The Journal of Juristic Papyrology
vol. XXXVI (2006), pp. 67–90

Piotr Hoffman
Bartek Klin

CAREFUL WITH THAT COMPUTER

ON CREATING MAPS BY MULTIDIMENSIONAL SCALING OF PAPYRI IN KATJA MUELLER'S RECENT STUDIES ON THE TOPOGRAPHY OF THE FAYUM

Katja Mueller[1] has recently advocated a mathematical, computer-aided approach to locating settlements in Ptolemaic Egypt.[2] Based on that approach, as well as on more traditional analysis, she reconstructed probable settlement maps for the Merides of Themistos and Polemon of the Arsinoite nome. In this paper, we review the validity of the proposed computer-aided approach. In particular, we assess:

[1] Katja Mueller, 'Places and Spaces in the Temistou Meris (Fayum/Graeco-Roman Egypt): Locating settlements by multidimensional scaling of papyri', *Ancient Society* 33 (2003), pp. 103–125, hereinafter 'Places and Spaces (Themistos)' and eadem, 'What's your Position? Using Multi-dimensional Scaling (MDS) and Geographical Information Systems (GIS) for Locating Ancient Settlements in the Meris of Polemon/Graeco-Roman Fayum', *Archiv für Papyrusforschung* 50/2 (2004), pp. 199–214, hereinafter 'What's Your Position? (Polemon)'

[2] A somewhat similar idea was pursued in Katja Mueller, 'Mastering matrices and clusters. Locating Graeco-Roman settlements in the Meris of Herakleides (Fayum/Egypt) by Monte-Carlo-simulation', *Archiv für Papyrusforschung* 49 (2003), pp. 218–254, but the technicalities there were very different. We do not analyse that study here, although some comments in the present paper would apply to ir as well.

- whether the approach is scientifically well-founded,
- whether it was correctly applied,
- whether the obtained results were properly interpreted.

What we do not venture to consider are the obtained settlement maps themselves: we are only interested in the approach and its application, not in the final product of that application.

I. MULTIDIMENSIONAL SCALING (MDS)

1. The Approach

The main idea of the discussed studies is to analyze existing data on social interactions between certain settlements, and to extract information on the actual location of these settlements from that data. To this end, the technique of Multidimensional Scaling (MDS)[3] was used. In this section we informally describe the goal of MDS, its advantages and limitations.

Suppose a list of settlements from a certain time and area is available, together with data, for any pair of settlements i, j from the list, on how strong a certain form of interaction between these settlements was; this **degree of social interaction between i and j** is usually represented by a non-negative number denoted by $d_{i,j}$.

For example, $d_{i,j}$ may be the number of marriages between the inhabitans of settlements i and j over a chosen period, or the number of postcards sent from i to j and vice versa, *&c.* Typically, the degree of interaction does not differentiate between interaction going from i to j and interaction going from j to i; in other words, usually $d_{i,j} = d_{j,i}$. Obviously, this need not always be the case: for example, one could separately consider postcards sent from i to j (denoted by $d_{i,j}$), and separately postcards

[3] Cf. J. B. Kruskal, 'Multi-dimensional Scaling in Archaeology: Time Is Not the Only Dimension', [in:] D. G. Kendall, D. Tautu & F. R. Hodson (eds.), *Mathematics in the Archaeological and Historical Sciences. Proceedings of the Anglo-Romanian Conference, Mamaia 1970*, Edinburgh 1970, pp. 119–132; J. B. Kruskal & M. Wish, *Multidimensional Scaling, Sage University Paper Series on Quantitative Applications in the Social Sciences* 11, Beverly Hills – London 1978; T. F. Cox & M. A. A. Cox, *Multidimensional Scaling*, (2nd ed.), London 2001.

sent from j to i (denoted by $d_{j,i}$). However, for simplicity it is often assumed that all considered interactions are symmetric.

From these data, one wants to find out where the settlements i, j, *&c.*, were located geographically. For any pair i, j of settlements, consider the **geographical distance** between the real location of i and the real location of j; this distance, which we shall denote by $D_{i,j}$, is, of course, also unknown. The task is to infer the real geographical distances, the $D_{i,j}$, from the known degrees of social interaction, the $d_{i,j}$. Having done this, it should be rather easy to locate all settlements on the map. Indeed, knowing all the correct $D_{i,j}$ and the real locations of at least three settlements (from, say, archeological data), one can locate all other settlements on the list using a straightforward mathematical procedure.

The focus is thus on extracting as much information as possible about the real distances $D_{i,j}$ from information on social interaction $d_{i,j}$. The question is: what can be said, by analyzing the latter, about the former? Or in other words: what is the **dependency** between the $D_{i,j}$ and the $d_{i,j}$?

The answer depends of course on the specific problem to be solved, and especially on what the social interactions are and how the degrees $d_{i,j}$ actually measure them.

As a purely theoretical example, consider the time of an airplane flight between settlements (admittedly, a rather unusual notion of 'social interaction'). Suppose that the degree of social interaction is measured by taking as $d_{i,j}$ the fraction $1/t$, where t is the duration of a flight from i to j at the time of interest, expressed in hours. Recall that $d_{i,j}$ and $d_{j,i}$ are always assumed equal.

In this case, constructing a reasonable dependency between the degrees $d_{i,j}$ and the distances $D_{i,j}$ should be relatively easy. One could assume that $D_{i,j} = v / d_{i,j}$, where v is the velocity of a standard airplane that was in use at the time of interest (hopefully not Ptolemaic Egypt). Even if v is unknown, from this formula one can compute all the relations between real distances: that is, for any four settlements i, j, k and l one can tell what is the relation between the distance from i to j and the distance from k to l was (formally, this relation is defined by the ratio $D_{i,j} / D_{k,l}$, and this ratio can be computed from the formula above). If at least one real distance $D_{i,j}$ is known from other sources, then the factor v may also be computed, so that all real distances are known.

Even in this trivial example, a number of simplifying **assumptions** are made, all of them potentially dangerous. For example, it is assumed that the data $d_{i,j}$ refer to exactly one model of airplane, i.e., that on all flights the same type of airplane is used. Further, the assumed type of dependency does not take account of the time an airplane needs for take-off and landing, which would add a constant c to the formula: $D_{i,j} = v\ (1/d_{i,j} - c)$. Finally, it was assumed that mountains, winds, *&c.*, do not influence an airplane's flight, so that the flight duration depends solely on the geographical distance between the origin and the destination.

All these assumptions may be wrong, and so the results obtained must be treated with care; what is needed is, on the one hand, explaining why these assumptions seem plausible or, at least, why they are not too far from reality, and, on the other hand, checking what would happen if they turned out to be false: whether this would have a minor, or maybe a major influence on the final results.

A second thing to keep in mind is that the initial **data** may be, and indeed typically are, incomplete and/or flawed. It is tempting to assume that the $d_{i,j}$ faithfully represent some reality (duration of flights, inter-settlement marriages,[4] number of postcards sent, *&c.*). Unfortunately however, the collected data are usually far from being precise. It is therefore crucial to explain, on one hand, why the data appear to be of reasonable quality, and, on the other hand, to check what influence potential errors would have on the final results.

Because of possible errors in the assumed dependency between the distances and the data, and because of possible errors in the data themselves, it is typically impossible to find distances that exactly fit the data as prescribed by the dependency. For example, in the case of airplane flights, it may turn out that for some settlements i, j, k there is $d_{i,j} = d_{j,k} = 1$ and $d_{i,k} = 1/3$. However, whatever the velocity v, there cannot exist geographical distances satisfying $D_{i,j} = D_{j,k} = v$ and $D_{i,k} = 3v$. This is because the distance $D_{i,k} = 3v$ from i to k cannot be greater than the distance from

[4] Cf. D. G. Kendall, 'Maps from Marriages: an Application of Non-metric Multi-Dimensional Scaling to Parish Register Data', [in:] *Mathematics in the Archaeological and Historical Sciences* (cit. n. 3), pp. 303–318.

i to j, and then from j to k, that is, than the distance $D_{i,j} + D_{j,k} = 2v$. This discrepancy shows that either the assumed dependency between the distances and the data was wrong, or the initial data were flawed, or – most probably – both.

Because such discrepancies are inevitable in practice, one is forced to look not for a set of real distances exactly fitting the given data according to the given dependency, but instead for a set of distances that fit those data as good as possible. In other words, one looks for real distances $D_{i,j}$ which do not necessarily fit the formula required by the dependency exactly, but which fit it with as small an error as possible. This error, or 'badness-of-fit', is a nonnegative number called **stress**. If the fit is perfect, then the stress is zero. The role of the computer procedure called **MDS** (Multidimensional Scaling) is to find distances $D_{i,j}$ such that stress is as small as possible.

It might not be clear how stress should be defined. After all, people may have different views on which set of distances fits a given dependency 'better'. A standard method of defining stress is the following formula (here, $f(d_{i,j})$ is the value that $D_{i,j}$ should have according to the chosen dependency):

$$\frac{\sum_{i,j}(f(d_{i,j}) - D_{i,j})^2}{\sum_{i,j} {D_{i,j}}^2}$$

In the airplane case (without taking landing and take-off time into account), this formula takes the form:

$$\frac{\sum_{i,j}(v/d_{i,j} - D_{i,j})^2}{\sum_{i,j} {D_{i,j}}^2}$$

In other words, the error is calculated by taking, for all pairs of settlements i, j, the squared difference between what the distance should have been according to the chosen dependency, *i.e.* $v/d_{i,j}$, and what the distance $D_{i,j}$ is according to the given approximation. All these squared differences are added up and divided by a scaling factor.

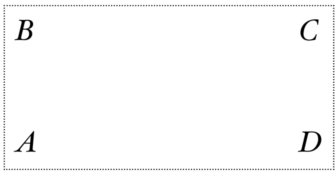
The MDS procedure arranges settlements in some n-dimensional space (that is, on the plane, or in 3-dimensional space, or in even more dimensions), and computes approximate distances $D_{i,j}$ such that the stress, as defined by the above formula and by a certain dependency, is as small as possible. Actually, the procedure does not even rely on a fixed dependency; instead, it attempts to choose a dependency from a prescribed class so that the stress can be reduced as much as possible. For example, the airplane velocity v need not be known to run the MDS procedure. Instead, it is enough to know that the dependency is of the form $D_{i,j} = v/d_{i,j}$ for some v and let the procedure try to choose v and the $D_{i,j}$ so that the stress is as small as possible. If the dependency is known to be of the form $D_{i,j} = v\, d_{i,j}$ (or, equivalently, of the form $D_{i,j} = v/d_{i,j}$), then the MDS procedure is called **metric MDS**. Of course, for geographical purposes one would typically be interested in the procedure performed in two dimensions (although see Section III.1 below).

The airplane example does not, of course, have anything to do with social interactions; it is mentioned here only because the dependency used there is so straightforward. In the case of real social interactions, dependencies don't have such a neat form. In such cases one can seldom assume more than that the dependency is monotonic, *i.e.*, that if the degree of interaction between settlements i and j is higher than that between settlements k and l, then the distance between the former is smaller than between the latter. Formally, this means that if $d_{i,j} > d_{k,l}$ then $D_{i,j} < D_{k,l}$. This assumption is weaker than the one in the airplane example, as there may be many ways of arranging the settlements on the plane so that all required inequalities hold. For example, for four settlements named A, B, C, and D, and for

$$d_{A,A} = d_{B,B} = d_{C,C} = d_{D,D} = 3,$$
$$d_{A,B} = d_{C,D} = 2, \text{ and}$$
$$d_{A,C} = d_{A,D} = d_{B,C} = d_{B,D} = 1,$$

one effectively knows only that A and B, as well as C and D, are closer to each other than A and C, A and D, B and C, or B and D. Thus all the following configurations of settlements fit the data perfectly for some monotonic dependencies:

or



or

In practice, however, even with the weaker assumption on the dependency it may still be impossible to fit data to real distances perfectly. When this happens, one again looks for as good a fit as possible, i.e., one with minimal stress, among configurations for which dependencies are monotonic. The procedure used to find the best fit is called **ordinal** or **non-metric MDS**. The name 'ordinal MDS' is appropriate, because only the order of the d_{ij} is relevant when choosing the best set of approximate distances; the exact numbers have no influence. Thus, in the above example, the number 100 could be used instead of 3, the number 12 instead of 2, and the number 0.25 instead of 1, and nothing would have changed in the final result.

2. Interpretation of Results

Both metric and ordinal MDS are well-established procedures, successfully used in social and medical sciences. They provide a conceptual, graphical presentation of some objects on a plane (or in an *n*-dimensional space), based on information about their similarities, or degrees of interaction. This way, the analyst can easily identify clusters, *i.e.*, sets of objects located close to one another. In many applications, one can also identify the number of dimensions needed for an adequate presentation with a small stress, thus finding the number of factors influencing the given notion of interaction between objects. In a celebrated example unrelated to geography[5] objects were Morse Code symbols and the data $d_{i,j}$ measured the subjective, perceived similarity of symbols, assessed empirically by a set of tests. Using MDS, the symbols were then placed in *n*-dimensional space for various *n*. It turned out that for $n=2$ one gets a map with reasonably low stress, where one axis determines the complexity of the Morse symbol (i.e., the number of all '–' and '.' in it), while the other determines whether there are more '–' in the symbol, or more '.' Based on this, it was concluded that the perceived similarity of Morse Code symbols depends on two factors (variables): the length of the symbol, and the number of '–' in it. Thus the similarity of Morse Code symbols seems to be a 2-dimensional phenomenon. In geographical applications, this ability of MDS to determine the dimensionality of a problem is less useful, since settlements are assumed and expected to fit in a 2-dimensional plane.

MDS is much less reliable when used to determine the precise location of objects in a space or on a plane, especially if the initial data $d_{i,j}$ are imperfect. Small errors in the initial data can substantially change the obtained locations of objects, even if the general picture of their configuration remains roughly the same. In some applications, as in the Morse Code example, only the general picture matters; however, when locating geographical settlements based on social interactions between them, it is important to remember that MDS cannot provide more than a rough

[5] Cf. R. N. Shepard, 'Analysis of proximities as a technique for the study of information processing in man', *Human Factors* 5 (1963), pp. 33–48, and Kruskal, 'Multi-dimensional Scaling in Archaeology' (cit. n. 3), pp. 119–132.

map. In particular, it would make little sense to plan any kind of archeological works based on MDS results. More reasonably, such results, when carefully obtained and interpreted, can be used to assess competing theories or maps obtained using other methods.

Moreover, it is important to know that MDS is a randomized procedure, and it can give different results when run multiple times on the same data. The procedure, trying to find a configuration of settlements matching the given social interaction data, begins by placing the settlements randomly on a plane and then tries to improve the configuration by small changes of locations. If the stress cannot be reduced by a small configuration change, the procedure stops. This means that there is a danger of finding a fake, local stress minimum instead of the best configuration of objects. For this reason, it is necessary to repeat the MDS procedure multiple times, starting from different random initial configurations, and report the configuration with the smallest stress as the best one.

For a meaningful interpretation it is also necessary to assess the influence of potential data imperfections on the final result of the procedure. To this end, the analyst might introduce intentional small changes to the data and check how they change the resulting configuration of settlements. She might also run the whole procedure on a subset of available data and compare the results with those obtained on the whole set of data. Without these steps it is hard to tell whether the settlement configuration obtained with the use of MDS is meaningful at all.

Finally, it must be remembered that MDS, a mere computer procedure, relies on a very simplified perspective of the world, encoded in a simple matrix of numbers $d_{i,j}$. In all but the simplest applications, this leaves aside plenty of knowledge about the problem domain. For example, locations of some settlements might be known from archeological data; other settlements might be known to be located near a big river; *&c*. Information of this kind is ignored by the MDS procedure, which means that its results must be treated with care and compared with knowledge obtained from other sources. In particular, one must not treat MDS results as decisive arguments that immediately falsify all theories inconsistent with them.

These caveats apply to all applications of MDS, however carefully designed and performed. In the following sections we attempt to assess the application studied in Mueller's articles and make some further comments related to that specific application.

II. THE DATA AND THE DEPENDENCY

The data used in 'Places and Spaces (Themistos)' is extracted from papyri in which names of 36 settlements from the Meris of Themistos arise; in the other work, papyri with names of 25 settlements from the Meris of Polemon are taken into account. In both papers, the assumption is made that if the names of two settlements occur in the same papyrus, then these settlements were probably located close to each other.

In both papers the degree $d_{i,j}$ of social interaction between settlements *i* and *j* is **binary**, i.e., defined to be:

- 1, if *i* and *j* occur together on some papyrus,
- 0, otherwise.

To these data *ordinal MDS* is then applied.

Applying ordinal MDS to these data means that the computer tries to distribute settlements on a map in such a way that if settlements *i* and *j* occur together on some papyrus, and settlements *k* and *l* do not occur together on any papyrus, then the distance between *i* and *j* is smaller than between *k* and *l*. Since finding such a configuration is impossible, the procedure looks for a map as close as possible to this ideal.

The choice of data and of the dependency is, for a number of reasons, controversial.

1. Choice of Data

The choice of data itself raises some questions. For instance, both in 'Places and Spaces (Themistos)' and in 'What's Your Position? (Polemon)' texts listing more than 5 settlements have been excluded from consideration. Excluding texts listing a large number of settlements may be justified, as they could introduce some misinformation. In an extreme case, a hypothetical papyrus listing all known settlements would immediately set all $d_{i,j}$ to 1, and all information contained in other papyri would be ignored. However, it is unfortunate that the reader has not been informed whether excluding texts with over 4, over 5, over 6 or over 7 settlements would change the results or not. An answer to this question is important: after all, there is no reason to believe that the number 5, chosen by the author, is any better than 4, 6 or 7. It should be mentioned that the harm due to papyri listing many settlements would disappear if the aggregate approach were used instead of the binary one (see Section 11.2).

Somewhat more questionable is the manner in which texts listing settlements from outside the Meris of Themistos have been excluded from consideration in 'Places and Spaces (Themistos)'. The problem is that 'some of these texts have been included' (see p. 108), while most have been excluded. No criterion of exclusion/inclusion is cited, so that it seems that the decision was taken with full discretion. Such an approach makes the results hard to reproduce. Again, it should also be checked how the obtained settlement configuration would change if other, all or none such texts were considered.

On the other hand, some additional preliminary analysis of the data would have been helpful. For instance, among the analyzed settlements Herakleia is special: in no used text is Herakleia connected to any other settlement. In other words, the used data effectively do not contain *any* information that could help locate Herakleia – with one exception. The only geographical information one could extract from Herakleia not being mentioned anywhere is that it is likely to have been far away from all other settlements. This information is not at all certain: Herakleia may not have been mentioned simply because it was too small, or for any other reason not dependent on its geographical location. In any case, Herakleia not being mentioned cannot be the base of any serious claim on its location. The MDS procedure did of course locate Herakleia somewhere (trying to place it as far as possible from other settlements), and based on this it has been placed in cluster 11, subcluster 2.4.[6] This cannot be accepted: there is nothing in the data that would allow one to place Herakleia close to

[6] 'Places and Spaces (Themistos)', p. 112

Chalkorychia and Archelais; these settlements are similar only in one respect: they all have little connections with other settlements. A careful preliminary analysis of the data should have induced one to remove Herakleia from the data.

The author has probably recognised the above problem, as in the final map of 'Places and Spaces (Themistos)' the three settlements are actually located quite far apart, reasonably based on classical analysis unrelated to MDS; in fact, in 'What's Your Position (Polemon)?' settlements such as Herakleia, *i.e.*, ones not connected to at least 2 other settlements, were excluded from the study.[7] This example shows that the results of the automated computer procedure must be validated using other approaches, and that whenever they contradict other evidence, one must be ready to discard them.

2. Aggregated vs. 0–1 Approach

Instead of the binary approach, where the $d_{i,j}$ are defined to be 0 or 1 only, one could adopt an aggregated measure of social interaction, where $d_{i,j}$ is the number of papyri that list both settlements i and j. This would provide more refined information about the configuration of settlements; indeed, a dozen papyri containing the names of two settlements provide a stronger indication of them being close to each other than just one such papyrus. However, in the binary approach such distinctions are ignored. In 'Places and Spaces (Themistos)' we read:

> In theory, it is also possible to generate an aggregated matrix. If a combination of two settlements reoccurs several times, the occurrences can, for instance, be counted and inserted into the matrix. For reasons explained below, no aggregated matrix was used. [p. 107]

Later, the reasons for not using the aggregated approach are stated as follows:

> Despite the mass already collected in PP Online, GEOREF[8] is incomplete and patchy at places. Not every published papyrus has been investigated for its spacial-geographical content and entered into GEOREF. Effectively,

[7] 'What's Your Position? (Polemon)' , p. 204, ft. 12.

[8] The database containing papyri used in the article.

> with the progress in publication of Greek and Demotic-Egyptian documents, it would be unrealistic to expect that PP Online could ever reach towards completion. Nonetheless, GEOREF in its present state gives a fair representation of the available data. What it cannot do is, however, to get near to provide us with safe grounds for a percentual or aggregated approach, as discussed above. [p. 109]

Unfortunately, this explanation is itself rather patchy at places. It is unclear why the database not containing all published papyri could have any influence on the choice of approach in the first place. Why should the ratio of papyri in the database to all published papyri be of interest? Would not the ratio of papyri in the database to all produced papyri be more important? This aside, it is true that with a highly incomplete database the aggregate approach might theoretically be more risky, as it depends on refined quantitative information that in the binary approach is simply ignored. If that information is seriously erroneous, it might be a good choice to ignore it. However, no analysis of the available data was presented in 'Places and Spaces (Themistos)' that would suggest that such harmful errors indeed do exist in the database used.

Instead, in 'What's your position? (Polemon)' the author attempts to estimate the 'completeness' of the available data on 25 settlements [pp. 202–203]. Rather surprising statements can be found in that part of the text. For instance, one reads:

> The traditional perception would be that more texts would add more settlements to the matrix. But this is not the case. The number of settlements attested in a matrix is not linearly correlated to the number of texts being used.

Apparently, the notion of monotonic dependency ('more texts implies more settlements') is confused here with linear correlation. The 'traditionally' expected monotonic dependency may exist, contrary to what the author seems to believe, even if the correlation isn't at all linear. In fact, it is quite obvious that indeed more texts imply more settlements, unless one takes the view that *all* settlements that have ever been mentioned in any text whatsoever have already been included in the matrix. Even more surprisingly, the author states:

> More texts do not generate more settlement entries in a matrix. It is probable that a different set of the same number of sources would contain other settlements, but not necessarily more settlements.

Here, the second sentence contradicts the first sentence. If two sets of texts generate different sets of settlements, say S_1 and S_2, then clearly taking both sets together, i.e., taking 'more texts', will generate the union of S_1 and S_2, and this union necessarily contains more settlements than S_1 or S_2 did.

Note that the author cites two older papers where settlements were located by using common occurrences in texts,[9] and that in both papers the aggregated approach was used. It is true that in Cherry's article the 0–1 approach was used as well, and then compared with the aggregated one, and that is was concluded that using the 0–1 approach 'results in little real information loss'.[10] Nevertheless, nothing in these papers supports the view that the 0-1 approach is in some sense better or less error-prone in practice. In any case, one could have simply applied both approaches, so that the results could be compared to assess their stability under changes in design decisions.

One should also mention that in the literature on ordinal MDS one often reads that data containing ties, i.e., data where the d_{ij} take only few distinct values (as here: 0 and 1) should be considered as dangerous and leading to relatively large errors. Thus Joseph B. Kruskal and Myron Wish write:[11]

> Unless there are a great many ties, their presence does not affect anything very much. The only common situation when there are so many ties that they have an effect is when the proximity values take on only a few distinct values.

All this makes the decision to apply the 0–1 approach rather arbitrary, especially if one considers the rather unconvincing motivation behind it.

[9] W. Tobler & S. Wineburg, 'A Cappadocian Speculation', *Nature* 231 (1971), pp. 39–41, and J. F. Cherry, 'Investigating the Political Geography of an Early State by Multidimensional Scaling of Linear B Tablet Data', [in:] J. Bintliff (ed.), *Mycenean Geography: Proceedings of the Cambridge Colloquium, September 1976*, Cambridge 1977, pp. 76–82.

[10] Cherry, 'Investigating the Political Geography' (cit. n. 9), p. 80.

[11] Kruskal & Wish, *Multidimensional Scaling* (cit. n. 3), p. 53.

3. The Dependency

Another, far more important, criticism, is that by feeding the MDS procedure in 'Places and Spaces (Themistos)' and 'What's your position? (Polemon)' directly with the $d_{i,j}$, the author neglects the size of the settlements considered.

It is trivial to observe that the degree of social interaction depends not only on the geographical distance between settlements, but also (and maybe foremost) on their size. This is usually taken into account by a formula of the form

$$D_{i,j} = c \cdot \sqrt{P_i \cdot P_j / d_{i,j}}$$

where P_i and P_j are the populations of settlements i and j, respectively, and c is some constant. This formula, used by Tobler and Wineburg,[12] and cited in a somewhat more general form by Cherry,[13] takes into account that a high number of interactions between i and j might be due not to the small distance between them, but to their big populations. If i, j, k and l are settlements, the first two big, the last two small, then a common occurrence of i and j on some papyrus contains less geographical information than a common occurrence of k and l. This difference is crucial, and not taking it into account makes the obtained results questionable. This is true not only in the aggregated approach, but also in the 0–1 approach applied here.

The only reason why populations could be omitted in the analysis would be that, in fact, all populations were roughly the same. It is for this reason that Cherry[14] omitted the population factor in his analysis. However, in the case of the Merides of Themistou and Polemon, this argument has not been put forward, and, for reasons presented below, probably could not be put forward at all.

[12] Tobler & Wineburg, 'A Cappadocian Speculation' (cit. n. 9), p. 40.

[13] Cherry, 'Investigating the Political Geography' (cit. n. 9), p. 78.

[14] Cherry, 'Investigating the Political Geography' (cit. n. 9), pp. 78–79.

If one tries to take populations into account, then one inevitably runs into the problem of estimating those populations (or, more precisely, of estimating them in relation to one another). A possible source of information on the size of settlements may be found in the analyzed papyri themselves: the more often a certain settlement is mentioned in documents, the bigger its population might have been. Analysis of the data used in 'Places and Spaces (Themistos)' shows that settlements appearing there might be of very different sizes. Some settlements appear in documents in connection with about half of the 36 settlements (*e.g.*, Alexandrou Nesos), while others appear in connection with only two other settlements (*e.g.*, Chalkorychia, Argias), or are not connected with any other settlement (Herakleia). Of course, these differences need not be due only to differences in size – Alexandrou Nesos may have been located in the middle of a highly populated area, while Chalkorychia, Argias and Herakleia might have been located in a remote part of the meris. Nevertheless it seems reasonable to assume that these settlements weren't all of similar importance. A similar remark applies to the Meris of Polemon; there, the author actually notices that 'the level of incidences a settlement generated with other settlements — hence a settlement's connectivity — can differ considerably'. For these reasons one could hardly argue that settlements considered in 'Places and Spaces (Themistos)' and 'What's your position? (Polemon)' were all of similar size and that for this reason the population factor could have been omitted.

4. Related Work

To our knowledge, the method applied in 'Places and Spaces (Themistos)' or 'What's your position? (Polemon)' has not been used *in this form* before. MDS in general does not normally serve as a tool for locating settlements; this is acknowledged by Mueller herself.[15] The first application of MDS to finding settlements by counting common occurrences in texts is, to our knowledge, the study by Tobler and Wineburg.[16] That paper differed from

[15] 'Places and Spaces (Themistos)', p. 107.

[16] Tobler & Wineburg, 'A Cappadocian Speculation' (cit. n. 9)

Mueller's work in a number of respects: the aggregated method was used, populations were taken into account, and from the formula used it appears that there were less ties. More importantly, the paper by Tobler and Wineburg can hardly be described as a scientific success, either theoretical or empirical. On the theoretical side, the paper contains no considerations proving the validity of the approach; it is more of an experiment. On the empirical side, there is no proof whatsoever that the locations obtained are correct. All this has been acknowledged by Cherry, who called the study 'somewhat unsuccessful'.[17]

The second application was Cherry's study[18] on settlements in Messenia, based on linear B tablet data. This is the study Mueller refers to in her work. Cherry used both the aggregated and the binary method, concluding that using the latter resulted in only a slight deterioration of the map [page 80]. In fact, he did not measure the number of common occurrences of settlements in one text; instead, he counted cases where settlements appeared adjacently in one text. He also clearly made the assumption that almost all settlements were of the same size [pages 78–79]. As a result of his work, Cherry obtained a map that resembled those obtained by conventional analysis. His paper is of exploratory character and cannot justify treating the procedure as an established application of MDS. In particular, the paper contains no estimation on how errors in the data or in the dependency would influence the results.

All this means that the MDS procedure as used by Mueller, even if other problems mentioned here are dealt with carefully, cannot at this point be treated as a basis for a definitive map. This would require a thorough numerical analysis of the method on existing and known data, as well as a confirmation of MDS findings with other methods. The successful studies by Kendall,[19] who (among other things) reproduced the map of modern France using an MDS procedure, are not enough in this respect, since they were based on errorless data of a form very different from those used in 'Places and Spaces (Themistos)'.

[17] Cherry, 'Investigating the Political Geography' (cit. n. 9), p. 79.

[18] Cherry, 'Investigating the Political Geography' (cit. n. 9).

[19] Cf. D. G. Kendall, 'Construction of maps from odd bits of information', *Nature* 231 (1971), pp. 158–159, and idem, 'Maps from Marriages' (cit. n. 4).

III. INTERPRETATION OF RESULTS

Up to now, the reader may have had the impression that in both 'Places and Spaces (Themistos)' and 'What's your position? (Polemon)' the MDS procedure was used to automatically generate, from the data and dependency described above, a map of settlements of the appropriate meris; after all, this is what Tobler and Wineburg, Cherry, and Kendall did. But this is not the case. We now sketch the actual manner of applying MDS in these two papers.

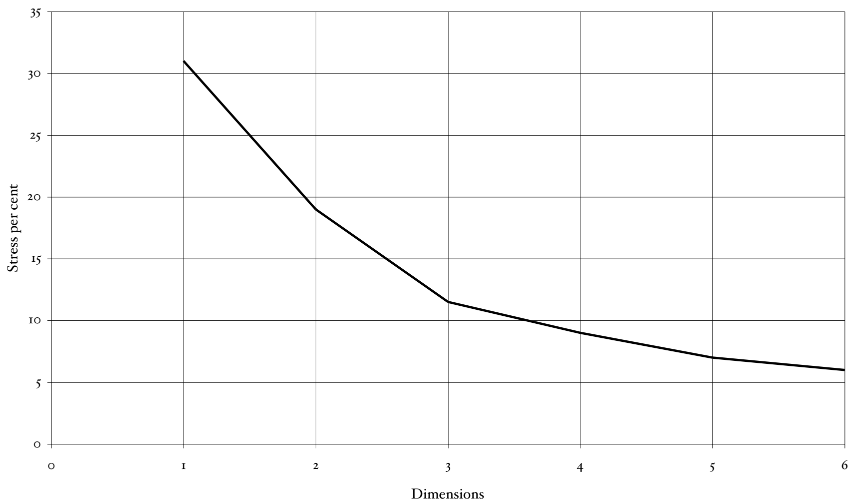
1. Meris of Themistos

In the paper on the Meris of Themistos, a modification of the method described in Sections I & II is used.

In a geographical context, MDS would normally be used to locate settlements on a map, that is, in a 2-dimensional space. However, MDS may also be used, and actually is typically used, for applications that have nothing to do with geography. In general, MDS is capable of placing objects in a space of an arbitrary number of dimensions, so that the real distances $D_{i,j}$ in the space resemble, as close as possible, the degrees of interaction $d_{i,j}$. The distances $D_{i,j}$ normally do not have anything to do with distances in a geographical sense, as in the study of Morse Code described in Section 1.2 above, where MDS was used to determine the dimensionality of available data.

Of course, the dimensionality of data is only a matter of interpretation. Even Morse Code data do not fit perfectly in 2-dimensional space. The stress obtained there is non-zero, and adding dimensions yields a better fit (*i.e.*, lower stress). But two dimensions seem a reasonable choice, because, first of all, the axes have a natural interpretation (length and the number of '–'), and, moreover, because moving to one dimension increases stress significantly, while adding dimensions decreases stress only by a small margin.

The latter reason is related to the 'elbow approach' to finding the dimensionality of given data. According to this approach, one considers the stress of the data after fitting it into 1, 2, 3, *&c.*, dimensions via MDS, as in the following figure:



One then finds an ‘elbow’ in the graph; the x-coordinate of this elbow is probably the true dimensionality of the data.

Mueller used the elbow approach to find the dimensionality of the 0–1 matrix used in estimating the locations of settlements in the Meris of Themistos, and she found an elbow suggesting that the data is 3-dimensional. This is quite surprising, as maps of Ptolemaic Egypt are just as 2-dimensional as all other maps. The data being truly 3-dimensional would mean that the social interaction studied (common occurrences on papyri) depends not only on two geographical coordinates but also on some other, third factor. An interpretation of the third factor would be an interesting problem itself: it could be time, an indication that papyrus data are heterogeneous, or a combination of several non-geographic factors. In ‘Places and Spaces (Themistos)’ no attempt was made to provide such an interpretation. Whatever the third factor though, its significance undermines the main assumption underlying the application of MDS to locating settlements: that the degree of social interaction depends solely, or at least mainly, on geographic location. This means that any 2-dimensional map obtained from data used in 'Places and Spaces (Themistos)' is unreliable, or at least seriously imprecise.

For this reason, the author did not rely on MDS as the source of an automatically generated map. Instead, she used the computer procedure for the much more modest purpose of finding geographic clusters, i.e., groups of settlements which were placed relatively close to one another by the MDS procedure in three dimensions. Under the assumption that these 3-dimensional clusters correspond to geographical clusters, she then arranges them on a 2-dimensional map using more traditional analysis.

This approach is much more cautious than a simple automatic use of MDS in two dimensions would be, but is still subject to some criticism. In particular, two settlements located far away from each other in three dimensions might be located quite close geographically, if they differ significantly with respect to the third, uninterpreted factor of the degree of social interaction. Worse, if the third factor bears much more significance than the geographical coordinates, even two settlements located in one 3-dimensional cluster could in fact be located far away geographically. It is hard to exclude either of these two unfortunate possibilities unless the third factor is properly interpreted.

After having divided all settlements into clusters, Mueller again uses MDS to find subclusters, and then arranges all the settlements, using this division and 'traditional' analysis, on the (2-dimensional) map. Using traditional analysis is probably inevitable, but it detaches the final map from the results of the MDS procedure. For example, the settlements of Herakleia, Chalkorychia and Archelais, put in the same cluster by the automatic procedure, end up quite remote to one another after the application of traditional analysis. This does not mean that the traditional analysis was flawed; instead it is an indication of the limited reliability of MDS in this context. In any case, it must be stressed that the final map of the Meris of Themistos presented in 'Places and Spaces (Themistos)' is only partially based on the automatic procedure of MDS, and that a different choice of archeological argument and traditional analysis might lead to quite a different map.

2. *Meris of Polemon*

A different approach is taken in 'What's your position? (Polemon)'. There, we find no mention of the 'elbow', or of plotting settlements in 3-dimensional space. Although this is not explained clearly, it seems that instead the author used the method described in Section 1.1 above, letting MDS directly produce a 2-dimensional map of the Meris of Polemon. It is interesting to note that there is no explanation whatsoever on why different methods were applied to the same problem in the case of both Merides.

In 'What's your position? (Polemon)', the above standard method is not applied to the whole binary data matrix of 25 settlements. Instead, 10 square submatrices are randomly chosen, and 10 maps of the Meris of Polemon are produced. Each of the chosen square submatrices contains 15 settlements, among which the following 3 are required to appear: Talithis, Tebetny and Tebtynis. The idea is that since the true locations of these settlements are known, it should be possible, for each of the 10 submatrices, to locate the absolute, geographical locations of the other 12 settlements.

This way, the author is able to produce 10 maps, each of 15 settlements. Unfortunately, even though much is said in the paper about Minimum Enclosing Rectangles, Geographical Information Systems, *&c.*, we do not learn how exactly these maps, with their absolute locations, are generated from the abstract maps produced by MDS. This makes the results impossible to reproduce.

If one now takes any of the 25 settlements except Talithis, Tebetny or Tebtynis, then such a settlement may have been placed in a number of different locations in each of the maps (this number may vary from 10 to 0; hopefully the latter possibility was actually blocked). The next step in the paper is to calculate, for every of the 25 settlements, a geometrical 'average' of all these different locations. This average, called centroid, is then plotted on the map.

The centroids are never compared to the results of performing MDS on the whole 25 by 25 square matrix. Instead, for each settlement the author computes the average distance between its possible locations and the corresponding centroid. This serves as a measure of the stability of the

MDS solution: the higher the average distance, the less reliable is the location of a settlement obtained via MDS.

This general idea must be appreciated, as this is the first time in both of Mueller's studies that the stability of MDS solutions under changes in initial data is tested at all. One might ask, however, why a more natural stability test is never performed, where intentional changes are made to the number and structure of the set of papyri considered, rather than to the set of settlements. This omitted stability test is especially important in the binary approach with sizes of settlements ignored (see Sections II.2–II.3 above). Indeed, in this approach even a single discovery of one new papyrus that lists together two otherwise unconnected settlements could change the initial data matrix quite substantially. If such changes influenced the final MDS arrangements of settlements, the entire approach would have to be considered invalid. It is unfortunate that this dangerous possibility is never excluded in 'Places and Spaces (Themistos)' or 'What's Your Position? (Polemon)'.

Let us also mention that the result of the centroid-based stability test, *i.e.*, the average distance of 5.01 km between a settlement location and its respective centroid, is never interpreted in the paper. The reader is left to decide whether this is a minor nuisance, or a major error invalidating the approach. Even though 5.01 km may seem little, it seems a considerable distance in the context of the Meris of Polemon. The Meris, as depicted on Fig. 4 in the text, stretches about 20 km North-South, and about 30 km East-West. This means that the average error of 5.01 km amounts to 15% to 25% (depending on the axis) of the size of the Meris, which seems rather significant.

IV. CONCLUSIONS

MDS is an established data analysis procedure, with many applications in the social and medical sciences, where large amounts of data need a conceptual, graphic presentation. Using MDS to locate settlements based on degrees of social interaction between them is another interesting application. However, results of this application need to be interpreted cautiously, for several reasons.

1) Any automated computer procedure uses only a limited amount of data about the analysed problem; in the case of MDS, all data used are simple matrices of numbers. This means that much existing knowledge about the analysed objects (settlements) is ignored, and that the resulting map must not be treated as a precise indication of the location of these objects. At best, it can only provide a rough picture of the real spatial arrangement.

2) To perform the MDS procedure on an existing body of data, one needs to make several design choices. In the papers studied here, the author selected papyri and settlements to be considered, preferred the binary approach over the more refined aggregated one, decided to disregard the size of settlements, *&c.* Moreover, there are some differences between the manner of application of MDS in 'Places and Spaces (Themistos)' and 'What's Your Position? (Polemon)'. Some of the design decisions appear objectionable (especially the disregard for the size of settlements), but even if they were not, one must be aware that different decisions might lead to significantly different results. To confirm the validity of the obtained maps, one should carefully check that this is not the case, or convincingly prove that the decisions made were optimal.

3) Another important question is the stability of final MDS results under small changes in initial data. This is especially important since the data used in this particular application, i.e., sets of papyri, are necessarily very far from being complete. Neither in 'Places and Spaces (Themistos)' nor in 'What's Your Position? (Polemon)' was it checked how a potential discovery of a new small set of papyri, connecting otherwise unrelated settlements, would influence the final result of MDS. It is therefore not easy to assess the reliability of those results.

4) These limitations make it necessary to use other evidence to confirm or reject maps obtained via MDS. In both papers the author, quite reasonably, analyzed existing sources in a more traditional fashion to locate some of the settlements more reliably. It must be noted that the results of that analysis differed considerably from the automatically obtained results of MDS. This means that the automatic procedure did not fully determine the final results, and that a different choice of traditional arguments might lead to very different maps.

To conclude, one should not be intimidated by the fact that maps in 'Places and Spaces (Themistos)' and 'What's Your Position? (Polemon)' were obtained with the use of computers. Computer programs are very powerful and robust, but when confronted with complex problems based on heterogeneous, hard to formalise, and incomplete data, their results need to be interpreted with care, especially if the manner of their application depends on some arbitrary design decisions. Maps of Ptolemaic Egypt obtained with MDS are suggestions rather than definitive solutions, and they must not be used to discredit arguments, evidence or theories that contradict them.

Piotr Hoffman

Institute of Informatics
Faculty of Mathematics, Informatics and Mechanics
Warsaw University
ul. S. Banacha 2
02-097 Warsaw
POLAND

e-mail: *piotrek@mimuw.edu.pl*

Bartek Klin

Laboratory for Foundations of Computer Science
School of Informatics
University of Edinburgh
The King's Buildings
Edinburgh EH9 3JZ
UNITED KINGDOM

e-mail: *bklin@inf.ed.ac.uk*