

Palaeo-Math 101

Principal and Partial Warps

While the topic of the last *Palaeo-Math 101* column, the thin plate spline (TPS), is used by most morphometricians as a technique for obtaining models of shape deformations, the mathematical machinery that stands behind those abstract little grid diagrams provides much more than a simple graphical device. To understand what thin plate splines are from an analytical point-of-view we need to back off a bit and consider the mathematical problem they try to solve.

Figure 1 shows the results of a *Procrustes* (GLS) superposition of landmarks from the trilobite genera *Acaste* and *Calymene* along with the resulting TPS grid. In this case *Acaste* was selected as the reference shape.

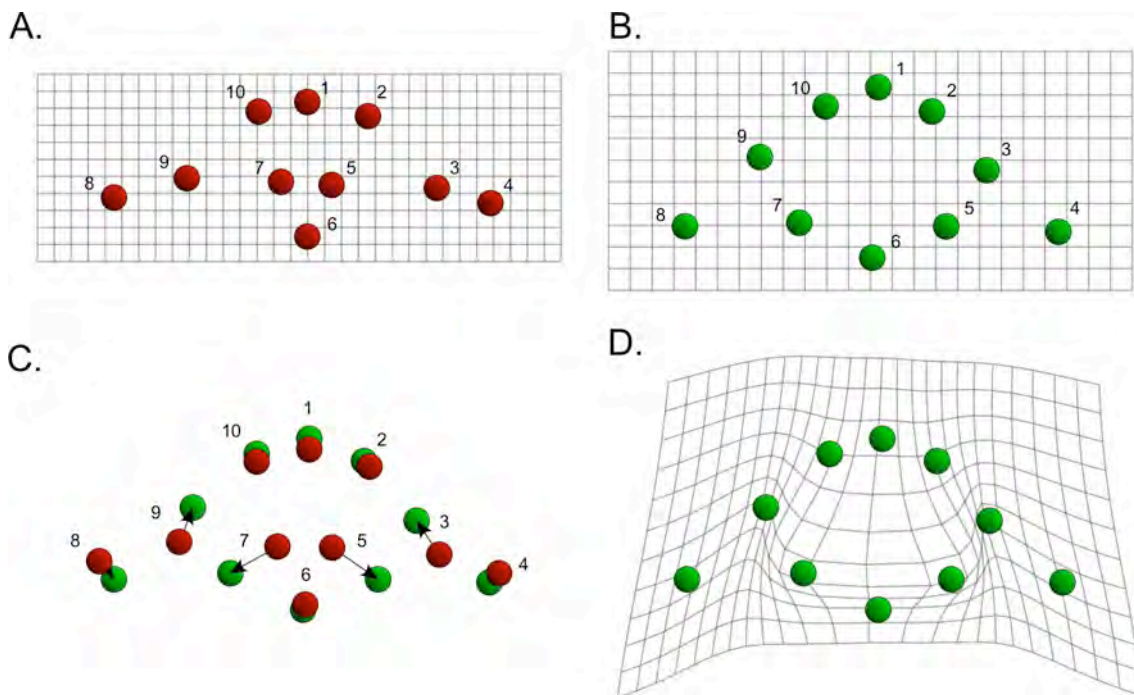


Figure 1. Graphic portrayal of the deformation implied by the transition between *Acaste* (A) and *Calymene* (B) based on an analysis of topologically homologous landmarks (1-10). *Procrustes* (GLS) superposition of landmarks with shape displacement vectors (C). Thin plate spline representation of the *Acaste* → *Calymene* deformation (D).

Note that the primary shape differences between these two forms reside in the locations of the eyes (landmarks 3 and 9), the position of the intersection of the glabella and posterior margin of the lateral projection (landmarks 5 and 7), and the position of the apex of the lateral projection (landmarks 4 and 8, see Fig. 1C). Accordingly, the TPS representation of this deformation shows strong displacement of the grid lines in the region of these landmarks and negligible grid deformations in the regions of the other landmarks. The important bit about the TPS representation, however, and the reason it's referred to as a spline, is that these strongly regionalized deformation patterns have been organized into a global model of non-linear displacements along the x and y axes that (1) mimics the character of a 3D surface in which the third axis (z) contains the displacement vector information and (2) appears to record the character of deformations in regions of the shape that have not been sampled by landmarks.

The TPS actually represents a method of solving a generalized problem in spatial statistics, namely the estimation of the value of a property (here length of the set of reference-target form displacement vectors) at an unsampled location based on the values of this property at neighbouring locations. In the field of spatial statistics this is a very common problem that is usually handled via a procedure called 'kriging' after its inventor, the South African mining engineer Daniel G. Krige.

Like the multivariate procedures I've discussed in previous columns, kriging is based on linear regression analysis. Unlike standard linear regression analysis though, it does not assume the

dependent variable is either completely random or distributed deterministically with respect to the spatial variables. Rather, it assumes the dependent variable is regionalized.

The idea of regionalized variables is a fundamental concept in spatial statistics.¹ These are, in a sense, variables that exhibit properties intermediate between those of a random variable whose pattern of variation obeys no rule and has no consistent structure, and a deterministic variable, whose pattern of variation is strictly rule-based and highly structured. Regionalized variables are continuous from point to point throughout the geometric space over which they are defined and can exhibit high correlations (= structure) over short distances. Nevertheless, the apparent consistency in the structure of their variation is inversely related to the distance between locations such that it's not usually possible to determine the rules by which variation is governed across the entire space. The solution to problems involving regionalized variables is to obtain a reasonable sample of variation at specific locations across the space of interest, use regression analysis-like strategies to estimate localized substructures in the dependent variable, and then to join these substructures into a single, continuous, global model.

As with all modelling procedures, the answer you obtain from a kriging analysis is, to a large extent, determined by a set of assumptions relating to the structure of covariances that exist between locations across the space of interest. In the case of the TPS, this set of assumptions is encoded by the bending energy matrix, which assumes that variations between regions of the shape are structured as though the deformation is mimicking the behaviour that would be expected from the physical deformation of an infinitely thin metal plate.

When metal plates are bent the physical energy that goes into deforming them is distributed over the entire plate in such a manner as to cause the energy required to hold the bend at any point on the plate to be minimized. In real metal plates flaws in their structure usually cause the bending energy to be focused in the region of the flaw. If this energy exceeds the strength of the material in the region of the flaw, the plate kinks or tears. But in a hypothetical, perfect metal plate of infinite strength the distribution of energy will be smooth and solely dependent on the spatial scale of the bend. In other words, it will take relatively little energy to achieve a broad bend that involves the whole plate, a larger amount of energy to achieve a small, but localized bend in one region of the plate, and quite a lot of energy to achieve a large, localized bend in only one small part of the plate. For those readers who recall the metalworking section of their 'Shop' or 'Practical Skills' classes in secondary school, this should accord with personal experience.

Although no one is so naïve as to pretend that organismal bodies are metal plates or that natural processes (e.g., development, evolution) are constrained to minimize the magnitude of deformations in a manner inverse to their spatial scale, the metal plate metaphor has desirable properties in terms of the standard statistical models we use to describe and model variation and change in many different contexts. Chief among these is the global minimization of deviations: the least squares model. Add this to the straight-forward assumption that spatial covariation across an area is structured in a manner that is uniform in all directions and conforms to a function that is the strict inverse of spatial scale and you have the essence of the TPS solution to the standard kriging problem. Additional mathematically convenient aspects of the TPS approach are that (1) all TPS interpolations form surfaces that are smooth at all scales, (2) the TPS model is completely determined, which is to say it needs no *ad hoc* manual tuning, and (3) all parameters needed to specify the TPS model can be estimated by solving a series of linear equations. As a general approach to fitting a continuous, global model of 3D point distributions to sparse data, the TPS is simple, elegant, visually striking, and consistent with the manner in which we're used to thinking about statistical descriptions of change in any number of parameters, including shape.

So, how can we get the TPS to produce an analytical — as opposed to a strictly graphical — model of shape change in a sample of landmarks and, once we've got that, what (if anything) can we do with it? In the last column I showed you how to calculate a TPS models for individual deformations between reference and target forms. In order to explain how to the TPS formalism has been used in an analytical context there are a couple of things I need to remind you of.

The first of these is that, in all instances, the geometry of the TPS of any landmark configuration is determined entirely by the reference form. A reference form is needed to serve as the basis for the calculation of the landmark displacement vectors in the *Procrustes* space on which the spline is calculated. To take the simplest example, the TPS of a reference form compared with

¹ Actually, morphometrics is a branch of spatial data analysis and spatial statistics (see Davis 2002).

itself is a perfectly flat, undeformed, rectilinear grid. This obtains because the lengths of all the displacement vectors in such a comparison are 0.0. Because all the displacement vectors in such a comparison are 0.0, the overall bending energy of the deformation is also 0.0. For all other landmark configurations irrespective of whether those configurations are realized in the manner or actual specimens or not, a set of displacement vectors will be specified. The geometry of these vectors relative to the reference form will allow a non-perfectly rectilinear, and in most cases non-flat, TPS grid to be calculated; in the case of the latter along with an associated bending energy.

As you've not doubt noticed I had to be careful with the wording of the sentence above. This is because of the hierarchy of geometric deformations that are possible in two and three-dimensional forms. Recall these can be subdivided into two groups: uniform and non-uniform (Fig. 2, see also Fig. 5 of MacLeod, 2010). Note that the two uniform deformation modes not removed by *Procrustes* superposition can be described mathematically by applying exactly same proportionate degree of deformation to each and every landmark location, such that the lengths of the implied landmark displacements are either exactly the same for all landmarks but oriented in different directions (compression-dilation) or are linearly proportional to the scale of the distance between non-displaced landmarks and oriented in the same direction (shear). Because of this regularity in the structure of the displacement vectors, the TPS grids resulting from uniform deformations remain strictly planar surfaces. Since these uniform deformation surfaces exhibit no global or localized displacements, their interpolated TPS surfaces are not 'bent' and so have no associated bending energy. However, outside these two special cases of geometric shape transformation, all others exhibit heterogeneous distributions of displacement vectors that give rise to variably bent or warped TPS grid geometries along with associated bending energies.

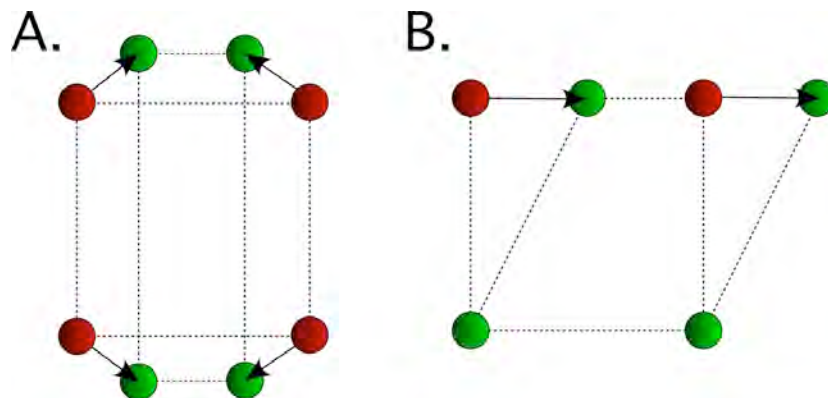


Figure 2. The two uniform shape deformation modes not corrected for by *Procrustes* superposition: compression-dilation (A) and shear (B). Arrows represent deformation vectors between reference (red) and target (green) forms. These classes of deformation will produce TPS grid geometries in which only the reference grid aspect ratio has been altered. See text for further discussion.

Perhaps the most unusual aspect of the TPS formulation is that it's not only the case that the spline is graphically dependent on the reference shape, all the standard bending energy calculations are referenced uniquely to the reference shape too. This makes sense because of the physical metaphor that lies at the heart of the TPS model — that bending energy is minimized across the space and that the spatial configuration of the reference form's landmarks controls the local vs. global deformation model. From an analytical point-of-view though, this places some subtle and easily overlooked constraints on the interpretation of TPS/bending energy analysis results.

The most critical of these constraints is an appreciation of the importance of selecting an appropriate reference shape. Recall in the column on shape theory (MacLeod 2009) I made the point that the reference shape controls the orientation of the tangent plane onto which the shapes that exist on the surface of the *Procrustes* shape hemisphere can be projected in order to obtain a linear, map-like ordination of shape variation based on their *Procrustes* distances. In principle, any shape that contains the same number of landmarks as the shapes in your sample could serve as the reference shape. But the single shape that best represents the distribution of

shapes in any sample is the mean shape. This is the shape that minimizes the overall deviation of landmarks from one another. As a result, the mean shape is also the shape that has the greatest overall similarity with all other shapes across the sample.

In some instances, and for some types of analyses, it might seem logical to choose some shape other than the mean shape to serve as the reference. For example, in a taxonomic study it might seem reasonable to use the shape of the holotype as a reference. Similarly, in a study on ontogenetic shape change it might seem appropriate to select the earliest or the latest developmental stage as the reference shape and compare all other shapes in the sample to that. Unfortunately, these will, in almost all cases, lead to a needlessly distorted ordination of shapes within the space of the plane tangent to the *Procrustes* shape hemisphere at the coordinate location of those potential reference shapes.

To illustrate the importance of this issue, let's take a simple example that involves use of the TPS to make a comparison of the structure of the bending energy matrix for alternative reference forms. You will recall that the TPS calculations are based on the bending energy matrix (L_p^{-1}) where L_p is as follows.

$$L_p = \begin{bmatrix} 0 & U_{1,2} & U_{1,3} & \cdots & U_{1,p} \\ U_{2,1} & 0 & U_{2,3} & \cdots & U_{2,p} \\ U_{3,1} & U_{3,2} & 0 & \cdots & U_{3,p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ U_{p,1} & U_{p,2} & U_{p,3} & \cdots & 0 \end{bmatrix} \quad (20.1)$$

In this equation, which is identical to equation 19.3 of the previous newsletter column (MacLeod 2010), recall U is a measure of the distances between landmarks in the reference shape (see Equation 19.1). The inverse of this matrix establishes the metaphor of pure, homogeneous bending energy in the sense that it is the simple inverse function of inter-landmark proximity.

The bending energy matrix can't be visualized in its entirety using a TPS grid because that technique requires a contrast in landmark configurations between reference and target forms in order to supply the landmark displacement vectors. However, since the bending energy matrix is a symmetric, square matrix, it is susceptible to linear decomposition via eigenanalysis in precisely the same manner as we've decomposed covariance, correlation, distance, and other sorts of similarity matrices throughout this essay series. There is an important difference between the eigenanalysis of the bending energy matrix and the eigenanalysis of those other matrices though, and it's this difference that really gets to the heart of the reference shape issue.

In all previous applications of eigenanalysis we've discussed we were decomposing a matrix that represented r -mode and/or Q -mode similarity/dissimilarity matrices between all pairs of objects in a sample. Eigenanalysis of such matrices results in the production of a set of orthogonal vectors that are aligned with directions of maximum variation or distance or similarity across the sample. If we've chosen our sample correctly, those directions also estimate the directions of maximum variation or distance or similarity in the parent population from which our sample was drawn.

The difference in the case of the bending energy matrix is that it's a matrix composed of distances between landmark positions drawn *from a single object or specimen*. Eigenanalysis of this matrix, when combined with the coordinate locations of the reference shape landmarks themselves, produces a set of orthogonal fields or modes of variation aligned with the directions of minimum landmark dispersion (= maximum bending energy) in the set of landmarks that describe this single object or specimen. Since the dispersion of landmarks is related directly to spatial scale, this means that, in addition to being aligned with the directions of minimum landmark dispersion, these modes of form or shape variation will also be ordered in terms of spatial scale. Eigenvectors of the bending energy matrix that account for the highest bending energies will represent modes of deformation characterized by large deviations over small spatial scales. Those accounting for the lowest bending energies will represent modes of deformation characterized by small deviations over large spatial scales.

Bookstein (1989, 1991; see also Rohlf 1993) have termed the eigenvectors of the bending energy matrix (L_p^{-1}) 'principal warps' drawing on the clear and compelling analogy with principal components analysis. These authors also referred to the eigenvalues associated with those vectors 'principal values'. In contrast, Slice *et al.* (1996) termed these same eigenvectors 'partial warps' in the sense that they describe parts of the deformation pattern inherent in the bending energy matrix. This dual terminology has led to much confusion especially insofar as Bookstein (1989, 1991) had already used the term 'partial warps' for the result of a decidedly different procedure (see below). Despite claims that the Slice *et al.* (1996) terminology has become 'standard' (e.g., Zelditch *et al.* 2004)², for the purposes of this discussion I will employ to the original terminology.

Because aspects of shape variation are removed from each landmark set during its conversion to shape coordinates, there are only $k-3$ positive principal values where k is the total number of landmarks used to sample the form. For the ten landmarks used to quantify cranial variation in our trilobite genera then, eigenanalysis yielded seven vectors with positive eigenvalues or seven principal warps. Four of these are shown for each of three reference configuration in Figure 3. Since any landmark configuration can be used as the basis for a principal warp calculation Figure 3 includes principal warps calculated for two real specimens (*Acaste*, *Calymene*) and one hypothetical configuration; the mean of consensus shape for the 18 trilobite specimens on which these ten landmarks can be located.

The principal values (λ) for each principal warp are shown below the TPS grids in Figure 3, expressed as a percent contribution to the overall bending energy for each alternative reference shape. Although the principal warps have no intrinsic deformation — after all, they are calculated for a single specimen — an external scaling factor is usually applied to supply the deformation magnitudes required by the TPS calculations. Setting this scaling factor to a constant allows the spatial heterogeneities implicit in the set of reference shape-specific principal warps to be displayed graphically. The arbitrary scaling factor selected for the calculation of all TPS grids in Figure 3 was 0.206.

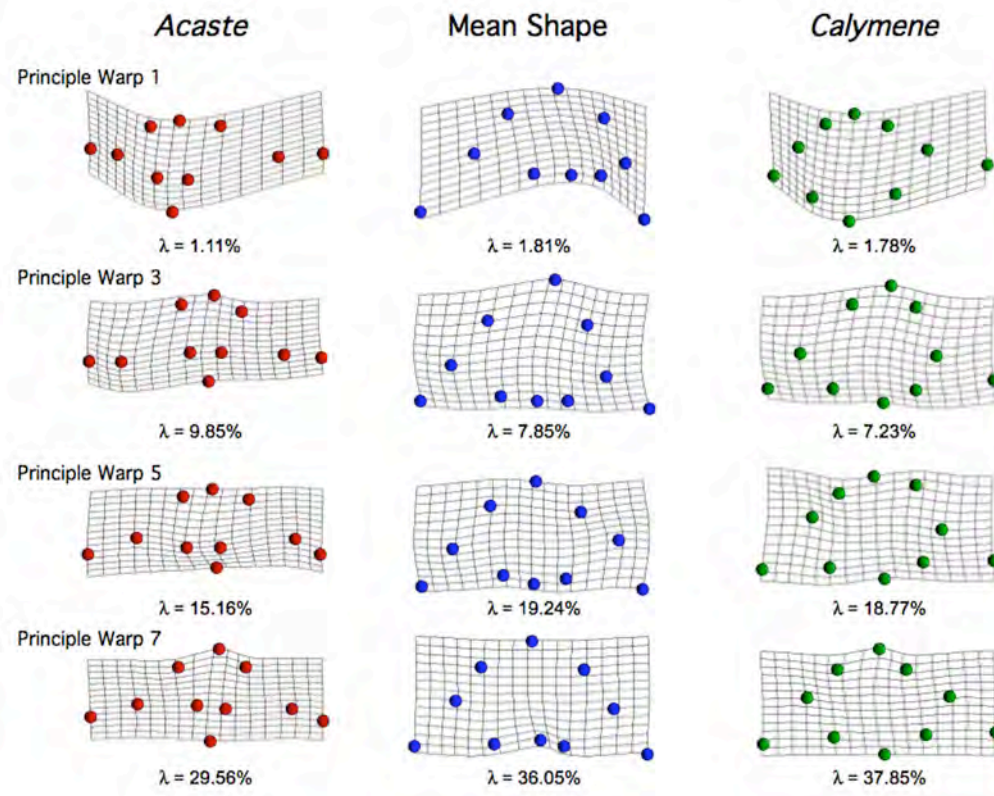


Figure 3. Selected TPS deformation grids for principal warps calculated from the *Acaste* (left), *Calymene* (right) and trilobite sample mean or consensus shape (centre). Note the wide range of variation inherent in the geometry of these TPC decompositions. See text for additional discussion.

² Jim Rohlf's morphometrics program packages for computing principal and partial warps (tpsSpln and tpsRelw) — which are the industry standards in this field — use the original terms for these procedures.

As you can see from these grids, the principal warp decomposition for each alternative reference shape yields a set of increasingly more localized deformational geometries. In each case principal warp 1 specifies a broad deformation that encompasses the entire landmark set. These high-level deformation patterns contrast strongly with the deformations expressed by principal warp 7, each of which predominantly involves relative adjustments in the positions of a pair of adjacent landmarks near the mid-line and at either the anterior (*Acaste*, *Calymene*) or posterior (mean shape) ends of the cranidia. The deformation patterns expressed by principal warps 3 and 5 are, in all cases, intermediate between these extremes.

By the same token though, it should be noted that, while the deformation sequences for *Acaste*, *Calymene*, and the mean shape are all consistent and make reasonable sense by themselves, there seems little geometric similarity between these deformation sequences. Principal warps 1 and 7 for *Acaste* and *Calymene* are somewhat similar geometrically, but both of these differ markedly from the mean shape's principal warps 1 and 7. In contrast, principal warps 3 and 5 of *Acaste* and *Calymene* exhibit marked differences whereas they are broadly similar for *Calymene* and the mean shape.

The point is that each potential reference shape — including the mean shape and especially for mean shapes calculated from samples containing high shape variation and low sample size — is going to incorporate atypical or idiosyncratic landmark placements to a greater or lesser extent. Because of the nature of the principal warps, these idiosyncratic differences will lead to broad and chaotic incompatibilities between the principal warp shape spaces calculated on the basis of individual specimens. Selection of the mean shape as the reference configuration will minimize this tendency to some extent depending on how well constrained and representative the mean is with respect to the shapes included in the overall sample. But even use of the mean shape as a basis for these calculations will not stabilize the principal warps space entirely. We will return to a discussion of the implications of the inherent instability and idiosyncratic nature of the principal warps sequence below.

Once specified, the principal warps of the reference configuration can be used as the mathematical basis for the creation of a linearized space within which any shape described by sets of corresponding landmarks may be projected. To make connection with the previous essay on shape theory (MacLeod 2009), the bending energy matrix represents the plane tangent to the *Procrustes* shape hemisphere at the point of the reference shape. The principal warps represent a set of orthogonal variables that re-describe the bending energy matrix as a series of spatially ordered modes of shape variation. Bookstein (1989, 1991), Rohlf (1993), and many others have referred to the projection of forms defined by comparable sets of landmarks and transformed into the principal warps space as 'partial warps'. The representation of these projections can take two forms.

The first, and possibly most analytically useful of these is to represent the projection in the form of a scatterplot of scores of projected shapes along a space defined by the x and y principal warp vectors³ (Fig. 4). These scatterplots represent ordinations of between-specimen shape similarity and/or difference with respect to those aspects of shape deformation being represented by the principal warp. The advantage of this sort of analysis is that, because the principal warp is referenced to a single specimen, the nature of the space so defined will not change with the acquisition of new specimens, removal of non-reference specimens due to taxonomic revision, etc. Recall this is not the case with the vast majority of standard multivariate data analysis methods (e.g., PCA, FA, PCoord, CVA, MDS) because they require a representation of similarity across a sample of objects. This, in turn, requires that the sample remain intact for the results of these analyses to remain meaningful. Change the sample in any way (e.g., drop some specimens out of the sample because of taxonomic revision, add some specimens to the sample because of new discovery) and you must re-compute all the results of these analyses for the patterns expressed to remain optimal and valid. This sample dependence is avoided in a principal warps analysis. So long as the reference shape remains valid, it defines the principal warps space. Any shape described by a comparable set of landmarks may be projected into and/or removed from this principal warps space without altering the nature of that space in any way.

The principle disadvantage of a principal warps analysis is exactly the same. Because the ordination space created is referenced to a single specimen, the influence of that specimen is

³ Or x , y , and z axes in the case of 3D landmark data.

absolute. Since the spaces so created are not optimized over a sample of specimens they represent nothing more (or less) than the shape characteristics of that single specimen albeit one than might be a hypothetical best single representation of a sample or population such as the mean shape. But even in this case, all that is being used in the analysis is the raw spatial configuration of hypothetical landmark points without any associated indication of presence, much less the extent or character, of within-sample shape variation.

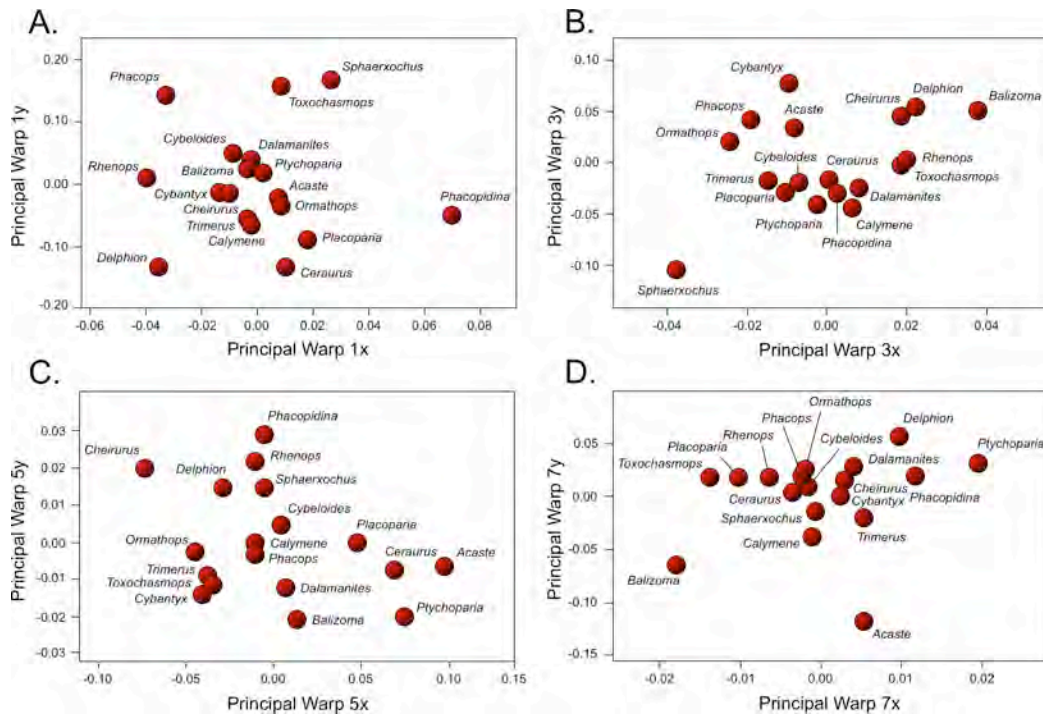


Figure 4. Scatterplots of trilobite partial warp scores on principal warps 1 (A), 3 (B), 5 (C), and 7 (D) calculate using the trilobite sample mean shape as the reference shape.

The other manner in which partial warps have been used is to gain a visual sense of the deformational geometries being expressed by the distribution of partial warps scores in the principal warps ordination space via TPS modelling. Figure 5 provides examples of such models for selected specimens whose ordination locations are shown in Figure 4. These models are heuristic devices that can be very useful in making qualitative interpretations of the shape ordination results and/or explaining the implications of those results to non-quantitative colleagues in a manner they can appreciate and understand.

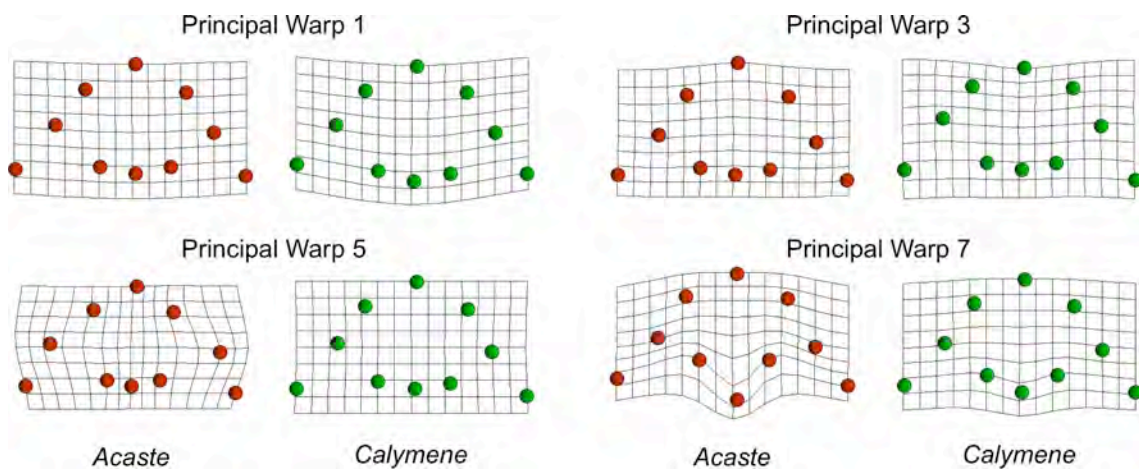


Figure 5. Partial warp TPS grids for the *Acaste* and *Calymene* landmark configurations on the principal warps shown in Fig. 4. This grid represent the TPS interpolations from the reference shape in four of the eigenvector decompositions of the mean shape's bending energy matrix.

Whereas the calculation of principal warps is quite easy computationally, and the manner in which they manage to support the creation of shape variables in which the deformational modes are ordered in terms of their spatial generality is quite elegant mathematically, their utility as analytical tools is, unfortunately, compromised by their inherent instability and absolute dependence on the configuration of a single set of landmarks. In the early days of geometric morphometrics it was more-or-less informally thought that the high-energy principal warps might be sufficiently localized spatially to represent taxonomic characters, developmental modules, and/or any of a number of other biological concepts based on the subdivision of a complex morphology into component parts, an assessment of whose shape would be useful.

An example of this was an attempt by Zelditch *et al.* (1995) to use principal warps analysis to define character states that could then be coded for use in a phylogenetic analysis. In a comment on this suggestion Rohlf (1998) noted that the inherent instability of principal warps spaces made ordinations of partial warp scores in the those spaces ill-suited for use in the context of the shape-based characterization of sets of morphometric data. Moreover, the *ad hoc* mathematical decomposition of bending energy matrices defined on the basis of the arbitrary selection of a single specimen conforms to no recognizable theory of biological homology; the theory that stands at the heart of the character concept. Rohlf went on to suggest that an analysis of geometric shape variation that was referenced to a sample of shapes under consider would be more a more appropriate approach to this general problem. Later, I followed Rohlf's suggestion, albeit in a slightly different shape-analytical context, in an explicit test of the ability of morphometric data to provide insight into phylogenetic character state definition (MacLeod 2002), Still later Zelditch and colleagues acknowledged the limitations of their previous use of principal warps analysis in this context (Zelditch *et al.* 2004).

Presently principal warps analysis represents something of a blind alley in morphometrics. From time-to-time you run across this strategy being used to ordinate shapes and test shape-related hypotheses (Naylor 1996, Rohlf *et al.* 1996). But these are usually example analyses whose real purpose is to illustrate the principal-partial warps technique than to use it as a tool to test biological hypotheses. Principal and partial warps concepts and calculations are also covered in most textbooks on morphometrics both older and new (Bookstein 1991, Reyment 1991, Dryden and Mardia 1998, Costa and Cesar 2000, Zelditch *et al.* 2004) despite their lack of a track record of clear and unambiguous utility and in the face of reasonably trenchant criticisms that have been levelled at the (comparatively few) investigations in which they have been employed. I suspect one of the main reasons interest in principal/partial warps survives is because their calculation is included in a number of standard morphometrics software packages. Prominent among these is Jim Rohlf's tpsRelw package in which the ordering of the calculation steps implies to many that determination of principal warps is a necessary precursor to the calculation of relative warps, which have always been considered far more useful than principal-partial warps. I'll use the next essay to explore this issue in the context of a description of the relative warps technique. In any event, the foregoing discussion is presented to inform the reader as to what principal and partial warps are, how they relate to TPS, to set the stage for the our discussion of how they relate to relative warps, and to emphasize that, if this approach to shape analysis is used at all, it should be with caution.

Norman MacLeod

Palaeontology Department, The Natural History Museum
N.MacLeod@nhm.ac.uk

REFERENCES

- BOOKSTEIN, F. L. 1989. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11, 567–585.
- BOOKSTEIN, F. L. 1991. *Morphometric tools for landmark data: geometry and biology*. Cambridge, Cambridge University Press, 435 pp.
- COSTA, L., da F., and R. M. CESAR, Jr. 2000. *Shape analysis and classification: Theory and Practice*. CRC Press, New York. 680 pp.
- DAVIS, J. C. 2002. *Statistics and data analysis in geology (third edition)*. John Wiley and Sons, New York. 638 pp.

- DRYDEN, I. L., and K. V. MARDIA. 1998. *Statistical shape analysis*. J. W. Wiley, New York. 376 pp.
- MacLEOD, N. 2002. Phylogenetic signals in morphometric data. In N. MacLEOD, and P. L. FOREY, eds. *Morphology, shape and phylogeny*. Taylor & Francis, London. 100–138 pp.
- MacLEOD, N. 2009. Shape theory. *Palaeontological Association Newsletter*, **71**, 34–47.
- MacLEOD, N. 2010. Shape models II: the thin plate spline. *Palaeontological Association Newsletter*, **73**, 24–39.
- NAYLOR, G. J. P. 1996. Can partial warps scores be used as cladistic characters? In L. F. MARCUS, E. BELLO, and A. GARCÍA-VALDECASAS, eds. *Contributions to Morphometrics*. Museo Nacional de Ciencias Naturales 8, Madrid. 519–530.
- ROHLF, F. J. 1993. *Relative warp analysis and an example of its application to mosquito wings*. In L. F. MARCUS, E. BELLO, and A. GARCÍA-VALDECASAS, eds. *Contributions to Morphometrics*. Museo Nacional de Ciencias Naturales 8, Madrid. 131–160 pp.
- ROHLF, F. J. 1998. On applications of geometric morphometrics to studies of ontogeny and phylogeny. *Systematic Biology*, **47**, 147–158.
- ROHLF, F. J., A. LOY, and M. CORTI. 1996. Morphometric analysis of old world Talpidae (Mammalia, Insectivora) using partial warp scores. *Systematic Biology*, **45**, 344–362.
- SLICE, D. E., F. BOOKSTEIN, L., L. F. MARCUS, and F. J. ROHLF. 1996. *Appendix I: a glossary for geometric morphometrics*. In L. F. MARCUS, M. CORTI, A. LOY, G. J. P. NAYLOR, and D. E. SLICE, eds. *Advances in morphometrics*. Plenum Press, New York and London. 531–551 pp.
- ZELDITCH, M. L., W. L. FINK, and D. L. SWIDERSKI. 1995. Morphometrics, homology, and phylogenetics: quantified characters as synapomorphies. *Systematic Biology*, **44**, 179–189.
- ZELDITCH, M. L., D. L. SWIDERSKI, H. D. SHEETS, and W. L. FINK. 2004. *Geometric morphometrics for biologists: a primer*. Elsevier/Academic Press, Amsterdam, 443 pp.

Don't forget the *Palaeo-math 101-2* web page, now at a new home at:
http://www.palass.org/modules.php?name=palaeo_math&page=1