A new superimposition method usable for comparing the shape of several maps

Abstract:

Bidimensional regression was developed by Waldo Tobler since the 1960s to evaluate the resemblance between two plans. This is the most used way in geography to highlight differences in shape between maps. However, bidimensional regression can only compare the shape of two background maps simultaneously. This article presents two ways of generalizing bidimensional regression in order to be able to compare more than two geographic maps using a geometric morphometric approach. For this purpose, we took inspiration from the generalization of the ordinary Procrustes algorithm. An example shows that the generalized bidimensional regression makes it possible to compare the shape, and therefore the structure of several cognitive maps thanks to the creation of a morphospace. More precisely, it makes it possible to produce a typology of maps based on their shape on the one hand, and to identify the main variations in shape of all the maps analyzed on the other.

Key-words: generalized bidimensional regression, geometric morphometrics, superimposition method, morphospace, cognitive mapping

Introduction

Bidimensional regression (BDR) is a statistical method of adjustment for measuring the degree of resemblance between two planes structured by two sets of homologous points (Tobler 1965; Tobler 1978a, b; Tobler 1994). BDR has obviously found numerous applications in cartography. For example, it was used to measure the level of precision between an ancient map and a modern map (Tobler 1994; Symington et al. 2002; Lloyd and Lilley 2009; Yabe 2024), or between two geolocation technologies (Helbich et al. 2012). But BDR is not limited to Euclidean spaces. It can be used to compare the distribution of places on a topographic map with those, counterparts of the first, located in a space constructed on non-Euclidean metrics such as cognitive distances (Cauvin et al. 1998; Cauvin 2002; Friedman and Kohler 2003; Kitchin 1993; Kitchin and Blades 2002), cost distances or time distances (Ahmed and Miller 2007; Cauvin 2005; Dusek 2011; Dusek 2012)**.** The non-Euclidean structure of such spaces are sometimes difficult to understand because the rules of Euclidean geometry, such as the triangle inequality[1], are not always respected. A major advantage of BDR is that it offers a solution to analyze the resemblance of shape through its Euclidean model (Tobler 1994). However, the shape reveals the structure (Cauvin 1997). That of non-Euclidean spaces can thus be highlighted thanks to the morphological comparison capabilities of BDR.

Yet, by its very principle, BDR is limited to the shape comparison of only two maps at a time. How then can we compare the shape of several maps to show similarities and dissimilarities in structure? How can we identify a typical shape model to which to connect several ancient maps, cognitive representations of space or time-space maps? How can we identify key trends in the shape variations of a set of maps using Tobler's regression? The disaggregated analysis of BDR results (Kitchin and Fotheringham 1997) provides a path of research but it only concerns statistical results and not geometric results.

The solution we propose is to generalize the BDR algorithm in order to integrate it into a relatively recent statistical approach to shape analysis called geometric morphometrics (GMM). This new field of research was defined by Bookstein (1992) as the statistical analysis of shape variations and their covariations with other variables, where shape is defined by Kendall (1977) as all the geometric information that remains when location, scale and rotation effects are filtered out of an object. In the eighties, this new area of research appeared to characterize shape using size-independent variables based on Cartesian coordinates of homologous landmarks, clearly identified on each object (Bookstein 1986; Rohlf and Slice 1990). The shape is then captured and analyzed as such directly by its coordinates and no longer, as previously, from measurements which included the size in addition to the shape.

---

[1] Triangle Inequality Theorem states that the sum of the lengths of any two sides of a triangle must be greater than the length of the third side.

This method is applied in many scientific disciplines (biology, paleontology, botany, entomology, archaeology, anthropology, etc.) for the comparative statistical analysis of several related shapes (Adams et al. 2013; Dryden and Mardia 2016; Mitteroecker and Schaefer 2022). The general principle and analytical potential of GMM in cartography was recently presented by Roulier (2023) who took as an example the comparison of the shape of several ancient maps of France with a modern map.

The objective of this article is then to present generalized bidimensional regression (GBDR) and its integration into a GMM process in order to compare several maps or spatial structures. The very general principal of the GMM is based on three main steps: 1) collection of homologous points called "landmarks" on several planes, 2) superimposition of these landmarks to the reference shape and 3) calculation of a morphospace by a principal component analysis based on the results of the superimposition. The central step of GMM is then represented by the superimposition step carried out by an iterative geometric adjustment procedure. The adjustment method usually used in GMM is GPA (generalized Procrustes analysis) which corresponds to the generalization of the ordinary Procrustes analysis (OPA). However, OPA has been considered equivalent to Euclidean BDR (Kern 2017). This is why we can present here a generalization of Tobler's algorithm in order to propose an alternative to GPA for GMM. This article will present the basic principles of bidimensional regression as well as the solutions envisaged to overcome its limitation to two maps. Two algorithms for generalizing the ordinary BDR will be detailed. A comparative test will be established between GPA and GBDR on the main results. Finally, an example relating to an investigation in spatial cognition will highlight the effectiveness of GBDR when used in a GMM analysis for the simultaneous comparison of several "mental maps".

2. BDR for comparing two maps at a time

2.1 General principle of BDR

Before explaining GBDR, it is important to understand ordinary BDR. In a strict sense, BDR refers to an adjustment function and is basically an extension of linear regression (Tobler 1994). From a general point of view, BDR compares two bidimensional variables, denoted Z and W corresponding to two set of n homologous points structuring two surfaces, with the objective of measuring the degree of resemblance between them (Fig. 1). BDR seeks first to determine the parameters of a function f which, applied to the bidimensional variable W (u and v values) makes it possible to obtain a new bidimensional variable Z' (x' and y' values), as close as possible within the meaning of least squares to the bidimensional variable Z (x and y values). In other words, $Z' = f(W)$, so that the residual values $Z - Z'$ are minimized by a least-squares solution. We postulate here that knowing the data of the variable W makes it possible to predict those of the variable Z' and that we fit the variable W on the variable Z. If we keep the filiation with the unidimensional regression, we then can say that the variable Z corresponds to the dependent (to be explained) variable and that the variable W corresponds to the independent (explanatory) variable.
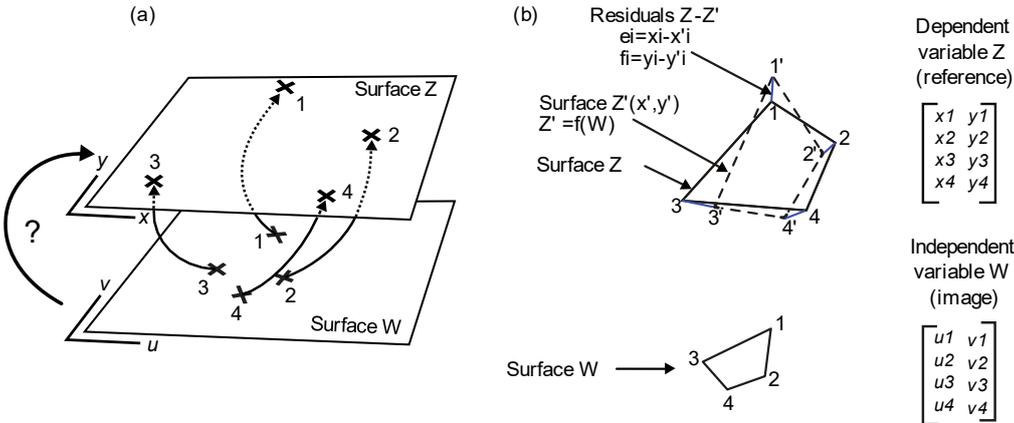
Fig. 1 The Euclidean BDR: (a) the problem to be solved; (b) fitting the surface W to the surface Z

The minimization of residuals can be obtained using a Euclidean model which apply a rigid transformation performed by simultaneous moving, rotating and scaling the bidimensional independent variable. Tobler (1994) developed two other linear transformation models which are the affine and projective models, but Kern (2017) established that only the Euclidean model of BDR was equivalent to the Procrustes method used in GMM. When

we fit the independent variable W on the dependent variable Z, the basic formula of Euclidean BDR is the following in matrix form:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \alpha 1 \\ \alpha 2 \end{pmatrix} + \begin{pmatrix} \beta 1 & -\beta 2 \\ \beta 2 & \beta 1 \end{pmatrix} . \begin{pmatrix} u \\ v \end{pmatrix} \tag{1}$$

Where (x', y') are the coordinates adjusted by BDR and (u, v) are the coordinates of the variable W. The parameters $\alpha 1$ and $\alpha 2$ apply a translation respectively on the x axis and on the y axis[2] while the parameters $\beta 1$ and $\beta 2$ are used to calculate a rotation and a scaling[3].

The BDR method is enriched with numerous statistics. The main indicator is the coefficient r^2 which represents the quality of the fit between the two sets of coordinates (explained variance varying from 0 to 1). The second main indicator is the root mean squared error (RMSE) which represents the average deviation between the position of the reference points and that of the adjusted points, expressed in the unit of the reference map.

$$r^2 = 1 - \frac{\sum \left[(x'-x)^2 + (y'-y)^2 \right]}{\sum \left[(x-\bar{x})^2 + (y-\bar{y})^2 \right]} \tag{2}$$

Where $\bar{x}$ is the mean of x and $\bar{y}$ is the mean of y.

$$RMSE = \sqrt{\sum \left[(x'-x)^2 + (y'-y)^2 \right] / n} \tag{3}$$

2.2 Comparing the shape of two maps

From a cartographic point of view, the dependent variable Z is usually a reference map, typically a topographic map and the adjusted independent variable W corresponds to a surface called image map that we wish to compare to the reference map[4] (an ancient map or a cognitive representation of space for example). In this case, it makes sense to express the values of the dependent variable in a geographic coordinate system, so that the RMSE can be expressed in a distance unit (meters, kilometers, miles, etc.). The adjustment produces new landmarks located as close as possible to the landmarks on the reference map. The BDR residuals, i.e. the positional differences between the reference points and the adjusted image points, correspond to the degree of structural resemblance between the two maps. These differences can be considered as displacement vectors whose origins are the reference points.

By removing the effects of position, rotation and size in its adjustment process, the BDR also makes it possible to show the difference in shape, as defined by Kendall in the introduction of this article. This is why the ordinary BDR can already be considered as a morphometric analysis tool. When the dependent variable is a reference map, an interpolation step can complete the adjustment step to visually highlight these differences in shapes between the two maps. The interpolation will generalize the residuals over the entire surface of the reference map, which results in a deformation, an anamorphosis making the differences in structure visually accessible. This is why Tobler closely associates a linear interpolation with the adjustment step (Tobler 1978a, 1994). Other methods were used to generalize the residuals. Symington and colleagues (2002) for example use a local polynomial regression. Multiquadric interpolation (Hardy 1971) has also been used in association with mapping adjustment (Jenny and Hurny 2011; Roulier 2018). Therefore, a major advantage of BDR in cartography is that it is able to show, through the visible change in shape of a background map, the underlying hidden structure of non-Euclidean spaces, such as that of cognitive spaces for example. Indeed, cognitive distances between places can be very different from physical distances. Many factors modify the cognitive representation of physical distances between places such as individual characteristics, environmental properties, social factors, etc. (Kitchin and Blades 2002). The choice of a reference place, the place of residence or a familiar place for example, is thus capable of modifying this cognitive representation. Ankomah and Crompton (1992) thus suggest that distances between places close to the reference place could be overestimated while distances between places far from the reference place would be underestimated (Fig. 2).

---

[2] Formulas of $\alpha 1$ and $\alpha 2$ are described in section 3.1
[3] Formulas of $\beta 1$ and $\beta 2$ are described in section 3.1
[4] It can be noted that it is technically possible to adjust the maps in the opposite direction (the reference map on the image map). However, adjusting the image map to the reference map allows us to know and maintain the unit of measurement of the reference surface. Furthermore, the reference being identical for several maps, it is possible to make comparisons.
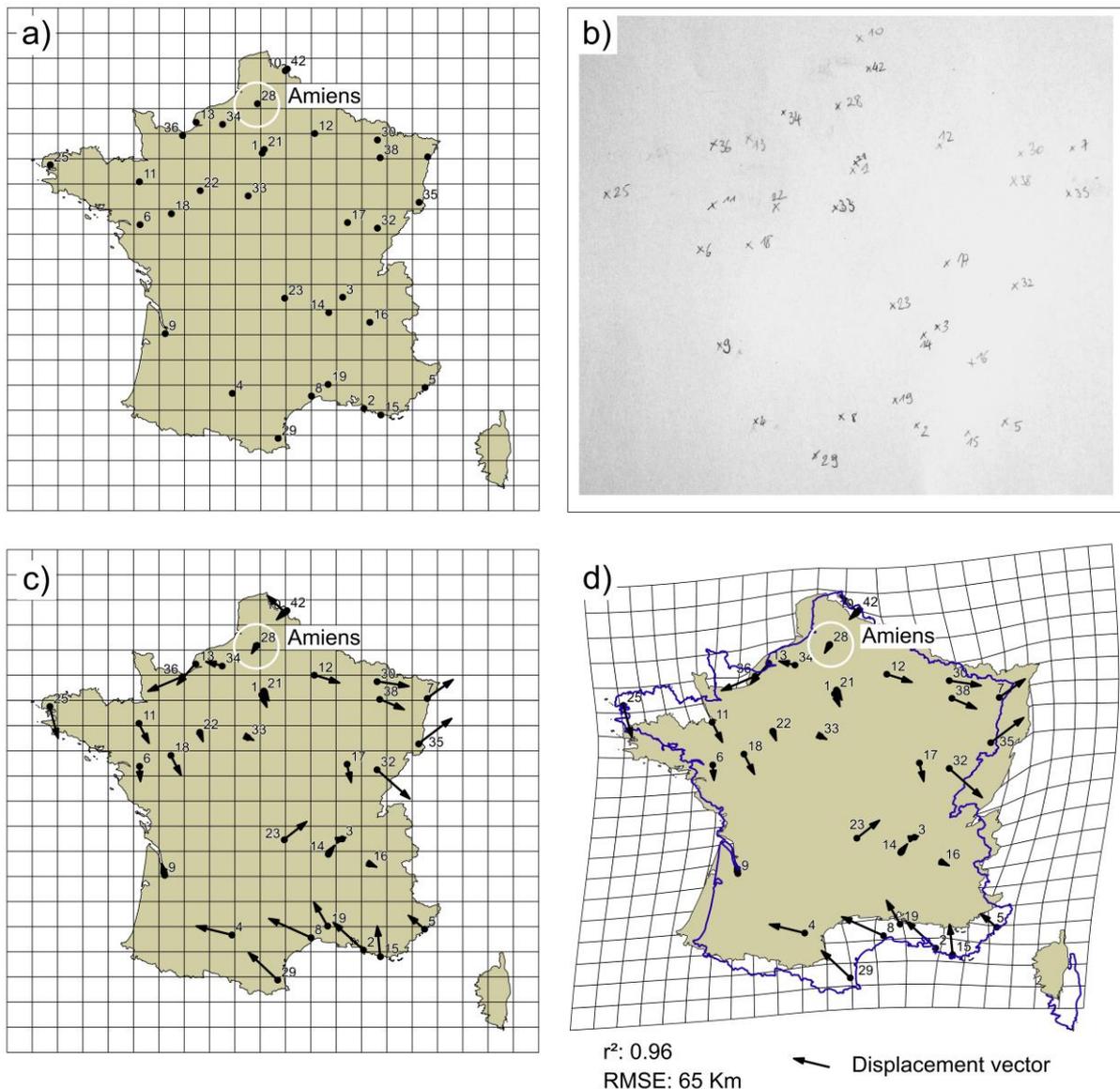
Fig. 2 An example of using ordinary BDR in spatial cognition. The cognitive representation of France by an inhabitant of the city of Amiens is highlighted by the cartographic transformation of the reference map: (a) The reference map with its 33 selected places; (b) the sketch map placing the homologous points on a blank page. (c) After the adjustment step, the displacement vectors express the position differences for each city. (d) The individual's cognitive representation of France is highlighted by the distortion of the reference map according to the displacement vectors

2.3 From BDR to GBDR

2.3.1 Limitations of the BDR aggregation strategy when comparing several maps

It can be useful to analyze and compare multiple shapes using BDR. Tobler (1994), for example, considers comparing a child's face to those of its two parents, by comparing the results of the BDR. Multiple comparisons have thus already been applied to several ancient maps to compare their precision or determine their lineage (Symington et al. 2002; Porter et al. 2019). Likewise, it seems relevant to seek to identify regularities in the "mental maps" of a set of subjects to draw general conclusions, for example in spatial planning (Cauvin et al. 1998; Roulier 2018; Llyod 1989). Maps can be compared by BDR, either from a global point of view or from a local point of view by comparing the geometric dispersion results at each landmark. This distinction leads, for example, in spatial cognition to distinguishing individual cognition from place cognition (Kitchin and Fotheringham 1998). This article focuses on the global comparison of shapes.

However, a significant limitation of BDR is that it can only compare two variables at a time, like unidimensional linear regression of which it is an evolution. It is, however, possible to compare more than two maps, either by

aggregation strategies or by disaggregation strategies (Kitchin and Fotheringham 1997). The aggregation strategy involves creating average groups and comparing the averaged results. Two aggregation strategies exist: the collective aggregation and the individual aggregation. The collective aggregation strategy involves averaging the source data for a given group, and then applying BDR. Conversely, the individual aggregation strategy consists of calculating all individual adjustments for a given group and then averaging the BDR statistical results. Kitchin and Fotheringham (1997) compared the two methods for some BDR results (r², distortion, scale, angle). The authors recommend individual aggregation for analyzing overall statistical results, because collective aggregation tends to lead to an inflated set of results for the average group. A second limitation is that collective and individual aggregation strategies do not allow individual results to be compared with each other because individual variations are removed (Kitchin and Blades 2002). In fact, identical aggregation results may hide very different individual results.

This is why the optimal strategy for comparing multiple maps using BDR is disaggregation where data are analyzed separately and not averaged but grouped together for comparison purposes only (Kitchin and Blades 2002). For example, Symington and colleagues (2002) apply BDR to 7 ancient maps of Suffolk (U.K.) and compare them all. All the 21 adjustments ((7*7-7)/2) are then statistically compared according to the Akaike's Information Criterion (A.I.C.). The results, assembled into a distance matrix, are then processed by multidimensional scaling. Each map is then identified in relation to the others according to its shape proximity in the graph. Similarly, Cauvin and colleagues (1998) calculated accuracy and consistency indices (r², RMSE, distortion index, etc.) for 96 cognitive representations of Strasbourg (France) and compared them using multiple factor analysis. The disaggregation strategy makes it possible to identify and position each individual in relation to the others based on statistical indices relating to the adjustment of the data. We propose to follow the same disaggregation approach but by applying multidimensional analysis to the geometric results and not to the statistical results of the adjustment. The disaggregation strategy involves comparing each image data set to the reference data set by BDR (as does the individual aggregation strategy), which can be tedious. The method we present proposes to establish all the comparisons in a single analysis. However, this approach requires a generalization of the BDR.

2.3.2 Procrustes methods and geometric morphometrics as inspiration for GBDR

To analyze the shape of several maps directly with BDR, we propose to generalize its algorithm based on the generalization method of the ordinary Procrustes analysis (OPA). OPA has its origins in quantitative psychology (Hurley and Cattell 1962) and aims at the same objective as BDR: transforming a bidimensional data matrix into a target matrix by reducing the residuals as much as possible, the only possible operations being translation, scaling, rotation and reflection. Although the two calculation principles are different, OPA (without reflection) and BDR have been considered equivalent to compare two matrices (Kern 2017). The OPA method performs the translation, rotation and scaling operations sequentially while the BDR adjustment calculation applies all three operations simultaneously. This is the main difference between the two methods because, like BDR, OPA is based on the least squares criterion to make this adjustment.

The OPA method has been generalized and integrated into GMM procedures in order to compare several objects. Generalized Procrustean analysis (GPA) is an iterative procedure used to eliminate non-shape-related geometric information (position, size, and orientation) between multiple configurations in a GMM shape analysis (Rohlf and Slice 1990). It produces both a multiple adjustment and a reference configuration called "consensus". Several GMM methods have been developed since the 1980s (Zelditch et al. 2004) but we can say that the most widely used method is the "Procrustean paradigm" (Adams et al. 2013) based on GPA.

A GMM analysis can use two types of GPA adjustment. The first type of adjustment, called "full GPA", translates towards the origin of the coordinate system and scales to unit centroid size only one of the configurations, the first for example. The unit centroid size corresponds to the square root of the sum of the squares of the distances between each landmark of the configuration and its centroid (Zelditch et al. 2004). The landmarks of all configurations are then translated, scaled and rotated iteratively until the best fit to the average shape (consensus) is found. The second type of adjustment, called "partial GPA", translates to the origin of the coordinate system and scales all landmarks to unit centroid size. The iterative procedure only applies rotations to the landmarks to minimize their distance from the consensus. In both cases, the alignment places each configuration as a point in a multidimensional curved space (Dryden and Mardia 2016). Both adjustment methods are very similar in their processing and results. The second is called "partial" because the iterative procedure only rotates the data. The reference is most often a consensus with the GPA. However, other references were used. For example, Zelditch

and colleagues (Zelditch et al. 1992), use the average of certain individuals as a reference. For its part, Roulier (2023) uses an external reference to the images (a topographic map) to compare several ancient maps.

3. Full and partial GBDR for comparing more than two maps

We chose to adapt BDR, the most used method of shape comparison in geography, to apply GMM to a set of several maps. The processing will produce, as with GPA, a set of adjusted points and a consensus reference. There are two kinds of GPA and therefore two kinds of GBDR fit. Here we present the detailed algorithms of full and partial GBDR adjustment taking five hypothetical triangular configurations as an example (Fig. 3). Note that the triangle is not just a simple theoretical shape useful for this demonstration. It can correspond to geographical realities that we wish to compare, for example in geomorphology (deltas) or in spatial analysis (Delaunay triangles) with the aim of testing the Central Place Theory.
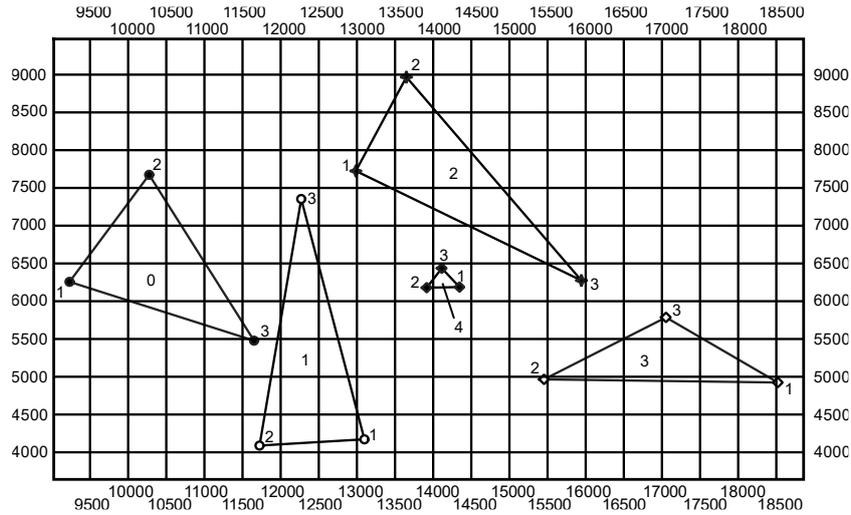
Fig. 3 The five triangles in their initial configuration space

3.1 The full GBDR procedure

For $k = 5$ configurations composed of $n = 3$ landmarks whose coordinates are (u, v), the GBDR iterative calculation procedure that we propose is as follows (Fig. 4):

3.1.1 Calculate the first reference configuration with coordinates (x, y) by translating the first configuration (configuration 0) towards the origin of the coordinate system $(0,0)$ and scaling it to unit centroid size. We first get the coordinates of the centroid of the first configuration this way:

$$\overline{u_0} = \left(\sum_{i=1}^{n} u_{0_i}\right)/n \tag{4}$$

$$\overline{v_0} = \left(\sum_{i=1}^{n} v_{0_i}\right)/n \tag{5}$$

We then calculate the centroid size of the configuration 0:

$$CS_0 = \sqrt{\sum_{i=1}^{n}\left(\left(u_{0_i} - \overline{u_0}\right)^2 + \left(v_{0_i} - \overline{v_0}\right)^2\right)} \tag{6}$$

Landmark coordinates of the first configuration are all translated to the origin of the coordinate system by subtracting those of the centroid. Scaling to unit centroid size is simply dividing the coordinates by the centroid size. The new configuration with coordinates (x, y) has a centroid size equal to the value of 1 and is centered on the origin, that is, the average of its coordinates becomes zero. For each landmark i of the configuration 0, the calculate is as follows (i varies from 1 to n):

$$x_i = (u_{0_i} - \overline{u_0})/CS_0 \tag{7}$$

$$y_i = (v_{0_i} - \overline{v_0})/CS_0 \tag{8}$$

3.1.2. Choose this first configuration as the initial reference configuration, with (x, y) coordinates. It corresponds to consensus n°0.

3.1.3. Initialize the Procrustes distance p_0 between two successive consensuses to the value of 1, before the iterative process. The Procrustes distance is a measure of the difference in shape between two objects[5]. The value of 1 corresponds to the maximum Procrustes distance in full GBDR.

3.1.4. The calculation of the regression matrix assumes the prior calculation of the parameters β_1, β_2 , $\alpha 1$, $\alpha 2$ which is established as follows for each configuration:

$$\beta_1 = \frac{\sum[(u-\overline{u})*(x-\overline{x})]+\sum[(v-\overline{v})*(y-\overline{y})]}{\sum(u-\overline{u})^2+\sum(v-\overline{v})^2} \tag{9}$$

$$\beta_2 = \frac{\sum[(u-\overline{u})*(y-\overline{y})]-\sum[(v-\overline{v})*(x-\overline{x})]}{\sum(u-\overline{u})^2+\sum(v-\overline{v})^2} \tag{10}$$

Where:
(x, y) correspond to the two-dimensional coordinates of the provisional consensus, i.e. the reference;
(u, v) correspond to the two-dimensional coordinates of each compared configuration, i.e. the images;
$\overline{x}$ is the mean of x and $\overline{y}$ is the mean of y for the provisional consensus;
$\overline{u}$ is the mean of u and $\overline{v}$ is the mean of v for each configuration.

$$\alpha 1 = \overline{x} - \beta 1 * \overline{u} + \beta 2 * \overline{v} \tag{11}$$
$$\alpha 2 = \overline{y} - \beta 2 * \overline{u} - \beta 1 * \overline{v} \tag{12}$$

All configurations are adjusted to reduce the difference between their homologous landmarks and those of the reference configuration. This quantity corresponds in this case to the full Procrustes distance. For full GBDR, the ordinary (full) BDR adjustment procedure is used (equation 1) between each configuration and the reference.

$$x' = \alpha 1 + \beta 1 u - \beta 2 v \tag{13}$$
$$y' = \alpha 2 + \beta 2 u + \beta 1 v \tag{14}$$

Where (x', y') correspond to the adjusted coordinates of each configuration with (u, v) coordinates.

3.1.5. Calculate for each homologous landmark of the adjusted configurations the average of the coordinates which produces the new provisional consensus configuration (xnc, ync).

3.1.6. As observed with GPA, there is a slight contraction of the data towards the value of zero when calculating the average. This contraction operates at each iteration. It is therefore necessary to resize the new provisional consensus to the value of 1 at each stage of the iterative process. This is done by dividing the coordinates of the new provisional consensus by its new centroid size.

---

[5] The Procrustes distance is a measure of the difference in shape between two objects because it is calculated after an adjustment that removes the differences in position, size and orientation between the two configurations of points associated with the objects. Only the geometric information on the shape is preserved, which corresponds to the definition of shape given in the introduction. At the computational level, the Procrustes distance corresponds to the square root of the sum of the squares of the differences between the positions of the landmarks in two adjusted configurations, i.e. optimally superimposed (by least squares).

3.1.7. Calculate the new Procrustes distance np between the previous consensus (with coordinates x and y) and the new consensus configuration (with coordinates xnc and ync).

$$np = \sqrt{\sum_{i=1}^{n}((x_i - xnc_i)^2 + (y_i - ync_i)^2)} \tag{15}$$

3.1.8. Perform the consensus stability test

When the adjustment no longer changes the difference in shape between the previous consensus configuration and the new consensus configuration, the latter is retained as the configuration minimizing the differences between all the configurations, that is to say as the definitive consensus, which ends the iterative process. This convergence is generally quite rapid. Rohlf and Slice (1990) suggest a tolerance of 10^{-3} or 10^{-4} for GPA. We use this threshold for GBDR. The stopping rule is as follows in our method:

If $p_0 - np > 0.0001$ start again at step 3.1.4. using coordinates (x', y') as coordinates (u, v), taking as reference coordinates (x, y), those of the new provisional consensus (xnc, ync) and assigning to p_0 the value of np ; otherwise go to the next step.

3.1.9. End of the iterative procedure. The new consensus has become the definitive consensus and all data are adjusted one last time by a complete BDR (Fig. 5).
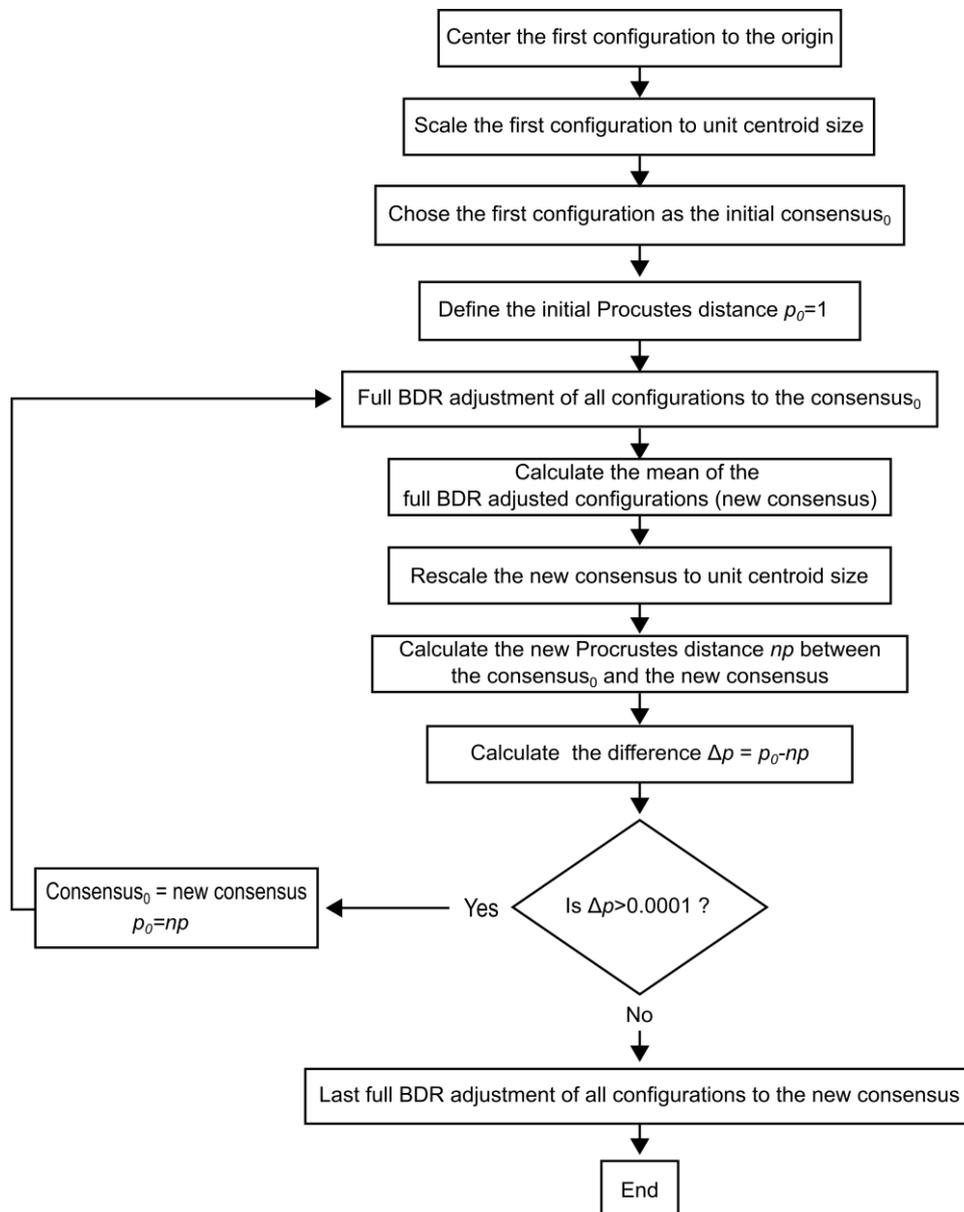
Fig. 4 The full GBDR algorithm
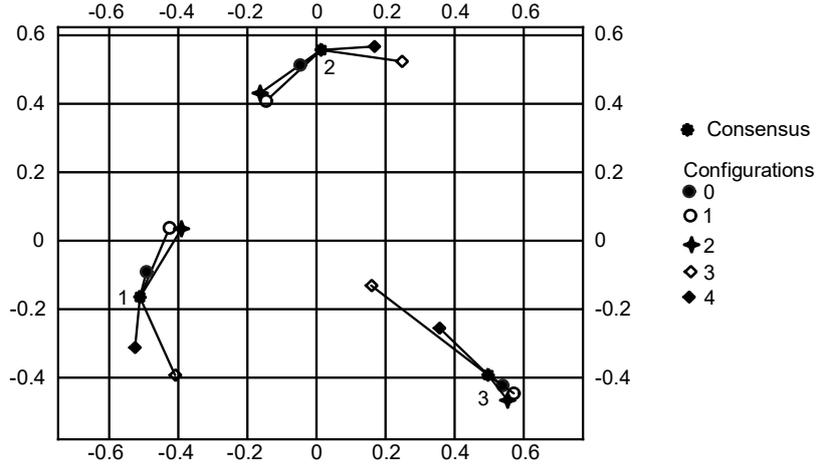
Fig. 5 The consensus and its adjusted data (full GBDR)

3.2. The partial GBDR procedure

3.2.1. Landmarks of all configurations (not only the first) are translated to the origin of the coordinate system and scaled to unit centroid size, following an iterative procedure close to the previous one.

3.2.2. Arbitrarily choose one of the centered and scaled configurations, the first for example, with (x, y) coordinates. This configuration becomes the initial reference configuration, that is to say the consensus n°0.

3.2.3. As with full GBDR, initialize the Procrustes distance p_0 to the maximum value. With partial adjustment, the maximum value reaches the value of 1.41 (square root of 2).

3.2.4. Configurations are then all adjusted so as to reduce the difference between their homologous landmarks and those of the reference configuration using an iterative process.

For partial GBDR, this adjustment simply corresponds to a rotation based on the ordinary BDR formula modified 1) so as not to translate the landmarks and 2) so as not to resize the data by dividing it by the BDR scale parameter (square root of the sum of the squared parameters β_1 and β_2). The parameters β_1 and β_2 are calculated in the same way as for the full GBDR.

The adjusted coordinates (x', y') on the reference configuration are then calculated as follows for each configuration:

$$x' = (\beta 1 u - \beta 2 v) / \sqrt{\beta_1{}^2 + \beta_2{}^2} \tag{16}$$

$$y' = (\beta 2 u + \beta 1 v) / \sqrt{\beta_1{}^2 + \beta_2{}^2} \tag{17}$$

3.2.5. As with full GBDR, we calculate for each homologous landmark the average of the adjusted coordinates which produces the new provisional consensus configuration (xnc, ync).

3.2.6. As with full GBDR, we resize to the value of 1, the size of the new provisional consensus.

3.2.7. We calculate the new Procrustes distance np between the previous consensus configuration (with x and y coordinates) and the new provisional consensus configuration (with xnc and ync coordinates) in the same way as for full GBDR.

3.2.8. As with full GBDR, we perform the consensus stability test. The stopping rule is the same as for full GBDR: If $p_0 - np > 0.0001$ start again at step 3.2.4 taking as reference coordinates, those of the new provisional consensus and assigning to p_0 the value of np; otherwise go to the next step.

3.2.9. End of the iterative procedure. The new consensus has become the definitive consensus and all data are adjusted one last time by a partial ordinary BDR (Fig. 6).
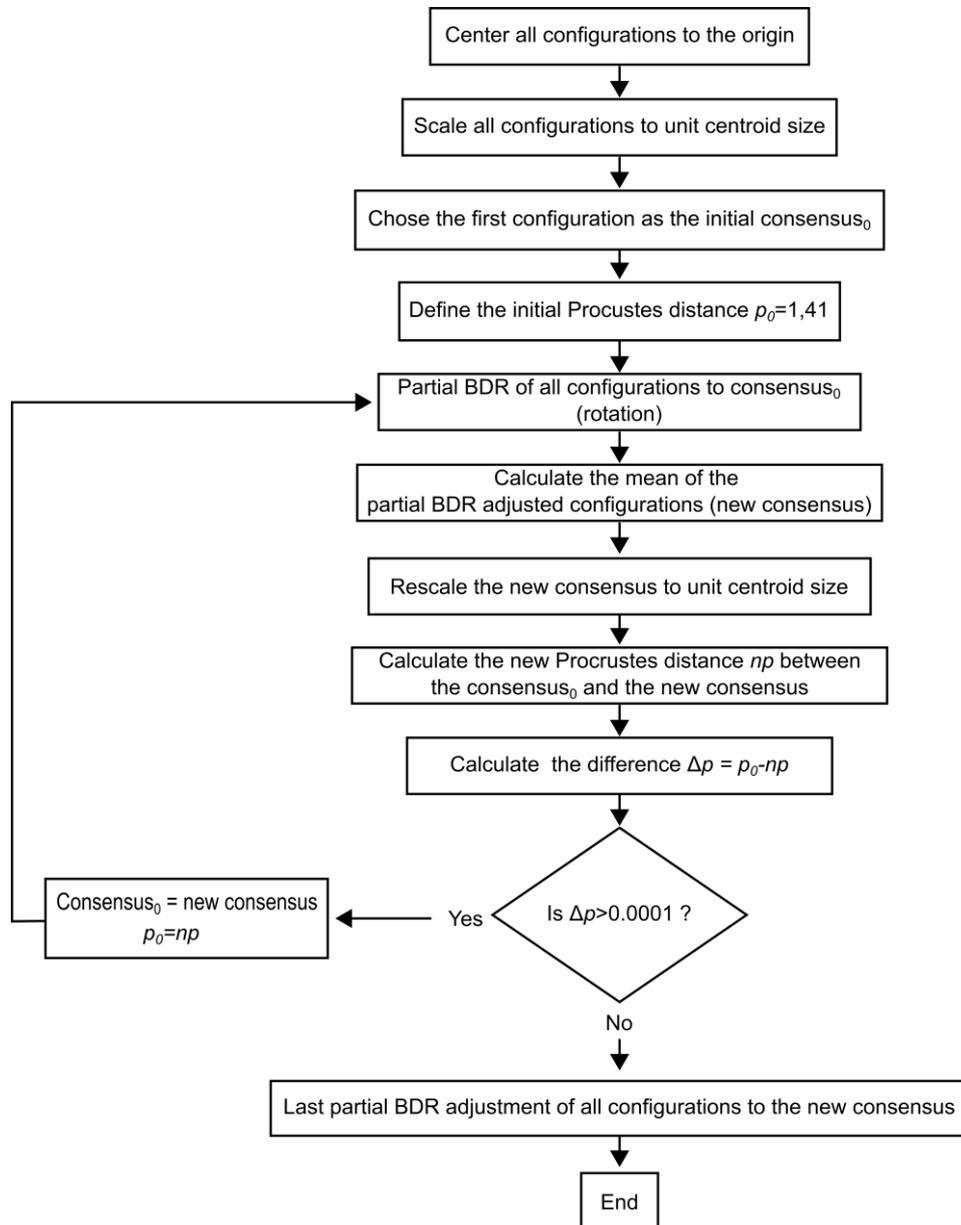
Fig. 6 The partial GBDR algorithm

We then notice that the main difference between the two algorithms lies in the management of the size of the configurations during the iterative process. Full BDR adjusts the configurations as close as possible to the consensus at each iteration while partial BDR maintains a constant size of one for all configurations adjusted to the consensus.

3.3 Comparison with GPA results

Both GBDR adjustment methods were implemented in the MapMorphy (https://mapmorphy.fr/) software. The calculation of the triangular consensus by GBDR could thus be compared to that of software carrying out GMM by GPA. The consensus result was chosen as the point of comparison because it is the outcome of the adjustment process. The software chosen for the comparison is tpsRelw (Rohlf 2015), whose first calculation step is a GPA.

TpsRelw was chosen because it is one of the few to offer both types of adjustment. The original coordinates of the 5 triangles were therefore recorded in tps format to be read by tpsRelw. The options chosen in tpsRelw are: "projection = none" on the one hand, and "scale aligned = Cos(rho)" for full adjustment and "scale aligned = 1" for partial adjustment, on the other hand. The two comparisons of consensus results were carried out by ordinary bidimensional regression using Darcy 2.1 software which provides r² results for full and partial adjustment (Table 1). Results show that the consensus calculated by GBDR and GPA are almost identical since the r² is at the value of 1 in both cases.

	Full adjustment				Partial adjustment			
r²	1,00000				1,00000			
	x		Y		x		Y	
	GBDR	GPA	GBDR	GPA	GBDR	GPA	GBDR	GPA
Point 1	-0,510982	-0,51098	-0,164782	-0,16475	-0,512323	-0,51234	-0,173184	-0,17318
Point 2	0,013815	0,01379	0,557155	0,55714	0,0211037	0,02111	0,5613523	0,56134
Point 3	0,4971671	0,49719	-0,392373	-0,39239	0,4912193	0,49123	-0,388168	-0,38816

Table 1 Comparisons of consensus results between GBDR (MapMorphy software) and GPA (tpsRelw software)

4. Using GBDR in a GMM analysis

GBDR is just an adjustment procedure. It must be integrated into an overall data processing methodology to reveal its usefulness. We chose GMM for this. Fitting is a crucial step in GMM but it must indeed be followed by other data processing for a complete geometric morphometric analysis. Three additional steps can show the potential of GBDR in the context of a GMM mapping application (Roulier 2023). The first is a PCA applied to the residuals of the adjustments. The result produces a morphospace in which each map is represented by a point. The distances between maps in the morphospace correspond to the Procrustes distance, that is to say the differences in shape (Fig. 7). The second is a clustering based on K-means and intended to identify a typology of cartographic shapes in the morphospace. The third is the production of models from the PCA results, using an interpolation function. The deformed grids of the figure 8 correspond to the models located at the intersection of the rectangular area of the PCA with its axes. Note that a model can be calculated at any point of the morphospace.
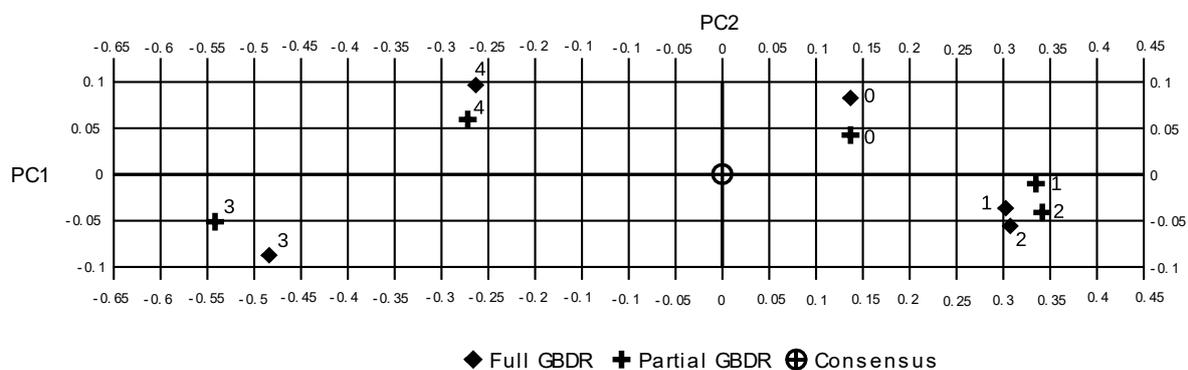
Fig. 7 The morphospace of the five triangles

Figure 7 presents the first two components of our example and shows that most of the shape differences are represented by the first, on the x-axis. PC1 indeed explains 99.39% of the total variance for the full GBDR and 99.43% of the total variance with the partial GBDR adjustment. Triangle n°1 and triangle n°2 (with an acute angle at landmarks n°3) are opposed to triangle n°3 (with an open angle at landmarks n°3). Triangle n°0 and triangle n°4 constitute intermediate variants related to one of these two extremes cases. This graph confirms what can be deduced intuitively from the observation of the initial configurations (Fig. 3).
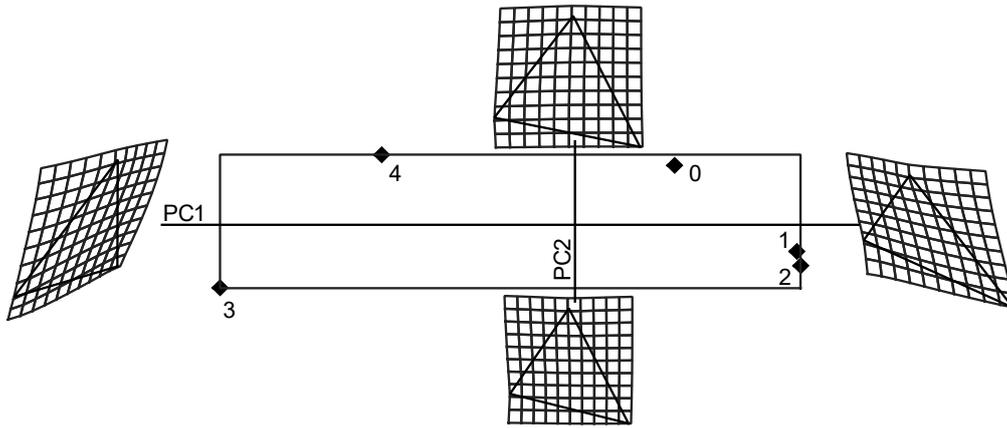
Fig. 8 Interpolated grids for 4 models (full GBDR adjustment)

5. A cartographic example of GBDR application

5.1 Individual spatial cognition analysis trough GBDR and GMM

The example chosen to show the usefulness of GBDR concerns a morpho-geometric analysis in spatial cognition, that is to say our internal knowledge of the structure of space. This knowledge of space is the result of a process called cognitive mapping. This is not the primary subject of this methodological article. We will therefore simply retain the definition given by Downs and Stea (2011): "Cognitive mapping is a process composed of a series of psychological transformations by which an individual acquires, codes, stores, recalls, and decodes information about the relative locations and attributes of phenomena in his everyday spatial environment". This process leads to the construction in each of us of a particular cognitive representation of space anchored in a personal cognitive space. Cognitive representations of space are guides for decision-making in spatial behavior (Kitchin and Blades 2002), which justifies their analysis.

Ordinary BDR is a proven method for analyzing the structure of cognitive representations of space by comparing the shape of two maps in geography (Cauvin et al. 1998; Cauvin 2002; Roulier 2018; Cauvin 1984; Roulier et al. 2021) or in psychology (Friedman and Kohler 2003; Kitchin 1993; Kitchin and Blades 2002; Giraudo and Pailhous 1994). If we indeed consider that a cognitive representation of space is two-dimensional in nature (Kosslyn and Pomerantz 1977), it can be considered as a map. To externalize this map, it is possible to ask the person to draw a sketch that we call "cognitive configuration". The structure of this cognitive configuration can then be, through differences in shape, analyzed by ordinary BDR using comparison of homologous landmarks (Fig. 2). We propose to generalize this approach to several maps using the GBDR method and geometric morphometrics with the objective of identifying a typology of cognitive spaces.

The objective of this example is then above all to highlight the variety and proximity of thought patterns in spatial cognition thanks to the integration of GBDR in a GMM analysis. Ideally, GMM should be used when shape variations are small (Rohlf 1999). However, cognitive representations of space are specific to each individual because they depend in particular on the personal experience of space. We can therefore observe more or less significant differences in intensity, orientation and coherence, between the cognitive representations of individuals on the one hand, and with a reference map (topographic map for example) on the other hand. Despite this, the GMM will attempt to find regularities and comparable cognitive representations of space based on the notion of shape. We propose a method that can be sequenced in two steps: 1. With GBDR, identification of similar cognitive representations of space, taking a consensus as reference. Choosing consensus as a reference reduces variations in shape, and 2. Calculation of average of similar cognitive configuration by data aggregation of each type and, using the ordinary BDR, comparison of the structure of these average configurations with that of a topographic map as a reference. Choosing in this second phase a topographic map as a reference for each cluster will make it possible to compare similar cognitive representations with geographical reality and to expose recognizable archetypes of cognitive spaces.

5.2 The small cognitive data set

Map sketching is a classic method of externalizing cognitive representations of space (Kitchin and Blades 2002). The data source comes from a survey carried out among 40 individuals who reported their cognitive representation of their territory of life, a window of 5 km by 5 km (Fig. 9) in the surroundings of the village of Froidestrées (Picardie, France). Respondents were asked to locate, using a point symbol and without outside assistance, a proposed list of 30 places on a blank page. Respondents could use boundaries that they had drawn from memory to position the requested locations. The morphometric comparison of data requires a homogeneous corpus. The dataset must therefore be composed of the same landmarks, which is why 7 cognitive configurations out of 40 and only 11 places out of 30 were retained (in addition, two outliers were removed from the sample). Each landmark is coded in a unique file with its X and Y coordinates, its identifier and that of the cognitive configuration. The sample may seem limited in size but it is not intended to generalize the results and is only of interest here as a methodological test. This simple small sample will allow us to demonstrate the effectiveness of our method by detailing and verifying the individual results (Appendix 1).
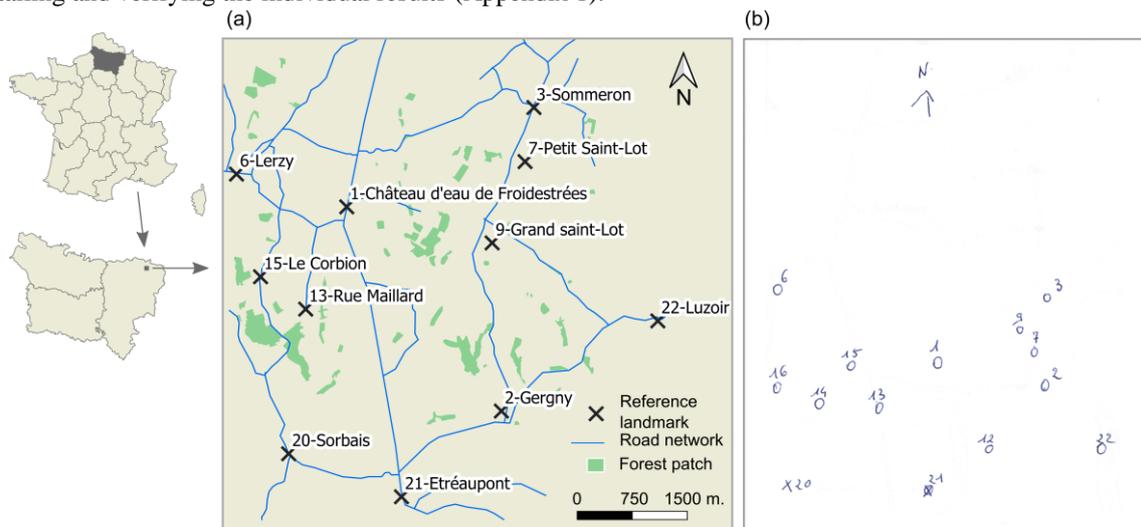
Fig. 9 The study area: (a) The reference topographic configuration composed of 11 selected places (b) An example of cognitive configuration collected on a drawing sketch map (individual n°34)

5.3 Step one: clustering of cognitive configurations adjusted by GBDR

The first step of the analysis compares the different cognitive configurations with each other in order to partition the data. In principle, cognitive configurations should be processed either with full GBDR or with partial GBDR. We apply them both for this demonstration. This step produces a new configuration: the consensus cognitive configuration (CCC) which is calculated as an average such that its Procrustes distance from each cognitive configuration is minimal. This CCC (i.e. the mean shape configuration) can be considered as an internal reference to all adjusted cognitive configurations (ACC). A PCA is then applied to all adjusted data, previously projected onto a tangent plane. To do this, the CCC is subtracted from each ACC to obtain residuals. This will center the axes of the PCA on the CCC. A covariance matrix is then calculated from the residuals. The PCA is applied to this matrix to obtain decreasing eigenvalues and eigenvectors.

Finally, the morphospace is constructed from the PCA scores (Fig. 11 and Fig. 12). The latter are obtained by multiplying the residuals of each ACC by the two eigenvectors chosen for the PCA graph. Each ACC is thus represented by a point in the morphospace and each axis represents a spectrum of cognitive configurations. Distances in morphospace reflect Procrustes distances, i.e. differences in shapes. The morphospace therefore highlights, through proximities, the resemblance with the CCC and the resemblance of the ACC to each other. The closer individuals are in the morphospace, the closer the shape of cognitive configurations, and vice versa. For our example, we will use the first two components which represent 61% of variance with the complete GBDR and 63% of variance with the partial GBDR.

We then create a typology using a k-means clustering. We can note in this regard that the two morphospaces are very similar in this example and that the adjustment by full GBDR and partial GBDR results in the same clusters.

The displacement vectors present a significant diversity both in intensity and in orientation (Fig. 10).
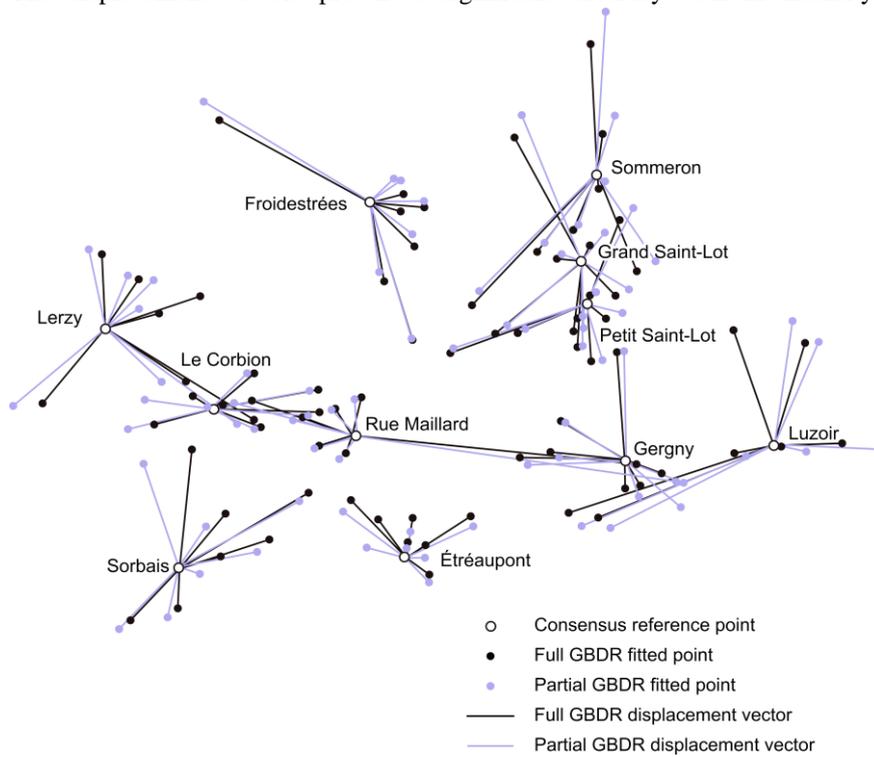
Fig. 10 GBDR adjusted places on the consensus

However, morphospaces manage to highlight oppositions and similarities compared to CCC. The morphospaces show that the three cognitive configurations of Cluster 1 are rather close to the consensus, that is to say the average. The main trend of PC1 (44 % of the total variance with full GBDR and 45% with partial GBDR) opposes a combination of deviations from the consensus, both lateral and vertical, to rather opposite deviations from the consensus on these two directions. PC1 thus opposes Cluster 3 to singleton Cluster 2. Furthermore, the second trend of PC2 (17 % of the total variance with full GBDR and 18 % with partial GBDR) opposes a strong extension of the central zone from bottom to top, with an opposite, but more moderate force. The singleton Cluster 4 falls under this second trend.
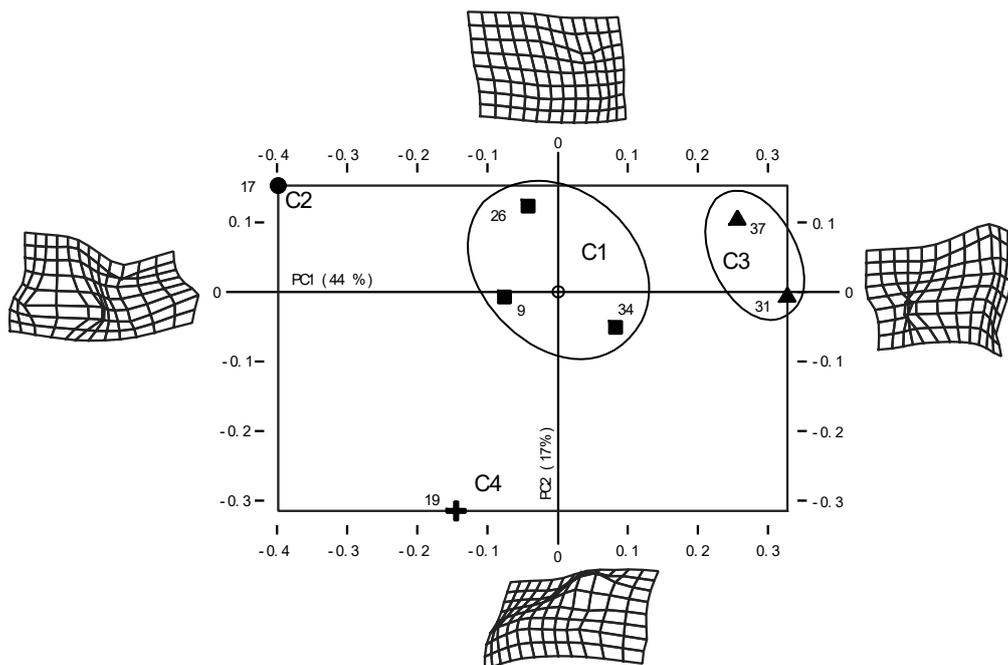
Fig. 11 Morphospace and clusters from the full GBDR adjustment
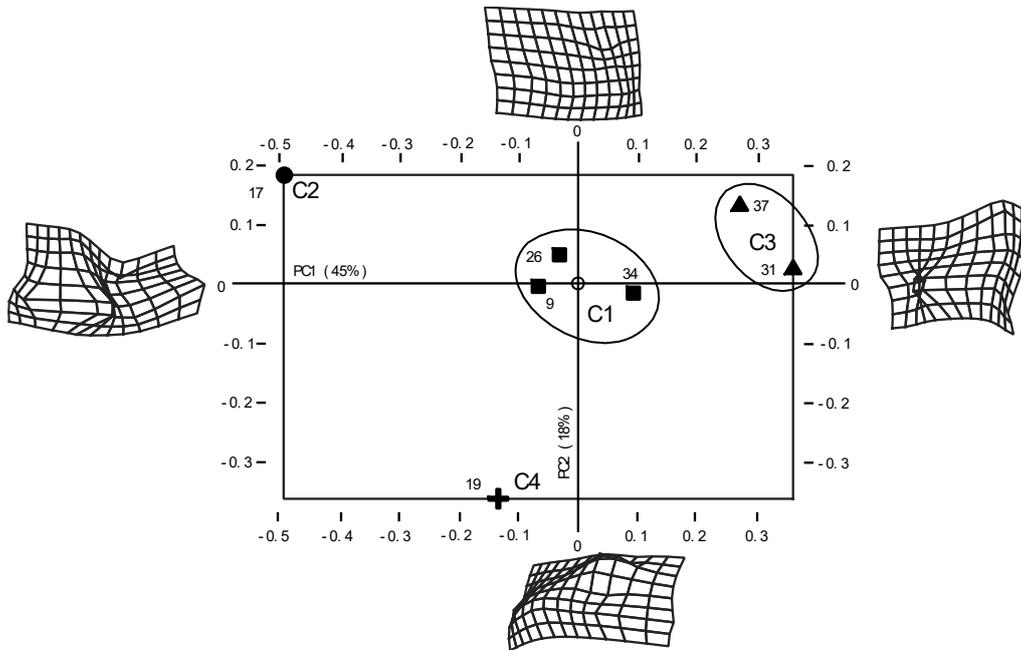
Fig.12 Morphospace and clusters from the partial GBDR adjustment

5.4 Step two: towards a typology of cognitive spaces

The second step of the analysis compares the related cognitive configurations with an external geographical reference map (Fig. 9a), using ordinary BDR. This can lead to geographical conclusions (including in terms of distance error) while the first step only leads to statistical conclusions. Let's keep in mind that this reference configuration is fundamentally the implicit reference to the origin of all cognitive configurations drawn. We can now talk about the visualization of cognitive spaces, understood as the deformation of objective topographical space. Illustrative geographic data can thus be added to better reflect the cognitive space. They must be expressed in the same coordinate system as the reference landmarks. These illustrations will be transformed by interpolation based on the displacement vectors. We choose to add a square mesh grid representing the continuous topographical space and the local road network which clearly reflect the structure of this space. This one to one comparison can be carried out for each cognitive configuration (Appendix 1). It is thus possible to carry out, for example, this comparison with the consensus. The resulting cognitive space can be considered as the average of our corpus, to which all individual cognitive spaces are compared (Fig. 13). The RMSE reveals an average difference of 844 meters with the topographic map. Some villages are positioned at too high latitude (Sorbais, Étréaupont, Gergny), which curves the cognitive space in the south of the sector. Conversely, the villages of Sommeron, Petit Saint-Lot and Lerzy located to the north of the sector have a location which underestimates the latitude. The result is a mean cognitive space "crushed" in the north-south direction.
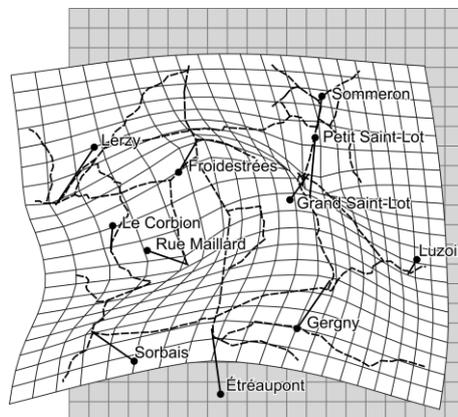
Fig. 13 The consensus cognitive space

But the comparison with an external reference by ordinary BDR can also be done for each cluster as a whole. We therefore propose to establish four average cognitive configurations from the clustering step in order to produce a typology of cognitive spaces (Fig. 14). The four types of cognitive spaces correspond to four ways of thinking about space for the sample and therefore to four ways of behaving there. The average coordinates of the configurations of a cluster then correspond to the means of the homologous landmarks of this cluster. The objective here is not to provide explanations on the constitution of groups, but just to show a method for identifying these groups. Additional data on the profile of those interviewed would be necessary to obtain elements of explanation. Cluster 1 illustrates the majority trend and groups together three cognitive configurations, quite close to the consensus, i.e. the average. Like the consensus, this cluster is characterized by a moderate underestimation of the north and south distances ($r^2 = 0.88$; RMSE = 814 meters). Then, Clusters 2 and 3 oppose each other on the first PC axis. Cluster 3 is characterized by erroneous positioning of certain villages towards the west (villages of Lusoir and Lerzy), while singleton cluster 2 is marked, conversely, by a strongly exaggerated positioning towards the east of the village of Rue Maillard. This causes a significant east/west distortion for Cluster 2 ($r^2 = 0.54$; RMSE = 1630 meters) because this error is accompanied by errors in the opposite direction, towards the west (villages of Sommeron, Petit-Saint- Lo, Grand-Saint-Lo, Sorbais, Etréaupont). The distortions for singleton cluster 4 are of a different order. The clear overestimation towards the north of certain positions located in the south of the sector (villages of Sorbais, Étréaupont, Gergny) and especially in the center of the sector (Grand-Saint-Lo) is combined with an error in the position towards the south of certain villages located to the north of the sector (Lerzy, Sommeron). This results in a significant north/south distortion of cognitive space ($r^2 = 0.70$; RMSE = 1313 meters) which moves it, like Cluster 2, away from the consensus in morphospace.

The significant variety of cognitive representations of space makes it difficult to create a typology based solely on drawn sketches. The GMM analysis by GBDR, however, makes it possible to identify the main trends among the group interviewed and to identify opposite and similar cases despite this diversity of cognitive representations. Additional analyzes could be carried out by reproducing the method on subgroups (on larger samples), and therefore by refining the results or creating groups based on more components for example with an Ascending Hierarchical Classification (AHC). Above all, it is possible to cross-reference the scores of individuals, considered as shape variables, with data on the location of these individuals, qualitative or quantitative attributes qualifying the different people who drew the sketches.
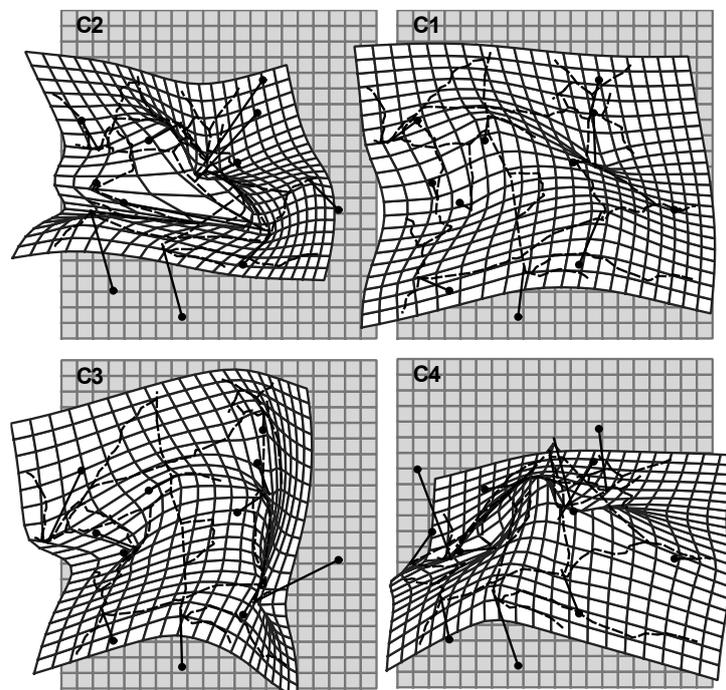
Fig. 14 The four types of cognitive space

6. Conclusion

We hope to have demonstrated in this paper that GBDR constitutes a credible alternative to GPA, in particular for spatial comparisons. Using GBDR in GMM avoids grouping data for comparison purposes. It proposes a strategy

for analyzing disaggregated data based on geometric results and not only on the statistical results of the BDR. This allows access to each individual result, developing geometry-based modeling, creating a morphospace, and producing models for any point in that space. Our demonstration concerns a geographical application of GBDR but this method could be used on classic applications of geometric morphometrics in natural sciences or archeology for example, or even on completely different areas of research such as FCP (free choice profiling). Other applications than spatial cognition are possible in cartography such as the comparison of ancient maps, spatio-temporal maps or spatial structures.

In this article, we have discussed two types of GBDR adjustment, partial GBDR and full GBDR. It is difficult to say which method to use in cartography, in the absence of comparative tests. On the one hand, full GBDR will tend to limit the weight of configurations presenting significant variations in the consensus because the method adjusts as closely as possible the size of each configuration to the reference at each step of the iterative process (Klingenberg 2020). This method could therefore be interesting for certain applications in geography such as the mapping of cognitive spaces, taken as an example in this article. However, it is advisable to project the adjusted data onto a tangent linear space before analyzing them by PCA. In this case, partial GBDR presents an important advantage because the distances between configurations in tangent space and PCA are more faithful to those of the distances on the initial curved shape space (Zelditch et al. 2004).

The GBDR algorithms were implemented in the MapMorphy software and used for the different examples presented in this article. The software includes, using a graphical interface, all stages of cognitive configuration processing: data loading, adjustment by GBDR, interpolation, PCA and clustering. It allows the user to query and represent data encoded in shapefile format, which facilitates their analysis in geographic information systems.

Declarations

Competing Interests:
Financial interests:
The author declares he has no financial interests.
Non-financial interests: none.

Other declarations:
Ethics statement: The data used for this research does not contain any personal information.
Informed consent: Informed consent has been obtained from all participants of this study.
Data: The datasets analyzed during the current study are available from the corresponding author on reasonable request.

References

Adams DC, Rohlf FJ, Slice DE (2013) A field comes of age: Geometric morphometrics in the 21st century. Hystrix, the Italian Journal of Mammalogy, 24:7-14. https://doi.org/10.4404/hystrix-24.1-6283

Ahmed N, Miller HJ (2007) Time–space transformations of geographic space for exploring, analyzing and visualizing transportation systems. Journal of Transport Geography 15:2-17. https://doi.org/10.1016/j.jtrangeo.2005.11.004

Ankomah PK, Crompton JL (1992) Tourism cognitive distance. Annals of Tourism Research 19:323.

Bookstein FL (1986) Size and Shape Spaces for Landmark Data in Two Dimensions. Statistical Science 1:181-222. https://doi.org/10.1214/ss/1177013696

Bookstein FL (1992). Morphometric Tools for Landmark Data: Geometry and Biology. Cambridge University Press. https://doi.org/10.1017/CBO9780511573064

Cauvin C (1984) A method for comparing localized data: Bidimensional regression. L'Espace géographique 13: 94-109. https://doi.org/10.3406/spgeo.1984.3909

Cauvin C (1997) About mapping transformations. Cybergeo: European Journal of Geography. https://cybergeo.revues.org/5385

Cauvin C (2002) Cognitive and cartographic representations: Towards a comprehensive approach. Cybergeo: European Journal of Geography. https://doi.org/10.4000/cybergeo.194

Cauvin C (2005) A systemic approach to transport accessibility. A methodology developed in Strasbourg: 1982-2002. Cybergeo: European Journal of Geography. https://doi.org/10.4000/cybergeo.3425

Cauvin C, Chernai Z, Daniilidis K (1998) Users and cognitive representations of the city: Examples in Strasbourg. In: The geographical space of cities, for multi-strata synergy. Anthropos, Paris, pp 300-346

Downs RM, Stea D (2011) Cognitive Maps and Spatial Behaviour: Process and Products. In: Dodge M, Perkins C, Kitchin R (ed) The Map Reader. John Wiley & Sons, West Sussex England, pp 312–317 https://doi.org/10.1002/9780470979587.ch41

Dusek T (2011) Comparison of air, road, time and cost distances in Hungary. In ERSA conference papers (ersa10p1025; ERSA Conference Papers). European Regional Science Association. https://ideas.repec.org/p/wiw/wiwrsa/ersa10p1025.html

Dusek T (2012) Bidimensional Regression in Spatial Analysis. Regional Statistics 1: 61-73.

Dryden IL, Mardia KV (2016) Statistical Shape Analysis: With Applications in R. John Wiley & Sons, Chichester

Friedman A, Kohler B (2003) Bidimensional Regression: Assessing the Configural Similarity and Accuracy of Cognitive Maps and Other Two-Dimensional Data Sets. Psychological Methods 8: 468-491. https://doi.org/10.1037/1082-989X.8.4.468

Giraudo M-D, Pailhous, J (1994) Distortions and fluctuations in topographic memory. Memory & Cognition 22: 14-26. https://doi.org/10.3758/BF03202757

Hardy RL (1971) Multiquadric equations of topography and other irregular surfaces. Journal of Geophysical Research (1896-1977) 76: 1905-1915. https://doi.org/10.1029/JB076i008p01905

Helbich M, Amelunxen C, Neis P, Zipf A (2012) Comparative spatial analysis of positional accuracy of OpenStreetMap and proprietary geodata. Proceedings of GI_Forum, 4, 24.

Hurley JR, Cattell RB (1962) The procrustes program: Producing direct rotation to test a hypothesized factor structure. Behavioral Science 7: 258-262. https://doi.org/10.1002/bs.3830070216

Jenny B, Hurni L (2011) Studying cartographic heritage: Analysis and visualization of geometric distortions. Computers & Graphics 35: 402-411. https://doi.org/10.1016/j.cag.2011.01.005

Kendall DG (1977) The diffusion of shape. Advances in Applied Probability 9: 428-430.

Kern JL (2017) On the Correspondence Between Procrustes Analysis and Bidimensional Regression. Journal of Classification 34: 35-48. https://doi.org/10.1007/s00357-017-9224-z

Kitchin R (1993) Using bidimensional regression to analyze cognitive maps. Swansea Geographer 30: 33-50.

Kitchin R, Blades M (2002) The Cognition of Geographic Space. IB Tauris, london

Kitchin R, Fotheringham AS (1997) Aggregation Issues in Cognitive Mapping. The Professional Geographer 49: 269-280. https://doi.org/10.1111/0033-0124.00076

Kitchin R, Fotheringham A (1998) The Effects of Spatial and Locational Cueing on the Analysis of Aggregate Cognitive Mapping Data. Environment and Planning A 30: 2245-2253. https://doi.org/10.1068/a302245

Klingenberg CP (2020) Walking on Kendall's Shape Space: Understanding Shape Spaces and Their Coordinate Systems. Evolutionary Biology 47: 334-352. https://doi.org/10.1007/s11692-020-09513-x

Kosslyn SM, Pomerantz JR (1977) Imagery, propositions, and the form of internal representations. Cognitive Psychology 9: 52-76. https://doi.org/10.1016/0010-0285(77)90004-4

Lloyd R (1989) Cognitive Maps: Encoding and Decoding Information. Annals of the Association of American Geographers 79: 101-124. https://doi.org/10.1111/j.1467-8306.1989.tb00253.x

Lloyd CD, Lilley KD (2009) Cartographic veracity in medieval mapping: Analyzing geographical variation in the Gough Map of Great Britain. Annals of the Association of American Geographers 99: 27-48.

Mitteroecker P, Schaefer K (2022) Thirty years of geometric morphometrics: Achievements, challenges, and the ongoing quest for biological meaningfulness. American Journal of Biological Anthropology 178: 181-210. https://doi.org/10.1002/ajpa.24531

Porter C, Lilley K, Lloyd C, Mcdermott S, Milligan R (2019) Cartographic Connections-The digital analysis and curation of sixteenth-century maps of Great Britain and Ireland. e-Perimetron 14: 97-109. http://www.e-perimetron.org/Vol_14_2/Porter_et_al.pdf

Rohlf FJ (1999) Shape Statistics: Procrustes Superimpositions and Tangent Spaces. Journal of Classification 16: 197-223. https://doi.org/10.1007/s003579900054

Rohlf F (2015) The Tps series of software. Hystrix 26: 1–4. https://doi.org/10.4404/hystrix-26.1-11264

Rohlf FJ, Slice D (1990) Extensions of the Procrustes Method for the Optimal Superimposition of Landmarks. Systematic Zoology 39: 40-59. https://doi.org/10.2307/2992207

Roulier F (2018) One way of adapting bidimensional regression to polygons: an example with spatial cognition. Cybergeo : European journal of geography. https://doi.org/10.4000/cybergeo.29082

Roulier F (2023) Geometric Morphometrics Applied to Cartography. Revue Internationale de Géomatique 32 : 17-37. https://doi.org/10.32604/RIG.2023.045458

Roulier F, Gallet-Moron E, Decocq G (2021). Cognitive Mapping of Forest Fragments. In: F. Pedrotti, E. O. Box (Eds.) Tools for Landscape-Scale Geobotany and Conservation. Springer International Publishing, pp 177-183. https://doi.org/10.1007/978-3-030-74950-7_9

Symington A, Charlton ME, Brunsdon CF (2002) Using bidimensional regression to explore map lineage. Computers, Environment and Urban Systems 26: 201-218.

Tobler WR (1965) Computation of the correspondence of geographical patterns. Papers of the Regional Science Association 15: 131-139. https://doi.org/10.1007/BF01947869

Tobler WR (1978a) Comparing figures by regression. ACM SIGGRAPH Computer Graphics 12: 193-195. https://doi.org/10.1145/965139.807389

Tobler WR (1978b). Comparison of Plane Forms. Geographical Analysis 10: 154-162. https://doi.org/10.1111/j.1538-4632.1978.tb00004.x

Tobler WR (1994). Bidimensional Regression. Geographical Analysis, 26: 187-212.

Yabe N (2024) Mapmaking Process Reading from Local Distortions in Historical Maps: A Geographically Weighted Bidimensional Regression Analysis of a Japanese Castle Map. ISPRS International Journal of Geo-Information 13, 124. https://doi.org/10.3390/ijgi13040124

Zelditch M, Bookstein FL, Lundrigan BL (1992) Ontogeny Of Integrated Skull Growth In The Cotton Rat Sigmodon Fulviventer. http://deepblue.lib.umich.edu/handle/2027.42/137530

Zelditch ML, Swiderski DL, Sheets HD, Fink WL (2004) Geometric morphometrics for biologists: A primer. Elsevier Academic Press, San Diego.

Appendix 1
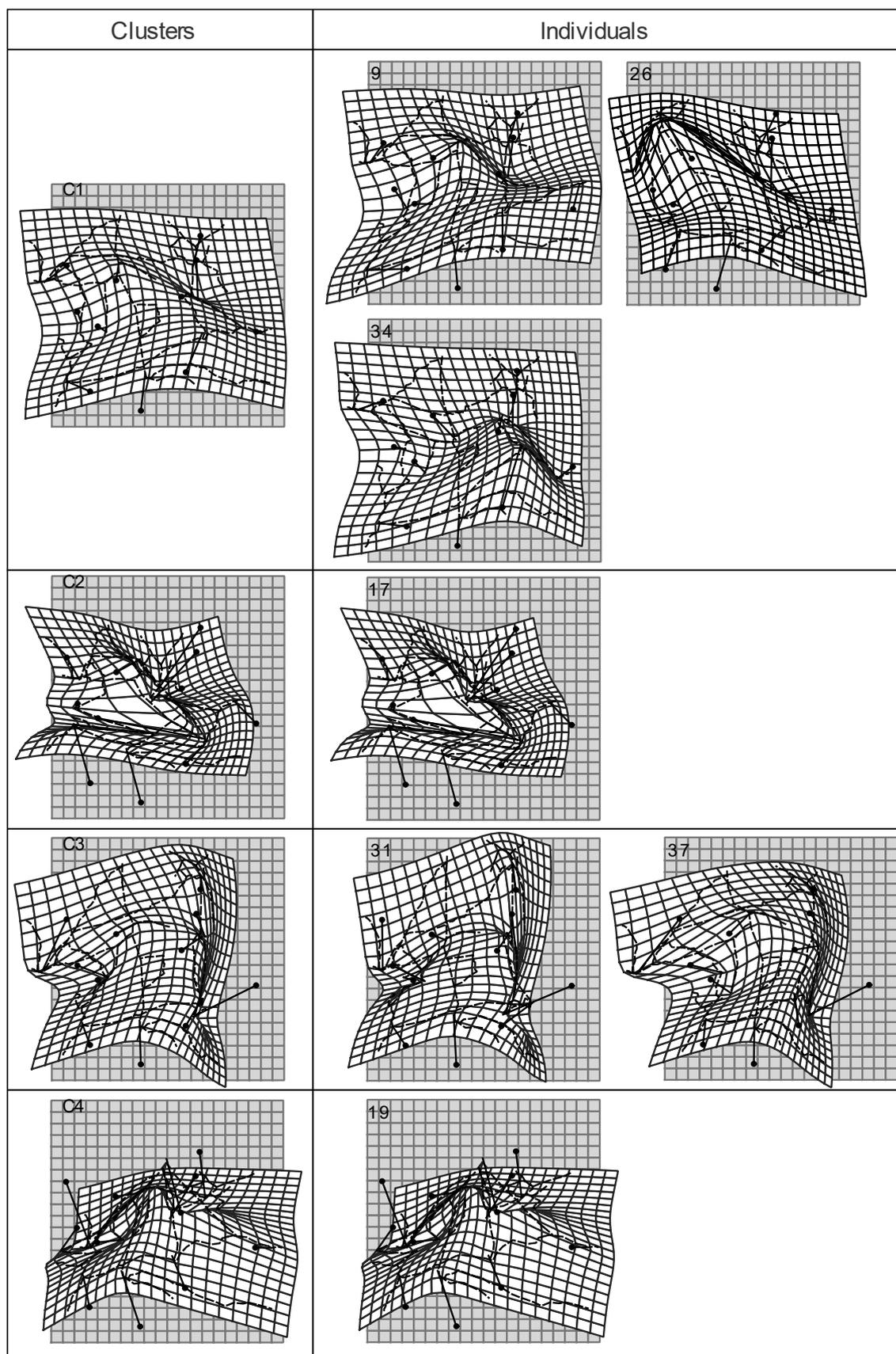
Fig. 15 The cognitive spaces by clusters