether this is the case. GC is right to note that Singh and McKechnie (2015) choose to present a non-linear allometric function for meat-weight predictions. However, we did this because length-weight regressions have been clearly shown to follow this type of relationship across multiple mussel species (Rodhouse et al., 1984; Zotin and Ozernyuk, 2004). Yet with respect to shell size, isometric and even linear relationships have also been observed in

---

* Corresponding author.
E-mail addresses: geraldsingh@gmail.com (G.G. Singh), iainm@uoregon.edu (I. McKechnie), tbraje@mail.sdsu.edu (T.J. Braje), breana@scic.org (B. Campbell).

http://dx.doi.org/10.1016/j.jas.2015.08.021
0305-4403/

bivalve shell morphologies, including multiple clam species as well as Atlantic blue mussels, *Mytilus edulis* (Aragon-Noriega et al., 2007; Mills and Côté, 2003; Rufino et al., 2006). Without knowing the specific relationship, we chose to express shell size parsimoniously as a linear relationship for the practical purpose of predicting shell length. We make no claim outside the range of shell sizes we measured.

A key aspect of our approach that GC does not seem to appreciate is that regression analyses can serve two purposes: to describe and/or to predict (Shmueli, 2010). Complete descriptions of shell growth, over the lifetime of an organism, or determining what variables are most responsible for shell growth, are of fundamental interest to zoologists seeking to demonstrate specific aspects of ontogeny and evolutionary development. Using regression to demonstrate such growth variation understandably requires a specific model structure that captures mechanism and theory for this zoologically oriented question. In contrast, what our method seeks to demonstrate is much less theoretical — that linear regression is a reasonable *predictor* of shell size. We have improved on existing methods of prediction, notably White's (1989) commendable first approach for estimating shell length from fragments. We provide methods to generate continuous estimates of length (rather than categories of length) with measured error rates that also quantify uncertainty of prediction (Wolverton et al., 2014). On principle, we do not presume a necessary relationship between umbo size and total length; rather, we use a simple relationship that works given our goal. In contrast, GC's critiques are focused on theoretical points that relate to using regression for description. That is, at no point do we claim that the morphological relationships presented are models *of* shell growth; rather we show that simple linear models perform well at *predicting* shell size with a reasonable degree of error. A cursory reading of our papers makes apparent this discrepancy.

As a further defense of simple linear regression, we note that linear regression has some distinct statistical advantages over other methods. First, associated R^2 values that give a sense of model fit enable a more accurate measure of explained variance than the pseudo R^2 values on which non-linear regressions rely (Smith and McKenna, 2013). Moreover, linear regressions do not have the same degree of loss of data resolution brought about by a log transformation (Cleveland, 1984).

We additionally note that GC's criticisms concerning statistical reliability can be disputed on practical points. For instance, GC's claim that our models are not “likely to be reliable” is incorrect as well as vague. Our results empirically demonstrate (through measures of model fit as well as double blind trials with training and test sets of shells) that our models are reliable for prediction with relatively low margins of error (Campbell and Braje, 2015:172; Singh and McKechnie, 2015:177). Elsewhere, GC's emphasis on non-zero intercepts highlights the possibility that some very small shells may produce negative size estimates reflects a valid but misplaced concern given that archaeologists are unlikely to recover many shell fragments with near-zero measurements (0–0.4 mm). Even if a few shells out of a series have negative values, these can be easily identified and the method can be reevaluated and improved rather than a hypothetical presumption of possible error, making this a relatively insignificant issue.

As a corollary, GC's criticisms of our statistical procedures include claims (in his “assumption 1”) that we misunderstand R^2 as a measure of performance yet he does not acknowledge our actual measurements of performance through double blind trials and error estimates (Campbell and Braje, 2015:172; Singh and McKechnie, 2015:177). Similarly, his criticism that Singh and McKechnie (2015) do not plot residuals seems moot given it is not common practice to provide diagnostic plots in publication and especially when raw data are provided in a supplemental spreadsheet as in this case. Belaboring this point however, GC criticizes Campbell and Braje's (2015) published residuals as somehow failing assumptions because they do not match figures from an introductory statistics book. It is unreasonable to expect a given dataset will fall exactly along a normal distribution, which would discount almost all empirical findings with linear models (because there are many cases where distributions may approximate normality without perfectly fitting normality). Further, GC neglects to consider Campbell and Braje's PP plots showing very high agreement with a theoretical normal distribution with no major deviations (Campbell and Braje, 2015:171). Finally, GC's references to statistical textbooks neglects more practical aspects of using linear models, namely the robustness of linear models to perform adequately in spite of deviations from normality, especially since the distribution of error at large sample sizes often conforms to a normal curve and thus provides the basis for the normality assumption, i.e., the Central Limit Theorem (Lumley et al., 2002).

1.2. GC's assumption #2: there is a single best dimension for estimation

CG's claim that we assert a single best dimension for size estimation is a major mischaracterization of our research effort. We do not suffer from the illusion that developing morphometric relationships between umbo measurements and total length is the single best measure of growth for the genus *Mytilus*. We chose to focus on the umbo solely because these structures are relatively well preserved in archaeological settings. Given that our methods seek to address a practical problem in coastal archaeological research, it was not necessary to exhaustively explore which of the myriad of potential shell measurements best predicts total shell length (i.e., a zoological question). What was important to us was to explore predictive relationships between total length and portions of shell that would be preserved in archaeological assemblages (namely around the umbo and sometimes including the hinge[1]).

Even as a critique of our method purely on predictive terms, GC's commentary is inaccurate. By stating that “a single umbo dimension is unlikely to predict shell size well” GC appears to be misreading our work as Campbell and Braje (2015) and Singh and McKechnie (2015) empirically show that this measurement performs well with low measures of error and in double blind trials.

Regarding GC's recommendations for multiple regression, he chooses not to acknowledge Singh and McKechnie's (2015) presentation of multiple regression formulas predicting total length. More significantly, we urge caution regarding GC's recommendation for backward elimination regression methods to determine strong predictors. Statisticians have long warned against the selective use of stepwise regression (including backwards elimination and forward selection) for a number of reasons, including: 1) biased parameter estimates; 2) the order of parameter loss in the procedure affects the resulting model; 3) stepwise procedures commit multiple comparisons (it is a suite of tests, not a single test) that can inflate Type I error; and 4) stepwise procedures assume a “single best model” when in fact there may be multiple models that fit the data to similar degrees (Thompson, 1995; Whittingham et al., 2006). If researchers are especially interested in determining the best models (multivariate or not) for predicting shell length, we recommend using information theoretic approaches that weigh

---

[1] GC's concern about our use of the term “hinge” rather than “resilial ridge” distracts from the fact that in general mussels have dysodont teeth morphology which includes weak teeth near the umbones and a strong ligament (Sturm et al., 2006). We opted to refer to these generally as “hinge teeth.”