Espen Uleberg

Scale Dependent Patterns in Large Museum Datasets

Abstract: The large dataset made available by the Museum of Cultural History, University of Oslo, is used to present broad patterns in the geographic distribution of all Stone Age finds from the museum. A set of metadata is introduced to describe the precision and accuracy of the geographic information. To incorporate most of the finds, the first presentation is done at the level of municipality. In a second analysis, only finds with more precise location are used, and Mesolithic/Early Neolithic and Late Neolithic sites are separated. This shows that a change in the distribution pattern from these large museum databases can be good starting points for analyses, a place to get new ideas and to see whether a hypothesis might be worth pursuing.

Norwegian Archaeological Databases

Large datasets that can show scale dependent patterns are becoming more readily available at the Norwegian university museums. Information on practically all objects becomes available on the net, and it can be expected that this will in many ways change the way research is done. It is easier to get an overview and take up research questions that cover large regions with large amounts of material. The database system and the datasets are the results of two consecutive national projects that ran from 1991 till 2006. In the early phases of the projects, an emphasis was placed on converting paper based texts to computer readable formats. After transcription, all text from the catalogues was tagged using SGML with a grammar developed for these texts. (Holmen / Uleberg 1996; Holmen / Ore / Eide 2004; Eide / Holmen 2006). The tagged texts were later converted to databases. The database system is compatible with CIDOC CRM (CIDOC), and is developed to be used for research purposes and not only for the daily artefact curation at the museums.

From the beginning of 2007, it is possible to search for objects in the databases of all the archaeological university museums simultaneously. This is a unique possibility, but at the same time it highlights the need for normalization. The terms used in the written catalogue are kept, but normalization has been done to facilitate better searches in the database. Modern terms have also been applied for parts of the material. In some cases, this can be done simply by replacing the old terms with modern ones. In other cases, it is necessary to examine the objects and reclassify them.

The artefact terminology changes not only with time, but also from museum to museum. Since each museum mainly curates artefacts from its own district, it is natural that not all types are present in all collections. Nevertheless, a search in the database can show very different results from Oslo, Bergen, Trondheim and Tromsø that may be more due to different traditions of cataloguing practices and terminology than prehistoric realities. It is a vital point that the researcher knows how the data is organized and what terms are used in the dataset. A series of differently structured searches in the base will make the researcher familiar with the data, and give a base for evaluating the results.

The museums, as owners of the data, want to provide the best possible dataset. At the same time they should not wait too long to give broad audiences access to the material. Systematic checking will eliminate most errors, but after a certain point, the quality improvement should be done through continuous use by many different users. Inconsistencies can best be detected and corrected when students and researchers with different background knowledge use the database. After all, good research presupposes a critical attitude to the sources whether they are paper based or electronic.

It is not only the artefact terminology but also the place names that are kept in their original form in the base. At the Museum of Cultural History, University of Oslo, we have added updated information of place names and coordinates. The starting point has been land registers from 1886, 1950 and 2000. By identifying and linking the farms mentioned in these registers, it has been possible to connect the place names used in the catalogues to the present land register. This is in turn the base for the distribution maps and will also make it possible to connect the museums data base with the Norwegian national register for sites and monuments, Askeladden.

Distribution Maps and Metadata

Maps showing artefacts and sites plotted in combination with other themes have always demanded a lot of time and effort. The distribution map has been part of a result, but can now to a larger extent be a starting point for archaeological research. When such distribution maps become available on the net, they can be downloaded and combined with other map elements. It will be possible to look for patterns in different scales, to get new ideas, and to see whether it seems meaningful to continue to elaborate on a certain hypothesis.

When using the information in different map scales, the coordinates must have metadata describing the level of precision and accuracy. The source of the coordinates, whether it is exact measurement in the field, measurement on paper map or a reading from a map on the internet, is recorded together with the original map scale, as it will indicate the precision. Artefacts that come to the museum from recent excavations can have very precise coordinates as modern techniques are used. Information about the provenance is more varied for items given to the museum by occasional finders. It is rather symptomatic that objects that have passed through a private collection have less information of their provenancethan any other group of objects. It is nevertheless important to treat the available information in such a way that it is possible to use all relevant objects in a spatial analysis.

Registration of metadata which provide information of the accuracy makes it possible to evaluate what kinds of analyses the objects can contribute to. I have introduced different accuracy classes. The best is when the exact coordinates are recorded at an excavation, then follows a class for objects from sites with known location. The next is when the artefact is related to a specific farm, and then follows a group of farms that share the same farm name. The next levels are parish, municipality and county. In each case, a representative point is chosen within each unit, e.g. near the farm houses, or at the church.

The metadata can also be used for scale dependent presentations on digital maps. Most finds, also those with low accuracy, can be included at a small scale level, while a zoom in will only leave finds with higher accuracy. A relatively imprecise location can give a good representation at certain scale levels, but be totally wrong when the map is more detailed. For a map showing the distribution of bronze swords in Norway, a correct presentation should include items with known municipality or parish. For a presentation of the same objects within one municipality, only finds with higher accuracy should be included.

It could have been possible to link the artefacts to farm polygons, but this does not turn out to be a good solution. The borders have changed over the years, and there is no reliable information available for historical farm borders. Starting in the 19th century, arable land was restructured so that each farm would have its property closer to the farmstead. Polygons would therefore in many cases only give a pretence of accuracy. Instead, a point near the farm houses will in most cases give a good approximation. The error will be greater for finds connected to activities like hunting, shieling and iron extraction sites, since these finds are quite often made at a larger distance from the farmsteads.

Comparing Numbers

Comparing numbers is an obvious thing to do when using a large data set as a starting point for the analysis. The question is how to count and compare across such a diversified material as a complete museum collection. Artefacts can be counted, but if one were to count each object, then how to compare a flint flake and a stave church portal or a group of sherds with a complete vessel? The availability of the data makes it easy to run through a lot of analyses and make comparisons that can be both meaningful and meaningless.

Catalogue numbers are easily counted. Before 1900 each object was allocated its own museum number. Since 1900, the main rule has been that a site, or a context within a site, is given a catalogue number and each object or group of objects a sequential letter or number within that context. Each of these entities is one entry in the database.

Sites can be counted. This is possible to do as far as one museum number is equivalent to one site but this is not always the case. However, concerning finds from the Stone Age, only one number is given to each site. This makes it possible to infer that each catalogue number can be seen as one or more activity areas. A map presenting the number of activity areas in each municipality should be a representation of the distribution of Stone Age settlement in Norway.

Norwegian Stone Age

The distribution of Stone Age sites in southeast Norway can be an example of how this large dataset can illustrate known patterns and also give rise to new questions. southeast Norway is chosen, because the collection in Oslo consists mainly of finds from this area. The rapid eustatic land rise after the Ice Age gives a good relative chronology around the Oslo fjord (Sørensen / Lie / Nybakken 1990). An interpretation of a hunter/gatherer Stone Age site in this landscape will often conclude that it is a coastal site, and then the age is given by the height above sea level.

From the end of the 1950s onwards, archaeological surveys and excavations became part of the work connected to the construction of dams for hydroelectric power in the mountains. This led to the discovery of large numbers of Stone Age sites. The general impression, that Stone Age sites were mainly connected to water, were of course maintained and strengthened as the surveys were mainly limited to the areas around the lakes. Flint found at the mountain sites are interpreted as indications of contact between the coast and the high mountains as flint is only found along the coast. On this background, different seasonal cycles between coast and inland have been proposed, often with autumn reindeer hunting sites in the mountains combined with winter, spring and summer coastal sites (e.g. Indrelid 1975; 1994; Mikkelsen 1989). A shift to transhumance in the mountains has been shown from the Late Neolithic/Bronze Age (Prescott 1995), and pollen indicating grazing from the Neolithic suggests that the mountains were used for shieling soon after domesticated animals were present in Norway (Moe / Indrelid / Kjos-Johanssen 1978).

Distribution Maps

Having this general description in mind, we will turn to some distribution maps of Stone Age sites and see how this general idea of the connection between coast, high mountains and Stone Age sites is reflected in the museum collection. The Stone Age sites are treated as one group because the museum catalogue is generally not more specific. When a more precise dating is made, it is more often Neolithic than Mesolithic. Neolithic finds tend to be dateable tools like daggers, axes and hatches, while dateable Mesolithic artefacts are seldom found. Stone Age sites often consist of flakes and debris that could be produced at any time during the Stone Age. Mesolithic sites are generally dated based on their height above sea level. Based on a curve for Skitrinnet (Sørensen / Lie / Nybakken 1990), close to Oslo, sites from the earliest Mesolithic period, 10000–9100 BP can be at 200–115 m above sea level and at the end of the Mesolithic coastal sites are at 50/42–38 m above sea level (Berg 1997).

Municipalities

This first analysis uses a total of 10,356 entries from the database. This is the museum numbers catalogued before 2000 dated to the Stone Age and covering the whole museum district. The number of finds is divided in three classes *(Fig. 1, left)*. There is one group up to 49, one intermediate from 50 to 149 and one with up to 495 objects in each municipality. The map corroborates the main idea that Stone Age sites are concentrated along the coast and in the high mountains. As for contact and routes between the coast and the mountains, there seems to be only one.

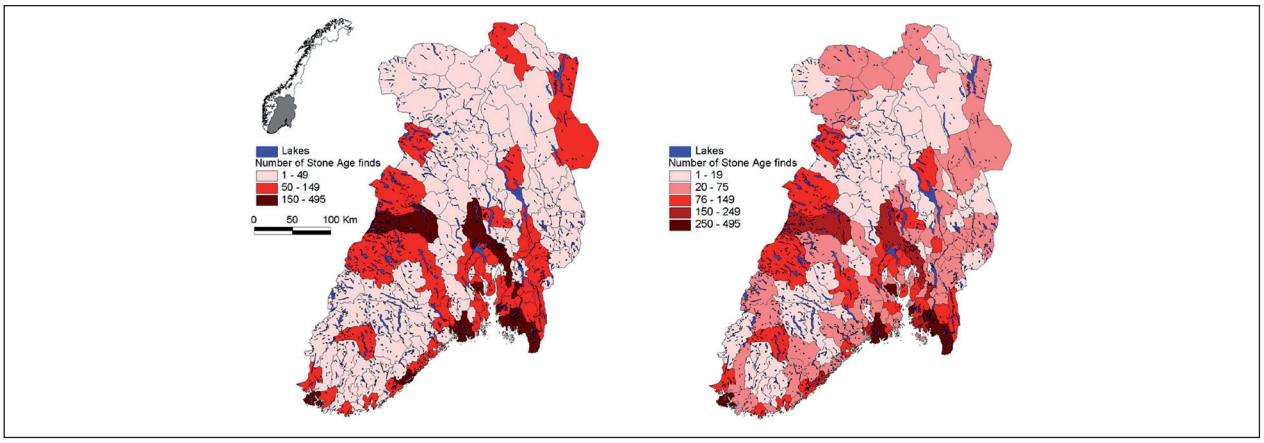

Fig. 1. Distribution of Stone Age finds in southeast Norway.

When the material is divided in five classes *(Fig. 1, right)*, the concentrations along the coast and in the mountains are still clear. At the same time, the map shows more inland areas with higher number of finds indicating several possible routes between coast and inland. It also shows that the sites are concentrated around lakes and also river systems. There is also a large area with medium number of finds around Oslo.

The map supports the general idea of site distribution, but also raises new questions concerning the lack of sites around some of the larger lakes. The low number of finds in two coastal areas is also interesting. One of them, on the south coast, is where one of the few submerged Mesolithic sites in Norway has been found (Sellevold/Skar 1999).

The maps show just the numbers for each area. From this starting point, the material could be normalized by square kilometres, acres of arable land or modern construction works in each municipality. A presentation of how the distribution has changed through time as more sites have become known would be interesting as research history. The number of finds is naturally not only connected to prehistoric activity but also to a large extent to activities in modern times. The distribution could show a real prehistoric distribution, or just reflect surveying intensity.

Farms

A more detailed map demands an accuracy of farm group or better. It is still broad patterns that can be seen. For this purpose, finds with an accuracy of a farm group or better can be used for the analysis, while those with less accurate information must be excluded. The map *(Fig. 2)* shows an area west of the Oslo fjord. 1146 find numbers are plotted on the map. Again, the overall impression is a concentration along modern rivers and lakes. The landscape is gently sloping, so the contour line 50 m above sea level goes far inland. This curve is roughly equivalent to the shore line 6300 BP (Sørensen / Lie / Nybakken 1990). When the sites are seen in relation to this level, even more of them can be interpreted as coastal sites. The sites lower than this line must be later than 6300 BP. A further elaboration would be to divide the material according to the different periods of the Stone Age and see the relationship to other prehistoric sea levels in more detail.

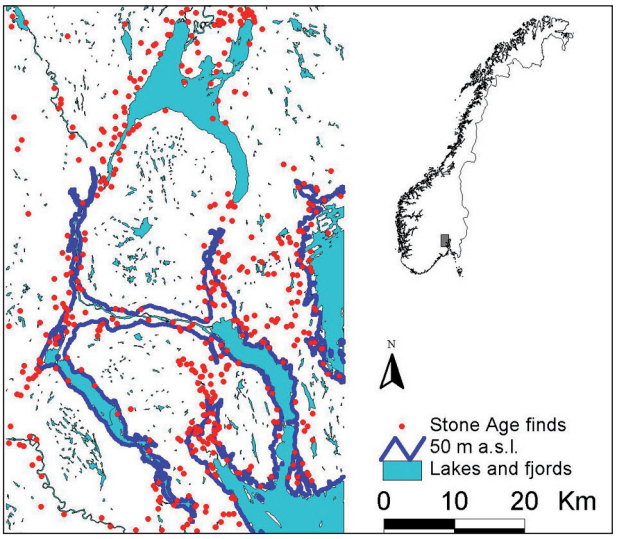

Fig. 2. Stone Age finds with an accuracy of farm group or better.

Precise Placement

An example of finds with more accurate placement is a group of Stone Age sites in the Lærdal Mountains. In these high mountains, it is possible to find sites as scatters on the surface, which give the impression that very little has happened and that the present reindeer hunt in the autumn is part of a continuous tradition (Johansen 1978). The larger sites are concentrated around the lakes, there are some concentrations around a few quarry sites, and in addition there are a number of smaller sites in the landscape. The larger sites can be described as attractors (Uleberg 2003). Seeing the larger sites as attractors make it possible to think of these sites in a way that transgresses the idea of time limited point allocations. They are areas with activity concentrations while other kinds of activity have left fewer traces around the sites and along the paths leading to and from them.

People with different subsistence economy will see and look for different elements in the landscape. People and landscape are intertwined. The landscape belongs to those who belong to it, with their experiences within the landscape. It is therefore hardly possible for us to understand why one specific location should be preferable to another, because we do not share the same experiences and the same way of looking at the landscape (Meløe 1989). The experiences made are dependent on the subsistence pattern. Hence, the subsistence pattern can be visible in the allocation pattern, and a shift in allocation pattern should reveal a shift from one type of economy to another. A hunter is looking for good places for the hunt, and needs an

understanding of the animal's movement in the landscape and will position himself according to this. As well as the hunter, the pastoralist needs an understanding of the animal's movement in the landscape, but the pastoralist is looking for favourable grazing grounds and places where it is possible to control and guard the animals. In addition, hunter/gatherers would go to the mountains in the

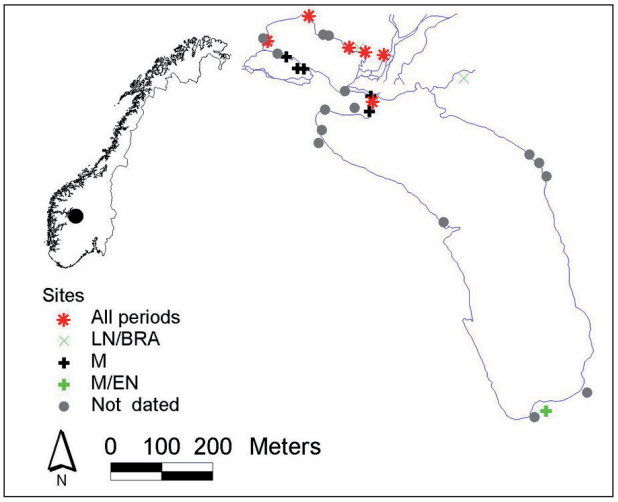

Fig. 3. Stone Age sites at Eldrevann.

autumn while pastoralists would stay in the mountains during late spring, summer and early autumn. The shift to a pastoralist economy that took place in the Late Neolithic could therefore be recognizable in the site allocation pattern.

The lake Eldrevann *(Fig. 3)* has the highest site density in the area. It was originally two lakes, with a height difference of only 20 cm, and separated by just a few metres of a small, shallow river. Tjørni had two outlets, creating a small island called Glitreøyni. The terrain leads the trail from the west coast up to this lake, creating one of the main routes from eastern to western Norway. A reindeer trail has also drawn hunters to the lake. The dateable artefacts cover the range from the Mesolithic to the Late Neolithic/Bronze Age, and the ^{14}C-dates are from 8500–2000 BP. There are also Iron Age finds and houses for modern shieling at the lake.

Artefacts were found all around Eldrevann except for a smaller part in the southwest, but only larger concentrations are marked as sites on the map. The dateable artefact concentrations were mainly in the northwest. Nine of the excavated sites could be dated. The sites can be grouped in areas around the lake. One area has only Mesolithic sites. That is the abovementioned small island, Glitreøyni, in the northwest. The earliest ^{14}C-date from the area, 8510 ± 110 BP, is from this group. The only dated site on the south shore is from the Mesolithic/Early Neolithic. One site at a distance from the lake is dated to Late Neolithic/Bronze Age while the rest have components from all these periods. It is also worth noticing that the Mesolithic/Early Neolithic sites are not close to buildings for modern shieling.

The distribution indicates that the first people in the area settled around most of the lake, since all sites have a Mesolithic component. The island in the northwest and one activity area in the south were used extensively in the Mesolithic period. People in later periods have chosen other areas, some of them correlate with modern shieling. This indicates that there has been a shift in landscape perception, and that the shift in subsistence activity in the Late Neolithic/Bronze Age is visible in the material. The distribution pattern also indicates that there have been no later shifts leading to similar changes in the exploitation of the mountain areas (Uleberg 2003).

Conclusion

I have now shown that large museum dataset can be used to present broad patterns. The large datasets are well suited to give an overview and to see broad outlines of the distribution of groups of objects. They can give new ideas that can be further studied, and it is an opportunity to see whether a hypothesis might be worth pursuing.

The data structure, which is CIDOC CRM compatible, can give access to both the original descriptions and results from later studies of the artefacts. One is nonetheless dependent on the information that was written into the database, and most of the time this will be the original catalogue supplied by normalization of the used terms.

The large datasets are very useful, but as always, students and researchers have to use several sources to get the necessary knowledge of the material they are interested in. The value of the database is as good as any paper-based source. One of the great benefits of the database is that it is easier to make search and validate results.

References

Aas / Faarlund 1995

B. Aas / T. Faarlund, Skoggrenseutviklingen i Norge. In: L. Selsing (ed.), Kilder for klimadata i Norden fortrinnsvis i perioden 1860–1993. AmS-Varia 24 (Stavanger 1995) 89–100.

Berg 1997

E. Berg, Mesolittiske boplasser ved Årungen i Ås og Frogn, Akershus. Varia 44 (Oslo 1997).

CIDOC CRM

CIDOC Conceptual Reference Model. Proposed ISO 21127. http://cidoc.ics.forth.gr/

Holmen / Uleberg 1996

J. Holmen / E. Uleberg, Getting the most out of it: SGML-encoding of archaeological texts. Paper at the IAAC'96, Iasi, Romania. http://www.dokpro.uio.no/engelsk/text/getting_most_out_of_it.html [31 Dec 2007].

Holmen / Ore / Eide 2004

J. Holmen / C. E. Ore / O. Eide, Documenting Two Histories at Once: Digging into Archaeology. In: K. F. Ausserer / W. Börner / M. Goriany / L. Karlhuber-Vöckl (eds.), Enter the Past, The E-way into Four Dimensions of Cultural Heritage. Computer Applications and Quantitative Methods in Archaeology 2003. BAR International Series 1227 (Oxford 2004) 221–224.

Indrelid 1975

S. Indrelid, Problems Related to the Early Mesolithic Settlement of Southern Norway. Norwegian Archaeological Review 8:1, 1975, 1–18.

Indrelid 1994

S. Indrelid, Fangstfolk og bønder i fjellet. Bidrag til Hardangerviddas førhistorie 8500–2500 før nåtid. Universitetets Oldsaksamlings Skrifter Ny rekke 17 (Oslo 1994).

Johansen 1978

A. B. Johansen, Høyfjellsfunn ved Lærdalsvassdraget (Oslo 1978).

Mikkelsen 1989

E. Mikkelsen, Fra jeger til bonde. Utviklingen av jordbrukssamfunn i Telemark i steinalder og bronsealder. Universitetets Oldsaksamlings Skrifter Ny rekke 11 (Oslo 1989).

Moe / Indrelid / Kjos-Johanssen 1978

D. Moe / S. Indrelid / O. Kjos-Johanssen, Environment and Early Man. Norwegian Archaeological Review 9:1, 1978, 32–36.

Prescott 1995

C. Prescott, Aspects of Early Pastoralism in Sogn, Norway. Acta Archaeologica 66, 1995, 163–189.

Sellevold / Skar 1999

B. Sellevold / B. Skar, The First Lady of Norway. In: G. Gundhus / E. Seip / E. Ulriksen (eds.), NIKU 1994–1999. Kulturminneforskningens mangfold. Temahefte 31 (Oslo 1999) 6–11.

Sørensen / Lie / Nybakken 1990

R. Sørensen / K.T. Lie / S.E. Nybakken, Drøbak 1814 II, kvartærgeologisk kart, M 1:50000 (Trondheim 1990).

Uleberg 2003

E. Uleberg, Settlement Patterns and Landscape Perception in Norwegian High Mountains in the Stone Age. In: T. Tsonev / E. Montagnari Kokelj (eds.), The Humanized Mineral World: Towards Social and Symbolic Evaluation of Prehistoric Technologies in South Eastern Europe. Proceedings of the ESF workshop. ERAUL 103 (Sofia 2003) 83–87.

Illustration Credits

All maps based on data from Statens kartverk (MAD 12002).

Espen Uleberg

Museum of Cultural History
University of Oslo
Frederiks gate 2
0164 Oslo, Norway
espen.uleberg@khm.uio.no