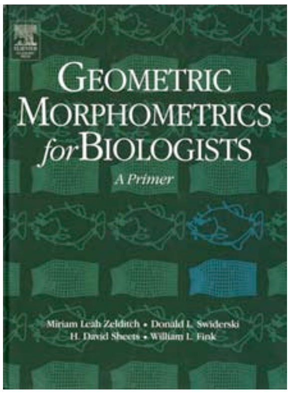

Geometric Morphometrics for Biologists: A Primer

by M. L. Zelditch, D. L. Swiderski, H. D. Sheets, and W. L. Fink (2004) Elsevier Academic Press, New York and London, 437p. ISBN 0-12-77846-08, £36.61/$69.95 (hardback)

Morphometrics is a seductive mistress. Despite its current popularity, and despite concerted attempts to explain the revolution that occurred in its practice during the 1980s and 90s, it still lulls nascent practitioners down the garden path of apparent mathematical sophistication. In return they are often left bereft and bewildered, having squandered their time carrying out simplistic analysis without engaging with the mathematics on which the discipline is grounded (limiting their ability to address novel future data-analysis challenges) and without gaining much insight into the biological processes that are supposed to be the focus of morphometric investigation. This, along with morphometrics' historical link to phenetics, leaves morphometrics often resembling an academic ghetto that graduate students and advanced undergraduates are sometimes encourage to visit—to pick up some street smarts and a bit of local colour—but cautioned not to linger in; especially after dark.

The current book, by three prominent practitioner-systematists and a physicist-software/methods developer, seeks to present an overview of geometric morphometrics in primer mode. Primers are supposed to keep detailed exposition to a minimum while conveying a sense of what can be done using particular methods. They are, in an important sense, written to be the first word in the presentation of a topic, not the last. These authors set out to write such a book for the audience best positioned to appreciate a primer: those with no more than undergraduate mathematical training who want an emphasis on applications rather than theory. Certainly there is a need for such a treatment, as was originally noted by Bookstein (1996). The canonical modern morphometric texts (e.g., Bookstein 1991; Small 1996; Dryden and Mardia 1998, Costa and Cesar 2000) are too technical and abstract to be fully understood by those not interested in making a commitment to the mathematics. At the same time, collections of generalized applications articles (e.g., Rohlf and Bookstein 1990; Marcus et al. 1993, Marcus et al. 1996; Elewa 2004) are too eclectic and lack the unified focus necessary to be used as comprehensive introductions, while the special-topics collections (e.g., Adrain et al. 2001; Zelditch 2001, MacLeod and Forey 2002) are too focused and contain too much non-morphometric material. But is this the primer we've all been waiting for?

At best, the authors' purpose has been only partially realized by the product of their labours. Many years ago Wilson and Bossert (1971) managed to write a brilliant primer of population biology in a scant 192 pages. More recently, Bryan F. J. Manley (1994) produced a highly useful primer of multivariate statistics in a mere 215 pages. Tipping the scales at 437 pages, *Geometric Morphometrics for Biologists* can hardly be called concise.

Part of the problem here may be that these authors intend their book to be used as a primary course text. This is not the traditional purpose of a primer. Nor is it a realistic course option in the biology or paleontology programmes of most universities, pressed as these are to integrate topics from a large number of specialist fields. This desire to be a textbook—instead of a genuine primer—also explains the curious weight given to mathematics (the text contains over 200 equations, many of them complex matrix-algebra expressions, but neglects to include any description of elementary matrix algebraic operations) and theory (encompassing 261 of the 407 text pages *sensu stricto*) in what the authors describe as a non-mathematical, applications-based treatment.

Topically, *Geometric Morphometrics for Biologists* is divided into three parts: (1) basics of shape data, (2) analyzing shape variables, and (3) applications of morphometric methods to complex hypotheses. A final, small section discusses the analyses of coordinates located in three dimensions (usually a trivial mathematical extension of the 2D case) and the analysis of outlines (an enormous subject in its own right that is barely touched upon in this book; see Lestrell 1997). In keeping with the exclusivist philosophy that has done so much to mar the intellectual reputation of morphometrics, the book is actually about landmark-based morphometrics, which the authors treat both as if this school encompasses all morphometrics and as if it is one school of thought among many. A similarly biased text appeared a few years ago from the outline school (Lestrell 2000). Geometric morphometrics is unified at a very basic level by the theoretical implications of Procrustes superposition. But neither this unity, nor the manner in with other approaches to morphometrics are related to geometric methods, is ever made clear here.

Part 1 of encompasses a discussion of landmarks (Chapter 1), shape variables (Chapter 2), shape theory (Chapter 3), superimposition methods (Chapter 4) and thin-plate splines (Chapter 6). The treatment of landmarks as a concept follows the more-or-less standard line that landmarks are both biological and topological homologues. Recent arguments disputing this stance (e.g., MacLeod 1999, Humphries 2002) are not considered other than to assert 'if discrete and recognizable structures are [biologically] homologous as structures, the discrete and recognizable locations on them are arguably [biologically] homologous as points' (p. 26). Such arguments from analogy confuse the ends of an analysis with the means by which these ends are accomplished. In the vast majority of cases it really doesn't matter whether topologically corresponding landmark locations are formal homologues or not, insofar as the landmarks need only be regarded as mathematical conveniences used to compare the relative positions of *structures*, which will always have a far less ambiguous claim to being actual homologues than do infinitely small point locations on structures. The simple fact that small random displacements of landmark locations have little effect on the overall analytic result is sufficient to demonstrate how 'beside the point' the assertion of 'point homologies' really are. Of course, it is also a simple matter to envision datasets composed of landmark locations collected from functionally equivalent structures (e.g., bird wing, bat wing, pterodactyl wing; porpoise body, shark body, ichthyosaur body) that could be subjected to landmark-based morphometric analysis, but which would specify demonstrably non-homologous comparisons. The authors seem to acknowledge this implicitly on the same page when they state, somewhat awkwardly 'For a deformation to make mathematical sense, the points on one form must correspond to the points on another.' Suffice it to say that topological correspondence is not necessary and sufficient evidence to infer homology and that the point of a morphometric analysis is not to understand deformations mathematically, but biologically. The difference is subtle and not widely appreciated, but important.

In this chapter we also learn that a collection of shapes too distinct to be represented by numerous landmarks is not suitable for 'morphometric analysis' (p. 27). That outline-based methods have been used profitably to summarize precisely such complex patterns of morphological variation for many years, even within the context of geometric morphometrics *sensu stricto* (e.g., Bookstein 1996a 1996b) is not mentioned. The fact that these authors also discuss non-Procrustes-based outline-analysis methods later in their book (see p. 395) as though they represent alternative, geometric morphometric approaches simply adds to the reader's confusion. This section closes with a lengthy digression into the optics of cameras and digital image processing that, while related to the general topic of morphometric data collection, has little to do with landmark theory *per se*.

The same type of organizational anarchy finds expression throughout Part 1; indeed throughout the entire book. For example, the section on the 'statistics of shape coordinates' (p. 57-58) refers to multivariate analysis of variance (MANOVA) and Hotelling's T^2 test without explaining what either these concepts are, or even cross-referencing them to (in this case) a chapter-long discussion provided later in the text (Chapter 9). At the end of the main methods chapters the smooth flow of the narrative is repeatedly broken by sections constituting an internally serialized user's guide for David Sheets' MathLab-based Integrated Morphometrics Programs (IMP) package. Few practicing landmark morphometricians use the IMP package and the decision to feature it over the simpler, more widely used, and more regularly updated public-domain software produced by F. James Rohlf (see http://life.bio.sunysb.edu/morph/) seems ill-advised; at least for the standard Procrustes-based methods.

On the positive side, the chapters on shape theory, superimposition methods, and the thin-plate spline are all quite good overall, with the shape-theory chapter being truly outstanding since much of this material has only been accessible to date in highly mathematical dissertations or (curiously) as part of the Help system in the Rohlf software packages. The thin-plate spline chapter also uses many useful analogies and examples that will greatly help students gain familiarity with the concept and confidence with the interpretation principal/partial warps. However, since this chapter is so successful in terms of the qualitative presentation, it seems more than a bit odd that the authors chose to drop a 37 expression-long algebraic derivation of partial warps into the middle of an otherwise excellent qualitative presentation. Such technical detail is unwarranted in a primer, superior mathematical discussions are available elsewhere, and, most seriously, the abrupt shift in presentation style will likely put many prospective readers off the entire chapter, much to their loss. Curious also is the authors' failure to provide any but the most rudimentary discussion of relative warps, which most practitioners find at least as useful as partial warps for achieving ordinations of objects in a linear shape space (but see Zelditch et al. 1995 and Rohlf 1998 for clues as to why the expected discussion of relative warps is virtually absent).

Part 2 provides an overview of multivariate ordination methods (Chapter 7), computer-based statistical methods (Chapter 8), multivariate analysis of variance (Chapter 9), regression (Chapter 10) and partial least-squares analysis (Chapter 11). Despite its title, Chapter 7 treats only principal components analysis (PCA) and canonical variates analysis (CVA). Inclusion of PCA is an obvious choice since so many of the 'warp' procedures are based on this numerical analysis method. The authors' presentation is basic, but good. Canonical variates analysis is a somewhat more unusual choice insofar as it is not the basis for any warp-based method and indeed violates one of the basic tenets of the shape theory on which geometric morphometrics is grounded.

As the authors discuss in Chapter 3, a—perhaps the—basic aspect of the morphometric synthesis that is geometric morphometrics was the realization that the correct basis of shape comparison was not the set of 'individual variables' as they would be construed in classic multivariate data analysis (see Manley

1994), but the abstracted shape itself as represented by *n* different landmarks. Thus, for the purposes of geometric shape comparison, a triangle of landmarks (Fig. 1A) is not a collection of three scalar distances (d_1, d_2, d_3) or six coordinate locations ($x_1,y_1; x_2, y_2; x_3,y_3$). It is a single variable that exists in a non-linear space or 'shape manifold'. This manifold quantifies the relative positions of all possible combinations of shapes that can be formed from three landmarks (Fig. 1B). In order to ordinate a set of triangular shapes correctly one must locate them on the shape manifold (an operation accomplished by Procrustes superposition), and project the positions of the observed triangles onto a linear plane oriented such that it is tangent to the three-landmark shape manifold at some position that makes sense with respect to the analysis (e.g., the mean shape). Operationally, this is accomplished in much then same way that three-dimensional coordinate positions on the Earth's surface a projected onto a piece of paper to construct a flat map. By taking advantage of this mathematical formalism, the resultant linear ordinations have the very desirable property[1] of allowing one to 'see through' their structure and 'sense' the geometry of shapes on the surface of the underlying shape manifold. However, the non-isotropic transformations used in canonical variates analysis to correctly represent distinctions between the within-groups and between-groups covariance matrices distort scaling relations within the shape tangent plane. Because of this it becomes impossible to make geometric interpretations of shape distributions within that plane that are accurate pictures of the underlying shape manifold. It's as if the canonical variates tangent plane had turned from a pool of clear water (the PCA-case that forms the analytic basis for principal, partial, and relative warps) into a funhouse mirror-like lens that stretches this space in one region and compresses it in another, giving a wholly inaccurate picture of the underlying—and already complex—non-Euclidean spatial relations. The authors fail to point this problem out. Readers might be left with the impression that there is less difference between PCA and CVA with respect to the way each handles landmark data than is actually the case.[2]

Chapter 8 is really about resampling methods (e.g., bootstrapping, jacknifing, Monte-Carlo simulation) that can be used to create probability density distributions for use with generalized shape-based hypothesis testing. Once again, these are basic, but competent reviews. The obvious omission here is the lack of even a mention, much less a discussion, of the Mardia-Dryden shape similarity test (Mardia and Dryden 1989), which was the first such test formulated under the rubric of geometric morphometrics and which is a far more simple and direct a method of assessing shape similarity than any of the techniques they discuss. Much the same can be said of the chapters on multivariate analysis of variance (MANOVA, Chapter 9) and regression analysis (Chapter 10, which includes sections on bivariate and multivariate regression). These are, again, competent reviews with well-chosen examples that show how these methods can be used to evaluate morphometric hypotheses. It does seem unusual, though, to see a more advanced topic like MANOVA presented before regression analysis; especially insofar as ANOVA and MANOVA are routinely used to evaluate the results of bivariate and multivariate regression analyses respectively (see my PalaeoMath columns for an example of a more traditional presentation). This section concludes with a discussion of partial least-squares analysis, another—albeit an even more esoteric—technique that can be used to assess the covariance between shape variation and variation in a variety of external variables (e.g., ontogeny, ecology, geography). This advanced numerical analysis method has huge potential, but it's a really adequate presentation—of which this is not as it lacks the background material necessary for true understanding—is well beyond the scope of what's needed in a primer.

The third section reviews example applications of morphometric analysis to the issues of morphological disparity (Chapter 12), evolution and development (Chapter 13) and systematics (by which the authors mean species and character recognition, Chapter 14). Instead of using examples from the technical literature the authors have opted for the presentation of example analysis. While this has the advantage of allowing the authors to go into more depth with the specific examples selected for presentation, to do this they have sacrificed much consideration of the breadth of studies that are undertaken in these areas; a deficiency that could have been mitigated by adding either introductory or concluding sections that reference other studies in these fields. Of these three chapters, all of which will repay a close reading, I think the chapter on ontogeny and evolution—the long-standing interest of Fink and Zelditch—holds up best. The text is rounded off by a short glossary of technical terms, a bibliography, and an index.

As I hope I've made clear, I have mixed views about the authors' success in achieving their stated aims. I'm quite sure *Geometric Morphometrics for Biologists: A Primer* will be popular among those who are predisposed to morphometric approaches to data analysis. This, of course, should include a fair number of palaeontologists. For those who plan to make a career in this field it should be regarded as required reading, as much for its failures as its successes. The danger here, of course is that some may regard the methods descriptions and example analyses as standard exemplars of good morphometric practice.

---

[1] For objects that may be represented by the same number of corresponding landmarks and are all basically similar in shape.

[2] See Klingenberg and Montiero in press for an attempt to address the problem of landmark-based CVA analysis within the context of geometric shape theory.

Suffice it to say that, in many cases, the advice and examples provided are controversial among other experienced practitioners.

Should members of the Palaeontological Association buy a copy? Probably not, at least not just now. Make sure your library orders a copy so you have access to one. Competition in this market will likely hot up in 2005 or 2006 when a much anticipated, and more comprehensive text, currently being written by F. James Rohlf, is published. Unless you are already committed to the field, I'd put off making a decision regarding the best general morphometrics reference book—for *Geometric Morphometrics for Biologists* is not a primer—to buy until a comparison can be made with the Rohlf book. If the latter turns out to be a specialist mathematical treatise, *Geometric Morphometrics for Biologists* will have to do for the moment. On the other hand, if the Rohlf book turns out to be written and organized along the lines of Sokal and Rohlf's (1995) highly respected *Biometry*, it will surely give *Geometric Morphometrics for Biologists* more than a run for your money. Still, both will probably be overkill in terms of what the causal practitioner really needs, which is that short, simple, low-cost primer Bookstein alluded to back in '96.

Norman MacLeod
Keeper of Palaeontology
The Natural History Museum

References cited

Adrain, J. M., Edgecombe, G. D. and Lieberman, B. S. 2001. *Fossils, phylogeny and form: an analytical approach*. Kluwer Academic/Plenum, New York. 402 pp.
Bookstein, F. L. 1991. *Morphometric tools for landmark data: geometry and biology*. Cambridge University Press, Cambridge. 435 pp.
Bookstein, F. L. 1996. Combining the tools of geometric morphometrics. 131–151. *In* **L. F. Marcus, M. Corti, A. Loy, G. J. P. Naylor and D. E. Slice** (eds.). *Advances in morphometrics*. Plenum Press, New York.
Bookstein, F. L. 1996. Landmark methods for forms without landmarks: Localizing group differences in outline shape. 279–289. *In* **A. Amini, F. L. Bookstein and D. Wilson** (eds.). *Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis*. IEEE Computer Society Press, San Francisco.
Bookstein, F. L. 1996. Landmark methods for forms without landmarks: Morphometrics of group differences in outline shape. *Medical Image Analysis* **1**(3), 1–20.
Costa, L., da F. and Cesar, R. M., Jr 2000. *Shape analysis and classification*. CRC Press, New York. 659 pp.
Dryden, I. L. and Mardia, K. V. 1998. *Statistical shape analysis*. J. W. Wiley, New York. 376 pp.
Elewa, A. M. T. 2004. *Morphometrics–Applications in Biology and Paleontology*. Springer-Verlag, Heidelberg, Germany. 263 pp.
Humphries, C. J. 2002. Homology, characters and continuous variables. 8–26. *In* **N. MacLeod and P. L. Forey** (eds.). *Morphology, shape and phylogeny*. Taylor & Francis, London.
Klingenberg, C. P. and Montinero, L. R. (in press). Distances and directions in multidimensional shape spaces: implications for morphometric applications. *Systematic Biology*.
MacLeod, N. 1999. Generalizing and extending the eigenshape method of shape visualization and analysis. *Paleobiology* **25**(1), 107–138.
MacLeod, N. and Forey, P. 2002. *Morphometrics, shape and phylogenetics*. Taylor & Francis, London. 308 pp.
Manley, B. F. J. 1994. *Multivariate statistical methods: a primer*. Chapman & Hall, Bury, St. Edmonds, Suffolk. 215 pp.
Marcus, L. F., Bello, E. and García-Valdecasas, A. 1993. *Contributions to morphometrics*. Museo Nacional de Ciencias Naturales 8, Madrid. 264 pp.
Marcus, L. F., Corti, M., Loy, A., Naylor, G. J. P. and Slice, D. E. 1996. NATO ASI Series. *Advances in morphometrics*. Plenum Press, New York. 587 pp.
Mardia, K. V. and Dryden, I. L. 1989. Shape distributions for landmark data. *Advances in Applied Probability* **21**, 742–755.
Rohlf, F. J. 1998. On applications of geometric morphometrics to studies of ontogeny and phylogeny. *Systematic Biology* **47**, 147–158.
Rohlf, F. J. and Bookstein, F. L. 1990. *Proceedings of the Michigan morphometrics workshop*. The University of Michigan Museum of Zoology Special Publication 2, Ann Arbor. 380 pp.
Small, C. G. 1996. *The statistical theory of shape*. Springer, New York. 227 pp.
Sokal, R. R. and Rohlf, F. J. 1995. *Biometry: the principles and practice of statistics in biological research (third edition)*. W. H. Freeman, New York. 887 pp.
Wilson, E. O. and Bossert, W. H. 1971. *A primer of population biology*. Sinauer Associates, Inc., Sunderland, Massachusetts. 192 pp.
Zelditch, M. L. 2001. *Beyond heterochrony: the evolution of development*. Wiley-Liss, New York. 371

pp.
Zelditch, M. L., Fink, W. L. and Swiderski, D. L. 1995. Morphometrics, homology, and phylogenetics: quantified characters as synapomorphies. *Systematic Biology* **44**, 179–189.

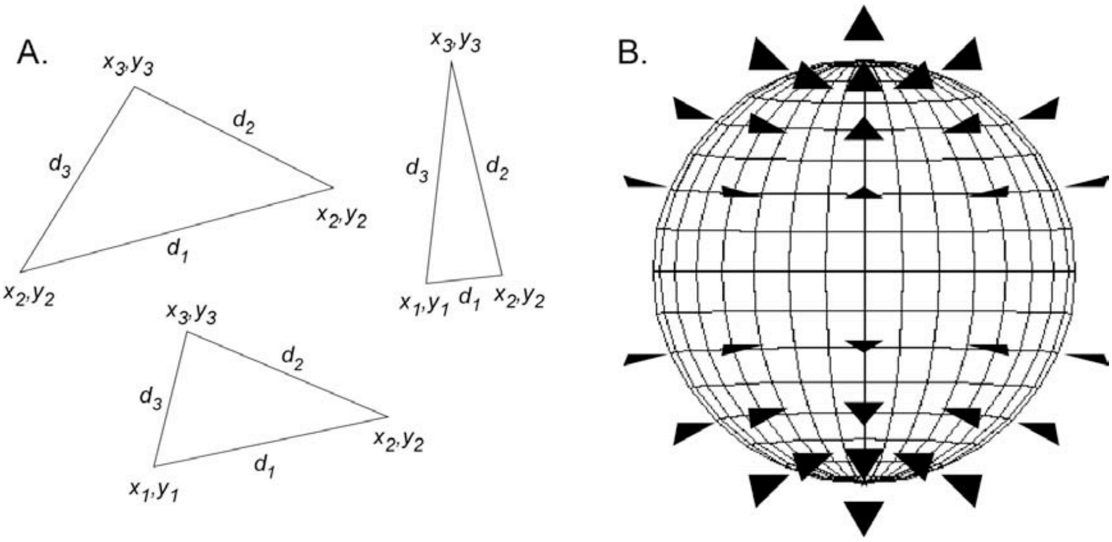

Figure 1. Alternative concepts of morphometric variables. A. traditional multivariate morphometric view of variables as either three inter-landmark distances (d_1, d_2, d_3) or six coordinate locations. This representation confounds size and shape variation. B Geometric morphometric concept of shape variation in which shapes themselves are regarded as the variables. Here, a representation of the shape manifold for all Procrustes registered triangles can be visualized as a sphere. On the surface of the sphere all possible combinations of triangles may be arranged such that distances from each other are scaled to the deviation from the mean shape in any direction summed over all three coordinates (= the Procrustes distance).