Hans-Hermann Bock · Peter Ihm (Eds.)

Classification, Data Analysis, and Knowledge Organization

Models and Methods with Applications

Proceedings of the 14th Annual Conference of the Gesellschaft für Klassifikation e.V., University of Marburg, March 12-14, 1990

With 148 Figures

Springer-Verlag
Berlin Heidelberg New York
London Paris Tokyo
Hong Kong Barcelona

The Use of Multivariate Statistics in Scandinavian Archaeology

Torsten Madsen
Institute of Archaeology, University of Århus
Moesgård, DK-8270 Højbjerg

Abstract: The use of statistics in Scandinavian archaeology started after the second World War only, and it has been slow to penetrate beyond a very basic level. To day, where complex multivariate techniques are within easy reach of all who care, the proper use of statistics in archaeology has become a pertinent problem. How does the use of statistics fit into the archaeological research process, and what are the demands on the archaeologists who apply complex multivariate techniques in their research. Experiences with multivariate statistics in Scandinavian archaeology in recent years has helped to clarify the archaeological research process, and been very informative on the use of multivariate analyses in archaeology.

1. Introduction

Statistics have always played a role in archaeology. By nature the archaeological material invites to quantification. Tables, especially, have been common since a very early date (i.e. SARAUW (1903)), even though they did not result in a more systematic application of the quantitative information. However, it was only by the middle of this century that formalised treatment of quantitative data beyond pure tabulations was seen. In Scandinavian archaeology it was P.V. GLOB (1944) who set the stage for a new epoch. He used combination diagrams to show closed finds and stratigrafic relations, line graphs to show depth of burials in relation to axe types, and some quite original "compass rose" diagrams to show the orientation of graves.

The actual breakthrough in Scandinavian archaeology for the treatment of numerical data by way of formalised statistical methods, however, came with the thesis of Mats P. MALMER (1962). For a new and rapidly increasing generation of young archaeologists this book became the source of inspiration from where the use of statistics in archaeology emerged. Beside the demands for formal type definitions it was MALMER's use of histograms, correlation diagrams and isoritmic diagrams (Fig. 1) for the treatment of formal measurements of artifacts that more than anything else became influential. As it turned out, statistics never gained the influence and importance that it should have. In England DORAN AND HODSON (1975) published their *Mathematics and Computers in Archaeology.* This important work ought to have exerted a marked influence in Scandinavia, but on the whole it was ignored. The only person deeply interested in the book seems to have been Carl-Axel MOBERG. In 1976 he produced a small paper, *"Structureville" A young person's guide to find analysis*, which was an interpretation and explanation to DORAN AND HODSON's book. MOBERG's paper ought to have been the bible laying on the night tables of all archaeologists. Instead it is largely unknown as it was distributed as handouts, only.

2. "The Good, the Bad and the Ugly"

David Hurst THOMAS (1978), clearly inspired from a well known movie, used the above caption to characterize the use of statistics in archaeology. It is indeed a caption that can be used in connection with Scandinavian archaeology over the last 25 years as well. And as in American archaeology it is not the good statistics that dominates. It is predominantly the ugly and to a lesser degree the bad statistics that hold the stage.

And what then do we mean when we speak of good, bad and ugly use of statistics? As for good statistics the concept itself is clear enough, but to the average archaeologist it is extremely difficult to decide whether an actual application of statistics found in a paper is to be considered good or not. There are two basic criteria that can be used. One is whether the statistical method is used technically correct. The other is whether the method fulfils a legitimate and sensible purpose in the actual context. These two criteria in reality are the negation of the bad and the ugly statistics, but then again this is not the whole truth of good statistics. A statistical method may be used correct in a technical sense and for a well defined purpose, and yet it may not be considered to be good if its use goes beyond what we, based on our theoretical-methodological background, will accept as meaningful in the study of human societies. Below I shall return to a category of statistics – the inductive – that I find hard to classify as good in most archaeological contexts. But this does not make it either bad or ugly, nor does other archaeologists need to agree with me, when I claim that it is not good in the actual context.

The bad statistics in principle is easier to detect, as it is a question of whether an actual statistical procedure is used in accordance with its specific preconditions, and in accordance with the general rules of using statistics. It may, however, be difficult for the individual researcher to decide whether an actual case of statistics is bad or not. The reason is that archaeologists in general have a rather shaky knowledge, when it comes to statistics. They have difficulties seeing through a case of bad statistics, and it is easy for them to produce a case of bad statistics themselves out of ignorance. I am not even referring to complicated statistics, where the average archaeological scholar has no chance at all. I am referring to relatively simple, straightforward statistics that most archaeologists have a fair idea about, but where they are too inexperienced to realize when it is misused. Bi-variate correlation statistics is here a classic. Everybody knows of scattergrams, where two variables are placed on separate axes in a perpendicular coordinate system, and where a point in the system reveals how an observation relates to the two variables. It is also commonly understood that if the points turn up in a pattern – for instance tend to lie on a straight line, then this somehow indicates a relationship between the variables. The problem, however, is that correlations can be odd and deceptive. If you measure the IQ of all children in a school, and at the same time note their shoe numbers, then you are going to find a high degree of correlation between IQ and shoe numbers. The correlation is real enough, but it has no explanatory value. It is a third value, age, which correlates with both IQ and shoe number, and binds both together in logical dependence. Logical dependences are the greatest pitfall in studies that use correlation, and it may turn up on many levels and deeply buried in the data. It often demands a deep insight into the data material to reveal logical dependence, and it is seldom or never as simple as the example with the IQ and shoe numbers. Earlier I have mentioned an example of the separation of thick butted flint axes of A and B type (MADSEN (1985): 19), and I should add that I am always very suspicious of studies showing nearly perfect linear correlations between variables, when these are of an index type. That is to say variables that have been calculated by a formula that includes a number of primary variables (for an example see HEDEAGER (1978)). It is often extremely difficult to understand what the individual indexes stand for, and even more difficult to evaluate if logical dependence is present.

The third category – the ugly statistics – is the easiest to trace, and at the same time the most harmless one, but also, when your eyes has been opened to it, the most annoying one. Ugly statistics are statistics with no purpose. It is the histogram that show the length of a set of flint axes, or a scattergram that show the relation between length and width of a set flint flakes, where neither the histogram nor the scattergram is used for any purpose in the further study of the material. Recent archaeological literature is often overburdened with this type of ugly statistics, primarily because it is considered to be sophisticated to use these types of diagrams, but also because archaeology, by tradition, is a very much describing and seldom problem oriented discipline. It is thus often considered legitimate to publish descriptions of data even if these do not lead to any conclusions, and sheltered by this view the ugly statistics thrives well.

3. The nature of the archaeological research process

It is common to view statistics – and computers for that matter – as neutral tools, that can be brought to use in a way which is unbiased by the view of the archaeological research process held by the researcher. This is not true. The understanding of what is the correct use of statistics is inseparably tied up with the way the research process is viewed, and with the way it is carried out in practice. It would be too much to expect that scholars should agree to the nature of the archaeological research process, let alone claim that archaeological research in general is done along uniform theoretical and methodological lines. With differing attitudes towards the research process the view of the role of statistics will also differ. Therefore, it is imperative that all researchers analyse their attitude towards the research process, and coin this into a stand-point regarding the role that statistics can and should play in archaeological research. Only by occupying oneself with the role of statistics, a researcher may gain a deeper understanding of the subject matter, and only then can we hope to get rid of the bad and the ugly statistics. In the following I shall give my own contribution to the understanding of the role of statistics in archaeology based on my personal view of the archaeological research process.

Archaeological research involves two very different realities. One is the reality of prehistoric societies. This is a reality that can no longer be observed even though it is the target of archaeological research. The other one is our present reality which we can observe. The archaeological record is an integrated part of this present reality, and it remains part of this reality no matter how intensely we observe it. The archaeological data belong to the present, and there is no way we can claim that we observe a past reality by observing archaeological data. There is no logical link that takes us from the archaeological record back into the past, and there is no way we can draw conclusions by rules of logic from the archaeological record to the nature and organisation of past societies. Even though the archaeological data can be observed in their present reality only, they are the key to knowledge of a past reality. In this connection it is an axiom to us that what we separate as archaeological data once were elements of a past human reality, and if only we had known this reality and had known the nature of all the successive transformations that shaped the archaeological record, we could have predicted this record in great detail. Thus our axiom allow us to claim that the archaeological record is structured by the past even if it is part of the present, and in consequence there must be a correspondence in structure between the past reality and the record in its present reality. This we may use as a guiding principle to evaluate propositions concerning the past, and indeed it is the only link we have to knowledge about the past reality.

In consequence, we may separate two obvious levels on which archaeologists work. One is the level of current reality, where we can observe, analyse and categorise with great precision. The other is the imaginary level of past human societies to which we ascribe qualities and causal relationships. The latter is as much a part of our current reality as the former, and the justification for claiming that our modelling on this level has relevance for the past depends on our ability to show that the structure of the propositions we put forward do not violate the structure of the archaeological record. It is worth noting that we can never prove a statement concerning the nature and organisation of past societies to be true beyond doubt. In simple cases we may feel very certain that our statements are right, even to the degree where we may be tempted to claim that we have drawn a logical conclusion from the archaeological record, but nevertheless it is only circumstantial evidence. For instance we have no witness to the thin butted flint axe being used as an axe for cutting down trees, and only a couple of hundred years ago these axes were considered to be remnants of strokes of lightning. To day we know "for certain" that they are axes, and this is a certainty that relies on a multitude of gathered informations, which viewed together in a coherent structure can be meaningful only if the thin butted axes were in fact axes (general form, shafted examples, find contexts, experimental use, chopped wood etc.) With more complex models and general explanations, however, we can never claim to be certain, and in my opinion these general statements are more a revelation of our current view upon present world realities, than they are statements of facts concerning past realities.

One important point should not be forgotten here. Although we cannot prove anything to be true, we certainly can falsify statements concerning the past. In theory, at least, we can outline the implications of a statement and compare these implications with the actual archaeological record. In simple matters this works quite well, as for instance when wear trace analyses of denticulated flint pieces with gloss from the Early Neolithic show that these cannot have been sickles as believed up till now (JENSEN (1989)). In connection with complex statements concerning the past, however, one may seriously doubt our ability to draw the right conclusions.

Butted in between these two work levels a third one appears. It is the process of linking assumptions concerning the past with actual data. It consists of statements and arguments of how the archaeological record was formed, with direct reference to a conscious model of the past society in question. The logic of this process runs from the model to the archaeological record. Yet it is not merely an intellectual exercise, as one might believe from deductive positivism. Empirical knowledge can and should indeed enter the linking arguments. Such a knowledge has definitely always been part of the linking process introduced through the "life experience" of archaeologists. However, realising the nature of the linkage it is far more profitable to rely on a formal study of present day formations of the "archaeological record". This type of formal studies in the shape of ethnoarchaeology and "action" archaeology has in recent years expanded to a degree, where one may truly speak of a third level of archaeological activity. Because of the empirical content of the linking argumentation, the linkage itself seldom appears as a deduction from our models, and there is no reason why it should. The linking argumentation may take any form we wish, as long as we are aware that we cannot make a link unless we have theories and models concerning the prehistoric past. In other words that the logic proceeds from these to the present day context; and in consequence that it is wasteful not to explicate theories and models in advance of an attempt to link.

4. The role of statistics in archaeology

Statistics, of course, are many things, but we may make a division into two major groupings — the inductive or inferential statistics, and the deductive or descriptive statistics. Taking this division as a starting point I would argue that the application of inductive statistics in archaeological research is very problematic. There are two main reasons for this, both of which stem from the nature of the archaeological record. Most inductive statistics require that we know in detail the distributional qualities of the populations to which we apply the inferences. At the same time, they require that we have complete control of the formation of the samples from which we infer. None of these requirements are met in archaeology. Whether we conceive of the "populations" as a material present in the past, or just as a material present in the earth today, we have to realise that the populations of archaeological material and their distributional qualities are basically unknown. Furthermore, if we speak of past populations, we have no way of knowing the exact history of the formation of the samples. The same, of course, does not necessarily apply if we speak of populations in terms of the hidden part of the archaeological record.

A more fundamental objection against inductive statistics in archaeology may, however, be raised. It is very doubtful whether the archaeological record can be considered to be a sample of anything at all in a statistical sense. That is, the archaeological finds and their composition cannot be viewed as samples that reveal what some larger background unit looked like. Each find - and each composition - springs from an actual historic event or sequence of events, and thus has a complete "as it is" meaning by itself. It is a unit of complete information. Even if we speak of excavation samples from, say, a large settlement site, we cannot claim that these are samples in a statistical sense, for the settlement site is not a population with some uniform theoretical structure. On the contrary, it is a statistically very haphazard phenomenon, and predicting what the rest will be like from an excavation "sample" in one part of the settlement is beyond statistics. It is of course possible to devise a sampling strategy that will reveal the structure of the settlement, but then we are not dealing with one, but many samples, and the statistics needed to reveal their information are not inductive, but deductive.

In sociology, there are no problems using inductive statistics, because it is possible to observe the populations and their qualities, and carefully define the extraction of samples in a way that makes it possible to use inference statistics in a meaningful way. As outlined above, the nature of archaeological research does not allow us to observe the original populations from which the archaeological record is extracted, nor can we follow – let alone define – the processes through which the record is extracted. Therefore, a use of inductive statistics in archaeology will for theoretical reasons be a misapplication, and I fear that in most cases it will also in practice lead to erroneous, or rather, nonsensical results.

Deductive statistics, in general, have no *a priori* assumptions that we cannot control. The specific methods do have limitations that impose restrictions on the data they analyse. However, these restrictions always apply directly to the observed input data. This leaves us in full control to use the methods properly. It is very important to keep in mind that deductive statistics are descriptive. They have no inferential value whatsoever. We can use them to clarify the contents of the archaeological record, and present it in a form that makes it easier for us to carry out our linking argumentation. They cannot in any way produce the conclusions for us.

5. The Medusa mirror

Among the deductive techniques, multivariate statistics forms a powerful group that although difficult to understand and hence difficult to use properly, has great potentials for archaeological research. Multivariate methods like Principal Component Analysis (PCA), Principal Coordinate Analysis, Correspondence Analysis (CA) and Discriminant Analysis, being data reducing, structure retaining methods have the capability of reflecting the essence of the data we submit to them. Provided that we do not feed them data that violate basic restrictions (like having relationships among variables in a PCA, where the correlation coefficient is nonsensical) then they will reflect the essence of the data back to us in some sense.

A point to bear in mind, when we begin to work with multivariate analyses, is the difference between the archaeological record and the description of the archaeological record. This of course is a distinction fundamental to archaeological research in general, regardless of whether multivariate analyses are used or not, but it is a problem that becomes more present when these techniques are adapted. In archaeology there has been a clear tendency to view a description of a set of data as an objective representation of these data. A material, comprehensively described in a publication, is often considered to be objectively documented for the use of other researchers. The descriptive publication becomes an aim in itself. Only when strongly formalised analytical techniques like multivariate analyses are brought to bear, does one realise how important that small, formal difference between data and the description of data really is. We should never forget that it is descriptions of data we analyse, and not the data themselves. Data are not God given. They are what the researcher has chosen, and carefully chosen one may hope, to illuminate a specific archaeological problem. And not only are they data chosen for a specific problem, they are also described for that specific problem to be solved. The success, or more often the lack of success, in choosing and describing data is mirrored back to us through the multivariate techniques. This is why I like to refer to the multivariate techniques as Medusa mirrors — they mirror back to us the ugliness, the failure, of our choices and our descriptions.

Let me give an example that I find very illustrating. In 1986 Per ETHELBERG published a 4th and 5th century Iron Age cemetery from Hjemsted in southern Jutland. In connection with this publication he also presented an analysis of the pottery in the graves with respect to both form and decoration, and with a chronological division as the objective. To analyse the pottery decoration he used a CA on presence/absence registrations of 45 attributes on 63 vessels from the graves. The first two axes of the analysis shows some rather clear groupings (Fig. 2). First of all, there is a clear main division along the first axis, with minor sub-divisions along the same axis to the right in the diagram, and along the second axis to the left in the diagram. Based on various types of datable material like brooches in the graves the main grouping to the left can be said to represent a set of

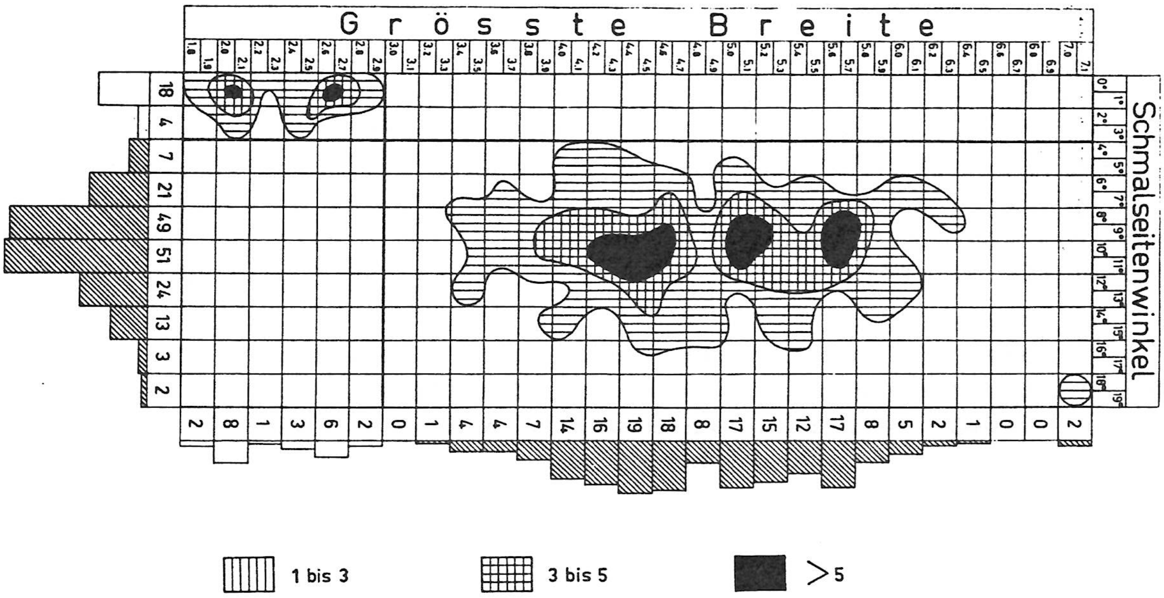

Fig. 1: Combined histogram, scattergram and isaritme diagram.
From MALMER (1962), Abb. 83.

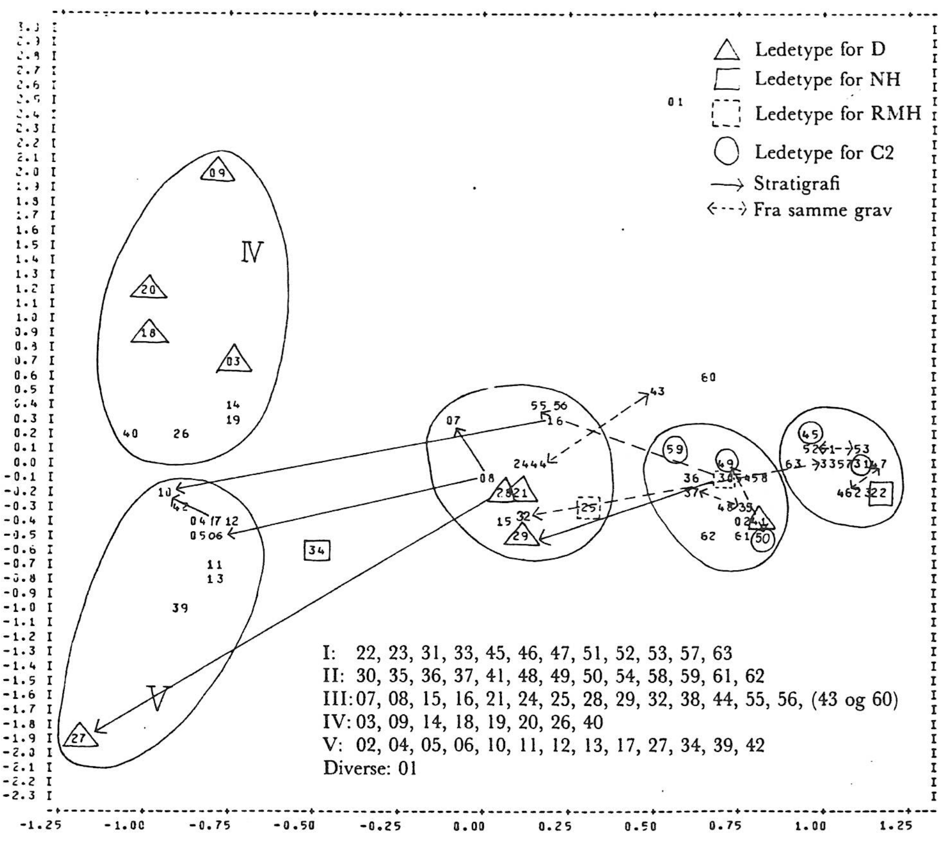

Fig. 2: CA of pottery (63 vessels) from the Hjemsted cemetery (first two principal axes).
From ETHELBERG (1986), Fig. 39.

clearly young pots in the assemblage, whereas the grouping to the right represent more of a mixture with predominantly older pots, but also with some younger ones.

Per ETHELBERG had hoped for and expected a time seriation to emerge from the analysis, but clearly this was far from being the result. As I stated in an appendix to his publication (MADSEN 1986) and on later occasions (MADSEN (1988, 1989)), the demands for a set of data to qualify for a seriation is the presence of a systematic, monotonous change of variables across units. In a CA this requirement is met when a parabola shaped layout is found in a plot of (primarily) the first two axes. Such a distribution was clearly not present, and even though the groupings that were the result of the analysis could be chronological significant, this had to be rendered probable by external arguments. The analysis itself did not support a chronological division.

When ETHELBERG's book appeared, one of my colleagues at Moesgd, Jytte RINGTVED, immediately reacted. She felt that the way P. ETHELBERG had described the pottery was mistaken if the objective was to isolate chronological aspects. The variables that he used would tend to emphasize attributes of decoration determined by form, rather than isolate style variation determined by time. And although an older and a younger group of pottery was separated through his description, the chronological potential of the pottery was far from utilised. Recently, RINGTVED has published these objections (1987), and at the same time treated the pottery from the Hjemsted cemetery anew. This was done in connection with a major study of the pottery from Late Roman and Early Germanic Iron Age in Jutland. Her description of the pottery is simple and selective, and she uses less than half the number of attributes that P. ETHELBERG did. Yet her results using CA are impressively clear. She obtains exactly what she wished to obtain — a clear-cut seriation of the pottery. It emerges in the plot of the first two principal axes as a neat parabola shaped formation. She also clearly demonstrates that it is her description that makes the difference to ETHELBERG's analysis — not the larger size of the material. She obtains this result by a separate analysis of the same pots that ETHELBERG used, resulting in a neat parabola shaped formation as well (Fig. 3).

The point made by this example is that the difference between success and failure lies with the choices and descriptions we make. If we want to seriate a set of data and the result we obtain from the analysis is not a seriation, then we should not automatically blame the data nor the analysis. First of all we should investigate our descriptions, as the reason for failure most probably lies here. Incidentally, this is also why many archaeologist today have become hesitant to voluminous descriptions of data. Why spend time and valuable pages on endless descriptions of data, if descriptions in general turn out to have value only if they are designed for specific purposes.

The descriptive multivariate techniques have excellent capabilities of reflecting structure in our descriptions of data. None of them, however, are robust in the sense that you can just pour in descriptions in one end and have whatever structure that lies hidden in the descriptions reflected back to you. Each technique incorporates a set of basic rules that govern what is to be considered as structure and how different informations are to be weighted against each other. To be able to use the techniques correctly, it is necessary to understand the rules that apply to each technique, partly in order to avoid direct mistakes, and partly in order to be able to optimize the use of the techniques. It is also important to know what happens with input data for a particular analysis, otherwise one would often be at a loss to understand the output data properly. For instance, it is important to know that the CA include weighting mechanisms that ascribe influence differently to units and variables, according to their size, and to the skewness of their distribution in the data set.

In the following I shall give an example of how a concern with the description of data seen in relation to the internal rules of the PCA may improve the outcome of its use. PCA is based on linear regression between measurement variables with a starting point in a correlation or covariance coefficient matrix. This means that variables with a high degree of correlation will dominate the results. Choices of variables and the way they are measured are therefore of extreme importance. Archaeologists are traditional with respect to measuring, as they normally take a series of measurements of length and width on various locations of the artefacts. Although this is a straightforward and legitimate way of

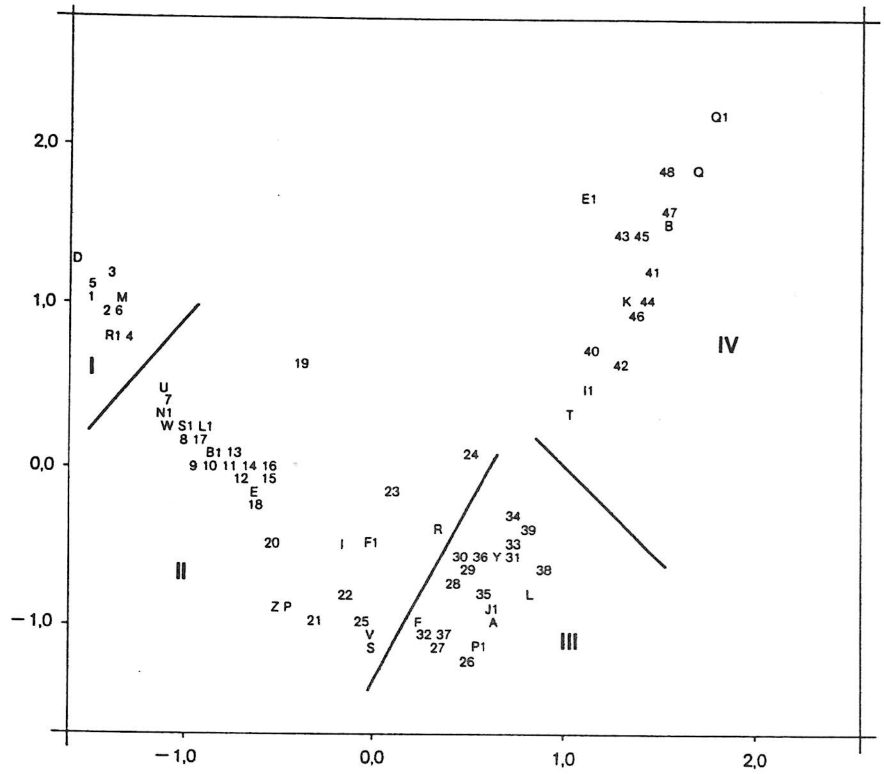

Fig. 3: CA of pottery from the Hjemsted cemetery (first two principal axes). From RINGTVED (1989): note 33.

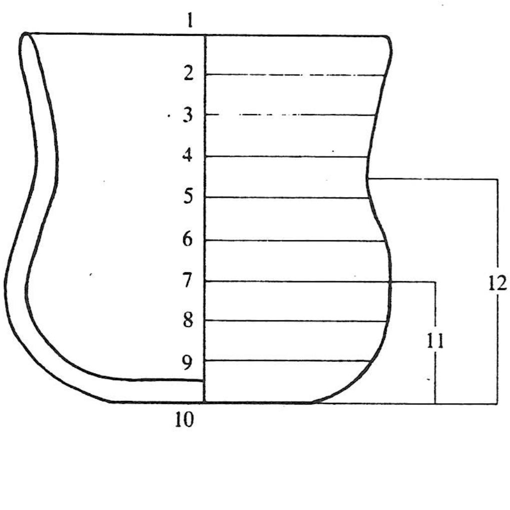

Fig. 4: The 12 measurements used to describe the Bell Beaker pottery. From SHENNAN (1988), Fig 13.17.

measuring, it often causes problems if a PCA is applied directly to these measurements. The problem is that the dominating relation between measurements of this kind is provided by their sizes. This relation is thus bound to dominate the results of the PCA completely, and may make other and more interesting relations less obvious.

On an earlier occasion I have discussed and demonstrated this fact in connection with analyses of Early Neolithic pottery forms (MADSEN (1988)). I shall demonstrate anew how vital the way measures are taken is to the result. This time I shall use data taken from SHENNAN (1988). SHENNAN's work is concerned with variation in form within a set of pots from the Bell Beaker Culture. The form of the pots is described through 12 measurements (Fig. 4). Ten of these were taken horizontally with fixed intervals from rim to bottom. In addition two vertical measurements were taken. To avoid that the size of the pots should become too dominant, all measurements were standardised using the heights of the pots. A PCA gave the result for the first and second axis shown in Fig. 5, and a closer investigation immediately reveals that the size factor has not been eliminated, as Shennan thought. The numbering in the plot is one that I have made based on a sorting by size on the sum of all measurements for each pot. Pot 1 is thus the smallest and pot 22 the largest. As can be seen the first axis reflects size almost perfectly.

The reason why SHENNAN did not succeed in eliminating size is that he used only the height of the pots for normalisation, and not simultaneously the width. He has normalised both height and width measurements with pot height, where he ought to have normalised height measurements with pot height and width measurements with say rim width. Unfortunately the original measurements were not published, and hence a correct analysis cannot be obtained, but an idea of the potential can be given. I have made an analysis where only the width measurements are used and where these, despite their previous normalisation with the height of the pots, have been normalised with the width of the rims (Fig. 6). An improved result with tendency for groupings can be seen, but how much better the results would have been with the original data can only been guessed.

It is important to realize that the "results" gained by the use of statistics and especially by the use of multivariate statistics are reflections of structure, and nothing else. Even if a result shows a clear and easy recognisable structure, it is not automatically an interpretation as well. Irrespective of whether a clear-cut structure is gained or not and whether an expected structure is obtained, then the interpretation as a whole rest on the understanding of the significance of the described attributes, and not on the structure in its own right. A good example of this is Flemming HØJLUND's analyses of Bronze Age pottery from Failaka in the Arabian Gulf (1988). The pottery stems from excavations of two contemporary, yet very different tells on Failaka. The excavations of tell F3 revealed small-scale architectural units, which were probably living quarters for merchants and their families with moderate facilities for production and storage. The nearby tell F6, on the other hand, contained a large building, called the Palace. The size and complexity of this structure indicates a type of social organisation very different from the one found in the family-based units of F3.

The chronology of the two tells is based on a study of the pottery, but naturally, the stratigraphy of the tells play an important role also. The same is true with the historical chronology of Mesopotamia, to which a correlation could be established by way of stamp and cylinder seals as well as pottery parallels. HØJLUND's work with Failaka was almost completed, when a CA of the pottery data was run for the first time. As this had been worked through for a very long time prior to the CA, and with the explicit aim of isolating chronologically sensitive attributes, a seriation was indeed expected. An analysis of the F3 data alone resulted in a nice parabola, with an internal order of units and variables that agreed quite well with the ideas that HØJLUND had of the chronological development (Fig. 7). The details of this parabola could provoke discussions on the finer details of the chronology established by HØJLUND, but as a hole there was a clear accordance. When data from F3 and F6 were analysed together, however, problems arose. The CA gave a neat parabola here as well, and an interpretation of the result as a time seriation could seem quite straight forward (Fig. 8). However, it was not as simple as that. The sequence of units and variables did not fit HØJLUND's chronology any more. Phase 3B from F6 had ended up after phase 3B from F3, phase 4A and 4B from F3 had

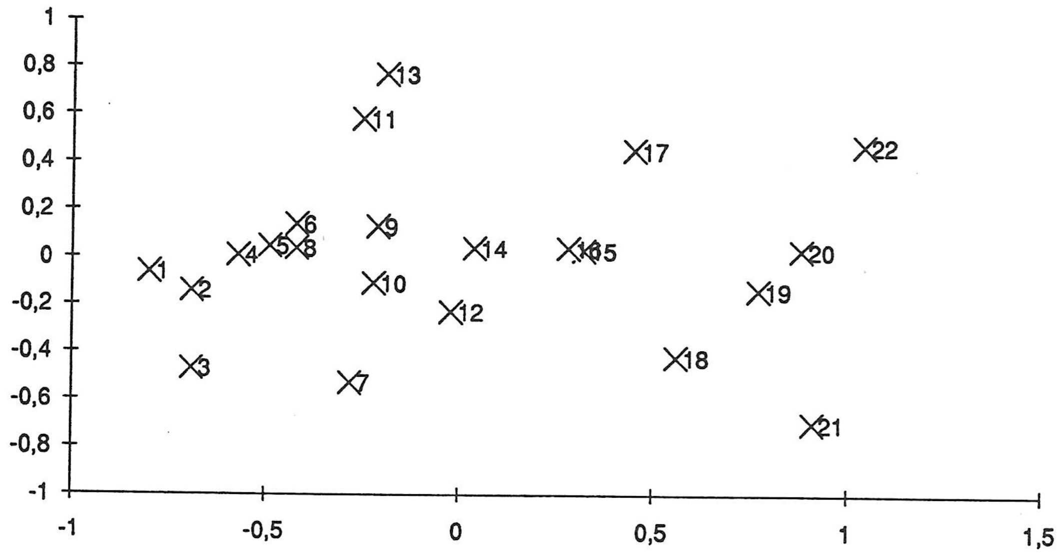

Fig. 5: PCA of measurements on 12 pots of Bell Beaker pottery. Numbers indicate a rank order by size. Original data from SHENNAN (1988), Table 13.12.

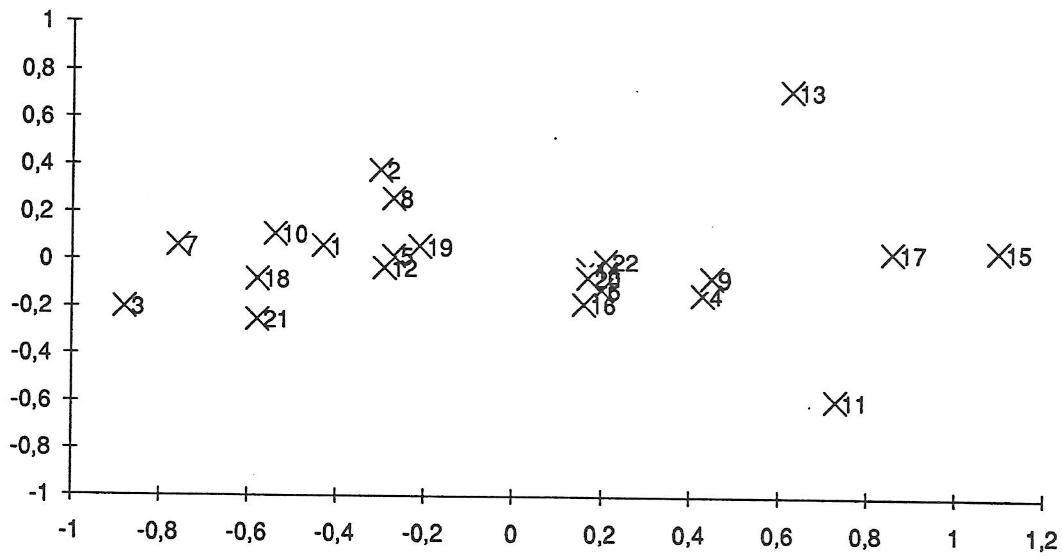

Fig. 6: PCA of measurements on Beaker pottery. Numbers indicate a rank order by size. Same data as in Fig. 5, but with measurements normalised by rim width.

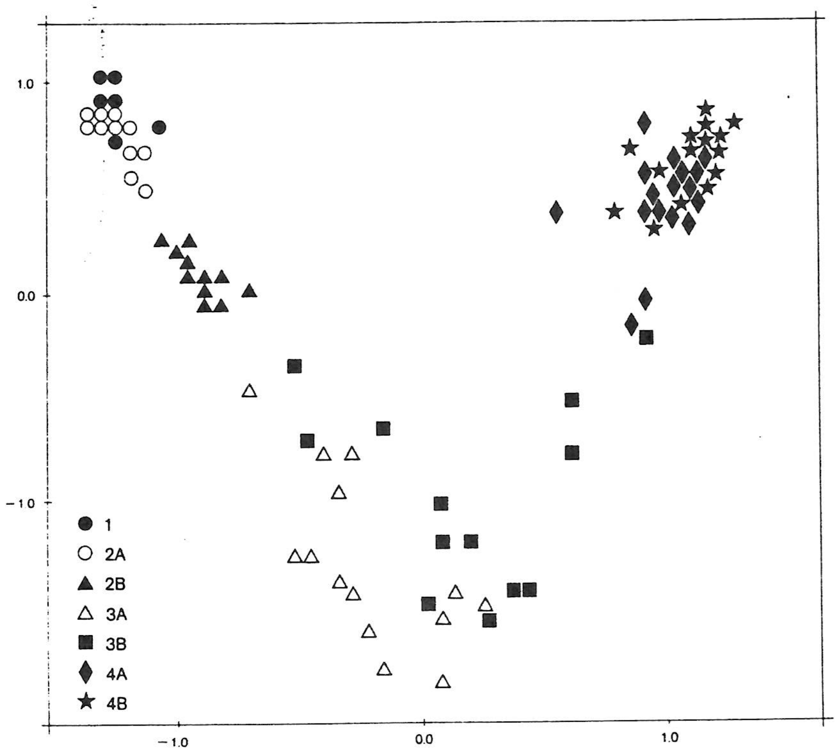

Fig. 7: CA of Bronze Age pottery from Tell F3, Failaka. From HØJLUND (1988), Fig. 1.

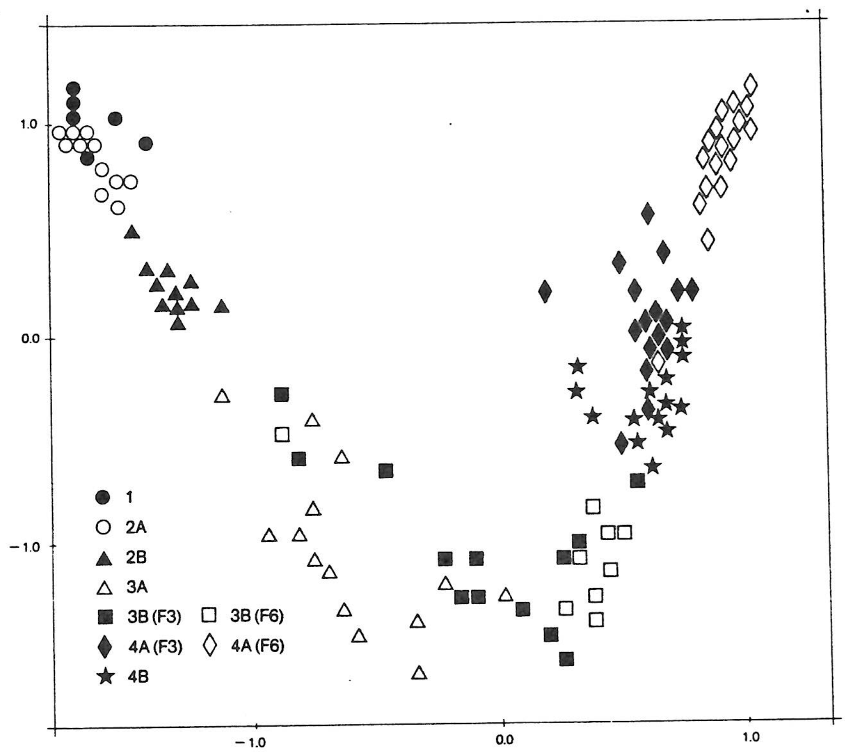

Fig. 8: CA of Bronze Age pottery from Tell F3 and F6, Failaka. From HØJLUND (1988), Fig. 4.

switched position in the sequence, and phase 4A from F6 had ended up last. Clearly, the structure was one we associate with a time seriation, but it was no time seriation, even though an analysis of F3 based on the same basic typology had resulted in a clear-cut time seriation. What had gone wrong?

The explanation given by HØJLUND is quite thought provoking. The Bronze Age pottery on Failaka is dominated by two major traditions. One is a locally developed tradition that dominated from the outset. The other is a Mesopotamian tradition that increase its influence over time, and becomes quite dominant towards the end of t he sequence. A closer examination of the data shows that it is this development with one tradition replacing another one which forms the basis of the successful seriation of F3. However, it is a paradox that it is the exact same development that causes the flaw, when F3 and F6 are analysed together. The crucial point is that the Mesopotamian tradition seems to have been finer and socially higher ranked than the local tradition, and this fact influenced the composition of material on the two socially different localities, F3 and F6. In the palace of F6 the Mesopotamian tradition predominated phase by phase compared to the contemporary layers at F3, where the local tradition persisted. This social difference means that the sequence obtained from a CA is not a true time seriation, but a mixture of a time and a social phenomenon. The result of the analysis gives no clue to this by itself. Only a knowledge of the data described and analysed can lead to the correct interpretation of the analysis.

6. Conclusion

The use of statistics in archaeology has grown considerably over the last few years. It is a development that I can only applaud. In my opinion archaeology can gain a lot of valuable support from statistics and especially from the group of multivariate techniques. What in my opinion makes statistics a valuable tool for archaeology is the ability to uncover and describe structure in data. The structures uncovered are essential to us, if we are to make efficient linking between assumptions concerning the past and the actual data.

However, it is not without problems to use statistics, and especially not multivariate statistics. It is a *must* that we know what the various techniques do in order for us to describe data properly, and in order for us to understand the output from the analyses. The advent of the personal computer has given us the possibility to use increasingly complex statistical analyses. This development increases the demands for the education of archaeologist in statistics. Not to make the archaeologist a statistician, but in order for the archaeologist to be able to understand what happens when a particular analysis is applied to a set of data. If we do not reach a point of better understanding, we will be inundated by bad an ugly statistics.

I do not intend to suggest that archaeologist should use advanced statistics no matter what. Archaeology has thrived excellently without and may continue to do so. Archaeologists who do not feel at home using statistics should not throw themselves to the lions. On the other hand I am not encouraging archaeologists to believe that they can manage as well or better without statistics. It is my firm belief that an archaeologist who has spent time to understand the use of statistical methods is far better off than one who has not.

References

DORAN, J.E. & HODSON, F.R. (1975): Mathematics and Computers in Archaeology. Edinburgh University Press, Edinburgh.

ETHELBERG, PER (1986): Hjemsted - en gravplads fra 4. & 5. årh. e. Kr. Skrifter fra Museumsrådet for Sønderjyllands amt 2,

HADERSLEV GLOB, P.V. (1945): Studier over den jyske Enkeltgravskultur. Aarbøger for Nordisk Oldkyndighed og Historie 1944: 1-283.

HEDEAGER, LOTTE (1978): Processes towards State Formation in Early Iron Age Denmark. In: KRISTIAN KRISTIANSEN & CARSTEN PALUDAN-MEER (eds.): New Directions in Scandinavian Archaeology. Studies in Scandinavian Prehistory and Early History Volume 1. The National Museum, Denmark. pp. 217-223

HØJLUND, FLEMMING (1988): Chronological and functional differences in Arabian Bronze Age pottery. A case study in correspondence analysis. In: TORSTEN MADSEN (ed.): Multivariate Archaeology. Numerical Approaches in Scandinavian Archaeology. Jutland Archaeological Society Publications XXI, Aarhus.

JENSEN, HELLE JUEL (1989): Plant Harvesting and Processing with Flint Implements in the Danish Stone Age. A View from the Microscope. Acta Archaeologica Vol. 59 - 1988: 131-142.

MADSEN, TORSTEN (1985): Numerisk dataanalyse for arkæologer. Institut for forhistorisk arkæologi ved Århus Universitet.

MADSEN, TORSTEN (1986): EDB analyserne. Appendix B i Per Ethelberg Hjemsted - en gravplads fra 4. & 5. årh. e. Kr. Skrifter fra Museumsrådet for Søerjyllands amt 2, Haderslev.

MADSEN, TORSTEN (1988): Multivariate statistics and archaeology. In: TORSTEN MADSEN (ed.): Multivariate Archaeology. Numerical Approaches in Scandinavian Archaeology. Jutland Archaeological Society Publications XXI, Aarhus.

MADSEN, TORSTEN (1989): Seriation - en grundlæggende arkæologisk arbejdsmetode. KARK Nyhedsbrev 1989 Nr. 3.

MALMER, M. P. (1962): Jungneolithische Studien. Acta Archaeologica Lundensia. Papers of the Lunds Universitets Historiska Museum Series in 8°. N° 2, Lund.

RINGTVED, JYTTE (1989): Jyske gravfund fra yngre romertid og ældre germanertid. Tendenser i samfundsudviklingen. KUML 1986.

SARAUW, GEORG F. L. (1903): En Stenalders Boplads i Maglemose ved Mullerup, sammenholdt med beslægtede Fund. Bidrag til Belysning af Nystenalderens Begyndelse i Norden. Aarbøger for Nordisk Oldkyndighed og Historie 1903 II Række 18 Bind: 148-315.

SHENNAN, STEPHEN (1988): Quantifying Archaeology. Edinburgh University Press, Edinburgh.

THOMAS, DAVID HURST (1978): The awful truth about statistics in Archaeology. American Antiquity Vol. 43,2: 231-244.