# STATISTICAL ANALYSIS OF THE INVENTORY OF LAYERS DESCRIPTIONS FROM THE ARCHAEOLOGICAL SITE OF OSLOGATE 6 IN OSLO

Dorota Stabrowska, Jacek Choloniewski and Kazimierz Lewartowski

Stratigraphical excavations on multi-layers archaeological sites imply a large number of data that has to be stored and analysed. The data cover information both on stratigraphy and on the strata themselves. Various models of documentation were designed to record the description of cultural layers. The inventory of layers used on the Medieval site of Oslogate 6 in Oslo (excavations conducted by Mr Peter Molaug), consisted of separate cards containing usual information: stratigraphical context, levellings, schematic plans of layers. In season 1988 was added a detailed description of layers. The design of the inventory made by Mr Andrzej Golembnik was based on archaeological experience and field practice on multi-layers sites, where the layers contained many mineral and organical components. The stratigraphic part of the inventory was filled in by the archaeologists working in trenches. The description of each layer was made by the same archaeologist for the whole site. It was done on the basis of a visual examination of the layer in situ as well as a careful analysis of a sample which had been rinsed on a four-level sieve. The analysis of the descriptive part will be the subject of our paper.

Parameters describing each layer are divided into 3 main groups (Il.1). The card begins with a label containing the inventory number, the Munsell color codes and shade. In a first group 13 physical features are characterized by means of a 5-points scale (0..4). The contents of the layers are described by 37 components of: mineral, animal and botanical origin as well as excrement and humus. Every component was marked as a percentage of the whole content, so the total of all entries in this group should give 100%. Additionally, the state of preservation of some of the components was noted. The card ended with an attribution of the given (described) layer to one of six possible categories of interpretation. These were written down as separate classes:

class A denote - natural layer;
class B - layer of levelling;
class C - building layer;
class D - habitation layer;
class E - destruction layer;
class F - other cases.

Each description was completed with an explanation of whether the layer was found "in situ", "spread" or "redeposited". Thus the assignment of each layer to one of the six predefined classes was the final step of the procedure of the analysis and description. The inventory of layers which after excavations was stored as a computer database contains a lot of informations in a form of recorded numbers. As a document of archaeological activities, it contains the record of numerous observations made in the field which can be of great help as a support of the author's memory when preparing the final publication.

Such a way of describing the layers allow for drawing a comparison between layers. Cards were filled by archaeologists according to their observations and intuitive interpretations but nothing was measured nor weighted. Though the classifications were judged by eye the descriptions should reflect relations between several components.

To uncover those relations hidden behind amounts of numbers one has to apply statistical methods. Our aim was to check if the Oslo inventory really contained anything more than observations and intuitive interpretations. Because the inventory was made by two archaeologists (A and B) working successively on the site, we could study also differences in methods of work of the two.

703 layers were recorded in the season 1988 (330 by archaeologist A and 373 by B). The numbers of layers by classes for the archaeologists, (first for A and then for B), are as follows: for class A - 4 and 0, for class B 104 and 63, for class C 55 and 22, for class D 152 and 278, for class E 11 and 10, for class F 4 and 0.

We have computed the average value and standard deviation for every parameter in order to discover characteristic features shared by the layers belonging to one class and also the attributes distinguishing the classes.

Since the archaeologists have estimated the degree of intensity of specific physical features, it would be expected that each class of layers should be determined by certain composition in the pattern of values.

Examination of computed average values of all physical parameters showed that they were obviously clear and distinct in the case of destruction layers (class E), especially for **burning**, **scorching** and **elasticity**. This observation was valid for both archaeologists. **Inclination** and **accumulation** took the highest values for all classes. For that reason the latter parameters could not been used as a tool for distinguishing between the layers. The chi-square test made in order to measure the association between the classes and specific values of physical features, gave generally negative results. Only the most numerous classes could be taken into account and the test indicated stronger association of **inclination** and **accumulation** with three classes: levelling, construction and habitation layers.

As in the case of the examination of physical features, the analysis of the components did not give satisfactory results, since all layers were built of the same main components: **humus**, **sand**, **clay** and **chips**. The only differences raised from the fact that they were present in various proportions and connected with divers minor components.

The levelling layers (class B) which are build mainly of **chips** and **humus** have noticeable portion of **excrement**, **sand** and **clay**. Archaeologist A recorded more **stones** and **straw**.

The components of the building layers (class C) are more differentiated, although here **chips**, **humus** and **clay** are also the most significant. Values of organic components are higher in the records of archaeologist A while archaeologist B noted larger portion of clay with distinct inclusion of **charcoal**. A surprisingly high mean of **shell** for A's layers is incidental, owing its value to large number of shells found in one layer of the C class. In the habitation layers (class D) the proportions of every basic components (**chips**, **humus** and **animal excrement**) is on the same level. This class is characterized by larger amounts of **nut** and **moss**. Significant addition of **straw** in the records of archaeologist A corresponds with similar amount of **straw** and **grass** in archaeologist B description.

Large differences between the two archaeologists occurred in records of the destruction layers (class E): A noticed mainly **stones**, **sand** and **clay** along with organic components, while B described the class as built of mineral components with predominance of **charcoal** and **lime**. High mean of **other botanic** components (A) was caused by one record again.

Since only one archaeologist assigned some of the layers to the A and F classes and in both cases a very few of them (4+4), it is difficult to draw any conclusion concerning these two classes. Natural layers (class A) was made of **sand** and **humus** which is what can be expected. The high mean of **lime** in class F was caused by one record only.

There was a constant difference between the two archaeologists irrespective of the class: in almost all cases the standard deviation of archaeologist A records was noticeably higher. It probably reflects mental distinctions between them, which influence the consistency and precision of the estimations.
We have also computed all mutual correlation coefficients. These coefficients are close to zero in most cases (only in a few cases they exceed 0.5) which suggests, that the parameters are statistically independent. The number of high coefficients decreases with the increasing number of layers indicating that they are mainly coincidental. One expected that **burning** and **scorching** should give a high correlation coefficient, but that was true only in class B for archaeologist A and in class D for both of them. Another expectation concerned a probable high negative coefficient of the pair **lamination** and **granularity**: that proved true for one of the archaeologists only (classes B, D, E). Most probably the other archaeologist did not obey the same rule as the first one, or was more inconsequent with his recording system.

The analysis of means proved that in some instances the classes of layers could be characterized by average values of certain parameters. The result concerns whole groups of layers and does not give indications for assignment of an individual layer to the proper class. We performed the so called discriminant analysis with the purpose of finding a tool which could test the relation of values of all parameters to the intuitively generated classification, and which also could help the archaeologist to classify newly explored layers. The discriminant analysis tries to find the mathematical function (*discriminant factor* or *discriminator*) that best characterizes groups of measurements within given classes of objects and then uses this function to allocate individual objects to the correct classes (see Appendix 1 for details). It is necessary to test the value of the discriminator by using it to classify the same data, and compare the resulting classification with the original one.

We made four experiments with the discriminant analysis: we performed it for each archaeologist separately, for all classified layers together and finally we let the computer find the discriminator on the base of the layers of archaeologist B and then use it to classify the layers of archaeologist A. Table 1 presents the results in absolute values and in percentage: the numbers of originally classified layers (columns) assigned to the classes by the computer (rows). In the case of an ideal identity of both classifications, all the extra-diagonal numbers should be equal to zero and the diagonal ones equal to 100%.

The results of the first two experiments are similar (tab 1a and b). They show real resemblance of the criteria used by the archaeologists and the computer to distinguish classes D (habitation) and E (destruction), while Classes B (levelling) and C (building) are not so precisely defined. Because archaeologist B had problems with distinguishing between his classes B and D, the computer erroneously assigned as many as 30% of levelling layers to class D (habitation). Generally, archaeologist B:s layers were a little better defined than those of archaeologist A, with exception of class B. Classes A (natural) and F (other) gave the best results: 100% of them were correctly classified by the computer.

Results of the third experiment, (tab 1c) made for all of the layers together, were expected to be worse than previous, which proved true (with exception of classes A and F). The earlier noted differences between the archaeologists were confirmed by a large number of wrongly classified layers.

The fourth experiment (tab 1d) shows that, at the moment, it would be rather dangerous to use a computer instead of an archaeologist's experience. The computer trained by archaeologist B was unable to classify correctly archaeologist A layers (except class D). Especially interesting is the complete mis-interpretation of class E through the assignment of 72% of the destruction layers to class D and none of them to their proper group(!).

Table 2 gives an answer to the question: which of the parameters are helpful in classifying the layers?. Only parameters having more than one case of high percentage are significant and can be used to distinguish between the respective classes. Setting aside classes A and F the best results for archaeologist A gave: **shade** (classes C and D), **inclination**, **twigs** and **seed** (D,E); for archaeologist B: **burning** and **scorching** (B,E), **accumulation** and **grass** (D,E). The results show that some of the physical features as well as some of the components can be used to differentiate between certain classes, and especially between habitation and destruction layers.

In order to check if the computer could be useful in establishing an independent classification of layers, we performed the clustering analysis. We applied K-means clustering algorithm, using minus logarithm of probability as a measure of distance (for details see Appendix 1 and J.A. Hartigan, *Clustering Algorithms* 1975). The optimal number of classes for the algorithm turned out to be 4. The procedure starts with random allocation of layers to 4 groups. Then the program works in a loop, moving layers from one group to another, trying to form clusters of the most similar objects. To measure the progress of the process the chi-squared test is made after each loop (iteration). The program was stopped when the two consecutive iterations gave the same result of the test. Table 3a shows the most interesting output. The new classification produced by the clustering program is similar to the intuitive one for well defined classes C, E: more than 50% of layers belonging to each of them was assigned to one cluster.

Results of this method depend on the initial randomly created groups of objects. The program produced a classification most similar to the original one when the initial groups were identical with our classes B-E (tab 3b). This experiment demonstrated that the intuitive classification has quite a good mathematical basis, again with exception of class B.

## Conclusions

The analyses of the data stored by the inventory of layers demonstrate that the intuitive classifications elaborated by the archaeologists can be supported on more scientific grounds. However, because the definitions of classes are lacking in precision it seems impossible to use a computer as an expert replacing the archaeologists. The machine can help to examine records, to point out human errors or to discover weak points of the system.

Our analysis showed that only 5 from among 13 physical features could be really used, at the moment, to distinguish between classes of layers. It seems that it would be better to apply a smaller

**Table 1a.** DISCRIMINANT ANALYSIS (52 variables used). model: A, data: A

| | Absolute values | | | | | | Percentage | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | E | F | A | B | C | D | E | F |
| A | 4 | 1 | 0 | 0 | 0 | 0 | 100 | 1 | 0 | 0 | 0 | 0 |
| B | 0 | 65 | 10 | 14 | 1 | 0 | 0 | 62.5 | 18.2 | 9.2 | 9.1 | 0 |
| C | 0 | 13 | 34 | 11 | 0 | 0 | 0 | 12.5 | 61.8 | 7.2 | 0 | 0 |
| D | 0 | 19 | 9 | 127 | 0 | 0 | 0 | 18.3 | 16.4 | 83.6 | 0 | 0 |
| E | 0 | 5 | 2 | 0 | 10 | 0 | 0 | 4.8 | 3.6 | 0 | 90.9 | 0 |
| F | 0 | 1 | 0 | 0 | 0 | 4 | 0 | 1 | 0 | 0 | 0 | 100 |
| TOTAL | 4 | 104 | 55 | 152 | 11 | 4 | Total number of classified layers: 330 | | | | | |

**Table 1b.** DISCRIMINANT ANALYSIS (52 variables used). model: B, data: B

| | Absolute values | | | | | | Percentage | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | E | F | A | B | C | D | E | F |
| A | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| B | 0 | 37 | 2 | 14 | 1 | 0 | 0 | 58.7 | 9.1 | 5 | 10 | 0 |
| C | 0 | 6 | 17 | 7 | 0 | 0 | 0 | 9.5 | 77.3 | 2.5 | 0 | 0 |
| D | 0 | 19 | 3 | 256 | 0 | 0 | 0 | 30.2 | 13.6 | 92.1 | 0 | 0 |
| E | 0 | 1 | 0 | 1 | 9 | 0 | 0 | 1.6 | 0 | 0.4 | 90 | 0 |
| F | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| TOTAL | 0 | 63 | 22 | 278 | 10 | 0 | Total number of classified layers: 373 | | | | | |

**Table 1c.** DISCRIMINANT ANALYSIS (52 variables used). model: A+B, data: A+B

| | Absolute values | | | | | | Percentage | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | E | F | A | B | C | D | E | F |
| A | 4 | 1 | 0 | 0 | 0 | 0 | 100 | 0.6 | 0 | 0 | 0 | 0 |
| B | 0 | 81 | 15 | 39 | 2 | 0 | 0 | 48.5 | 19.5 | 9.1 | 9.5 | 0 |
| C | 0 | 30 | 48 | 38 | 0 | 0 | 0 | 18.0 | 62.3 | 8.8 | 0 | 0 |
| D | 0 | 43 | 12 | 348 | 1 | 0 | 0 | 25.7 | 15.6 | 80.9 | 4.8 | 0 |
| E | 0 | 11 | 2 | 5 | 18 | 0 | 0 | 6.6 | 2.6 | 1.2 | 85.7 | 0 |
| F | 0 | 1 | 0 | 0 | 0 | 4 | 0 | 0.6 | 0 | 0 | 0 | 100 |
| TOTAL | 4 | 167 | 77 | 7430 | 21 | 4 | Total number of classified layers: 703 | | | | | |

**Table 1d.** DISCRIMINANT ANALYSIS (52 variables used). model: B, data: A

| | Absolute values | | | | | | Percentage | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | E | F | A | B | C | D | E | F |
| A | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| B | 2 | 38 | 24 | 33 | 3 | 3 | 50 | 36.5 | 43.6 | 21.7 | 27.3 | 75 |
| C | 0 | 4 | 1 | 2 | 0 | 0 | 0 | 3.8 | 1.8 | 1.3 | 0 | 0 |
| D | 2 | 62 | 30 | 116 | 8 | 1 | 50 | 59.6 | 54.5 | 76.3 | 72.7 | 25 |
| E | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0.7 | 0 | 0 |
| F | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| TOTAL | 4 | 104 | 55 | 152 | 11 | 4 | Total number of classified layers: 330 | | | | | |

Table 2a DISCRIMINANT ANALYSIS (52 variables used). model: A, data: A

| | Absolute values | | | | | | Percentage | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | E | F | A | B | C | D | E | F |
| COLOR | 3 | 0 | 0 | 0 | 1 | 3 | 75 | 0 | 0 | 0 | 9 | 75 |
| SHADE | 0 | 0 | 34 | 91 | 0 | 2 | 0 | 0 | 62 | 60 | 0 | 50 |
| LEAKAGE | 3 | 0 | 0 | 0 | 9 | 1 | 75 | 0 | 0 | 0 | 82 | 25 |
| BURNING | 0 | 0 | 55 | 0 | 6 | 2 | 0 | 0 | 100 | 0 | 55 | 50 |
| SCORCHING | 0 | 0 | 0 | 138 | 6 | 3 | 0 | 0 | 0 | 91 | 55 | 75 |
| HOMOGENEITY | 2 | 37 | 0 | 0 | 2 | 3 | 50 | 36 | 0 | 0 | 18 | 75 |
| COHESION | 0 | 38 | 0 | 0 | 6 | 3 | 0 | 37 | 0 | 0 | 55 | 75 |
| LAMINATION | 2 | 47 | 0 | 34 | 1 | 4 | 50 | 45 | 0 | 22 | 9 | 100 |
| GRANULARITY | 0 | 0 | 0 | 110 | 1 | 4 | 0 | 0 | 0 | 72 | 9 | 100 |
| MOISTURE | 3 | 0 | 0 | 34 | 1 | 4 | 75 | 0 | 0 | 22 | 9 | 100 |
| ELASTICITY | 0 | 56 | 4 | 60 | 5 | 2 | 0 | 54 | 7 | 39 | 45 | 50 |
| COMPACTNESS | 2 | 0 | 10 | 1 | 3 | 3 | 50 | 0 | 18 | 1 | 27 | 75 |
| HUMIFICATION | 2 | 0 | 19 | 0 | 6 | 2 | 50 | 0 | 35 | 0 | 55 | 50 |
| INCLITATION | 0 | 0 | 1 | 106 | 7 | 2 | 0 | 0 | 2 | 70 | 64 | 50 |
| ACCUMULATION | 1 | 0 | 0 | 80 | 8 | 2 | 25 | 0 | 0 | 53 | 73 | 50 |
| STONE | 0 | 29 | 0 | 0 | 7 | 4 | 0 | 28 | 0 | 0 | 64 | 100 |
| GRAVEL | 0 | 0 | 0 | 124 | 4 | 1 | 0 | 0 | 0 | 82 | 36 | 25 |
| SAND 1 | 2 | 0 | 0 | 46 | 0 | 3 | 50 | 0 | 0 | 30 | 0 | 75 |
| SAND 2 | 0 | 0 | 1 | 48 | 0 | 4 | 0 | 0 | 2 | 32 | 0 | 100 |
| SAND 3 | 1 | 0 | 0 | 59 | 0 | 4 | 25 | 0 | 0 | 39 | 0 | 100 |
| LOAM | 2 | 0 | 0 | 69 | 7 | 0 | 50 | 0 | 0 | 45 | 64 | 0 |
| CLAY | 0 | 0 | 0 | 122 | 2 | 3 | 0 | 0 | 0 | 80 | 18 | 75 |
| LIME | 0 | 0 | 0 | 2 | 11 | 1 | 0 | 0 | 0 | 1 | 100 | 25 |
| ASH | 0 | 0 | 0 | 3 | 2 | 4 | 0 | 0 | 0 | 2 | 18 | 100 |
| CHARCOAL | 0 | 0 | 95 | 6 | 0 | 0 | 0 | 0 | 63 | 55 | 0 | 0 |
| OTHER MIN | 0 | 0 | 0 | 1 | 0 | 4 | 0 | 0 | 0 | 1 | 0 | 100 |
| INSECT | 0 | 0 | 0 | 0 | 1 | 4 | 0 | 0 | 0 | 0 | 9 | 100 |
| BEETLE | 0 | 0 | 0 | 48 | 0 | 4 | 0 | 0 | 0 | 32 | 0 | 100 |
| BONES | 0 | 27 | 0 | 0 | 3 | 4 | 0 | 26 | 0 | 0 | 27 | 100 |
| HORN | 0 | 10 | 0 | 0 | 0 | 4 | 0 | 10 | 0 | 0 | 0 | 100 |
| HAIR | 0 | 36 | 0 | 0 | 1 | 4 | 0 | 35 | 0 | 0 | 9 | 100 |
| SHELL | 0 | 16 | 1 | 0 | 0 | 4 | 0 | 15 | 2 | 0 | 0 | 100 |
| OTHER ANIM | 2 | 0 | 0 | 4 | 0 | 4 | 50 | 0 | 0 | 3 | 0 | 100 |
| EXCREMENT H | 0 | 0 | 0 | 4 | 0 | 4 | 0 | 0 | 0 | 3 | 0 | 100 |
| EXCREMENT A | 0 | 21 | 0 | 26 | 0 | 4 | 0 | 20 | 0 | 17 | 0 | 100 |
| HUMUS | 3 | 0 | 1 | 53 | 4 | 0 | 75 | 0 | 2 | 35 | 36 | 0 |
| CHIPS | 1 | 0 | 12 | 118 | 6 | 0 | 25 | 0 | 22 | 78 | 55 | 0 |
| TWIGS | 0 | 0 | 15 | 104 | 7 | 0 | 0 | 0 | 27 | 68 | 64 | 0 |
| BARK | 0 | 25 | 32 | 2 | 0 | 4 | 0 | 24 | 58 | 1 | 0 | 100 |
| CONIFER | 0 | 0 | 0 | 34 | 1 | 4 | 0 | 0 | 0 | 22 | 9 | 100 |
| LEAVES | 0 | 0 | 0 | 16 | 0 | 4 | 0 | 0 | 0 | 11 | 0 | 100 |
| CATKIN | 0 | 0 | 1 | 0 | 0 | 4 | 0 | 0 | 2 | 0 | 0 | 100 |
| MOSS | 0 | 21 | 0 | 22 | 0 | 4 | 0 | 20 | 0 | 14 | 0 | 100 |
| ROOTS | 0 | 0 | 2 | 0 | 11 | 1 | 0 | 0 | 4 | 0 | 100 | 25 |
| CHAFF | 0 | 0 | 0 | 44 | 0 | 4 | 0 | 0 | 0 | 29 | 0 | 100 |
| STRAW | 2 | 0 | 0 | 99 | 0 | 3 | 50 | 0 | 0 | 65 | 0 | 75 |
| CORN | 0 | 0 | 3 | 1 | 0 | 4 | 0 | 0 | 5 | 1 | 0 | 100 |
| GRASS | 0 | 0 | 2 | 0 | 1 | 4 | 0 | 0 | 4 | 0 | 9 | 100 |
| NUT | 1 | 41 | 0 | 0 | 0 | 3 | 25 | 39 | 0 | 0 | 0 | 75 |
| FRUIT STONE | 0 | 0 | 0 | 3 | 0 | 4 | 0 | 0 | 0 | 2 | 0 | 100 |
| SEED | 0 | 0 | 0 | 90 | 11 | 0 | 0 | 0 | 0 | 59 | 100 | 0 |
| OTHER BOT | 0 | 0 | 0 | 0 | 3 | 4 | 0 | 0 | 0 | 0 | 27 | 100 |
| TOTAL | 4 | 104 | 55 | 152 | 11 | 4 | Total number of classified layers: 330 | | | | | |

Table 2b DISCRIMINANT ANALYSIS (52 variables used). model: B, data: B

| | Absolute values | | | | | | Percentage | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | E | F | A | B | C | D | E | F |
| COLOR | 0 | 5 | 5 | 127 | 6 | 0 | 0 | 8 | 23 | 46 | 60 | 0 |
| SHADE | 0 | 0 | 0 | 171 | 10 | 0 | 0 | 0 | 0 | 62 | 100 | 0 |
| LEAKAGE | 0 | 0 | 9 | 180 | 6 | 0 | 0 | 0 | 41 | 65 | 60 | 0 |
| BURNING | 0 | 50 | 8 | 0 | 7 | 0 | 0 | 79 | 36 | 0 | 70 | 0 |
| SCORCHING | 0 | 50 | 5 | 0 | 9 | 0 | 0 | 79 | 23 | 0 | 90 | 0 |
| HOMOGENEITY | 0 | 4 | 9 | 210 | 0 | 0 | 0 | 6 | 41 | 76 | 0 | 0 |
| COHESION | 0 | 38 | 13 | 0 | 6 | 0 | 0 | 60 | 59 | 0 | 60 | 0 |
| LAMINATION | 0 | 15 | 0 | 185 | 6 | 0 | 0 | 24 | 0 | 67 | 60 | 0 |
| GRANULARITY | 0 | 0 | 13 | 172 | 7 | 0 | 0 | 0 | 59 | 62 | 70 | 0 |
| MOISTURE | 0 | 27 | 16 | 3 | 3 | 0 | 0 | 43 | 73 | 1 | 30 | 0 |
| ELASTICITY | 0 | 25 | 10 | 49 | 7 | 0 | 0 | 40 | 45 | 18 | 70 | 0 |
| COMPACTNESS | 0 | 28 | 16 | 0 | 2 | 0 | 0 | 44 | 73 | 0 | 20 | 0 |
| HUMIFICATION | 0 | 43 | 1 | 104 | 3 | 0 | 0 | 68 | 5 | 37 | 30 | 0 |
| INCLINATION | 0 | 7 | 15 | 0 | 6 | 0 | 0 | 11 | 68 | 0 | 60 | 0 |
| ACCUMULATION0 | 1 | 0 | 222 | 8 | 0 | 0 | 2 | 0 | 80 | 80 | 0 | |
| STONE | 0 | 0 | 3 | 39 | 9 | 0 | 0 | 0 | 14 | 14 | 90 | 0 |
| GRAVEL | 0 | 12 | 0 | 262 | 0 | 0 | 0 | 19 | 0 | 94 | 0 | 0 |
| SAND 1 | 0 | 0 | 19 | 53 | 2 | 0 | 0 | 0 | 86 | 19 | 20 | 0 |
| SAND 2 | 0 | 0 | 21 | 29 | 4 | 0 | 0 | 0 | 95 | 10 | 40 | 0 |
| SAND 3 | 0 | 0 | 22 | 0 | 2 | 0 | 0 | 0 | 100 | 0 | 20 | 0 |
| LOAM | 0 | 0 | 17 | 66 | 3 | 0 | 0 | 0 | 77 | 24 | 30 | 0 |
| CLAY | 0 | 0 | 18 | 167 | 0 | 0 | 0 | 0 | 82 | 60 | 0 | 0 |
| LIME | 0 | 62 | 0 | 14 | 2 | 0 | 0 | 98 | 0 | 5 | 20 | 0 |
| ASH | 0 | 0 | 0 | 271 | 2 | 0 | 0 | 0 | 0 | 97 | 20 | 0 |
| CHARCOAL | 0 | 0 | 22 | 71 | 2 | 0 | 0 | 0 | 100 | 26 | 20 | 0 |
| OTHER MIN | 0 | 0 | 0 | 11 | 10 | 0 | 0 | 6 | 0 | 4 | 100 | 0 |
| INSECT | 0 | 0 | 1 | 59 | 10 | 0 | 0 | 0 | 5 | 21 | 100 | 0 |
| BEETLE | 0 | 0 | 0 | 0 | 10 | 0 | 0 | 8 | 0 | 0 | 100 | 0 |
| BONES | 0 | 0 | 0 | 103 | 10 | 0 | 0 | 5 | 0 | 37 | 100 | 0 |
| HORN | 0 | 0 | 0 | 0 | 10 | 0 | 0 | 5 | 0 | 0 | 100 | 0 |
| HAIR | 0 | 0 | 0 | 22 | 10 | 0 | 0 | 2 | 0 | 8 | 100 | 0 |
| SHELL | 0 | 0 | 0 | 43 | 10 | 0 | 0 | 0 | 0 | 15 | 100 | 0 |
| OTHER ANIM | 0 | 0 | 0 | 120 | 9 | 0 | 0 | 0 | 0 | 43 | 90 | 0 |
| EXCREMENT H | 0 | 0 | 0 | 29 | 10 | 0 | 0 | 0 | 0 | 10 | 100 | 0 |
| EXCREMENT A | 0 | 0 | 0 | 139 | 10 | 0 | 0 | 0 | 0 | 50 | 100 | 0 |
| HUMUS | 0 | 0 | 21 | 68 | 4 | 0 | 0 | 0 | 95 | 24 | 40 | 0 |
| CHIPS | 0 | 0 | 3 | 256 | 5 | 0 | 0 | 8 | 14 | 92 | 50 | 0 |
| TWIGS | 0 | 0 | 4 | 248 | 5 | 0 | 0 | 0 | 18 | 89 | 50 | 0 |
| BARK | 0 | 0 | 5 | 216 | 6 | 0 | 0 | 0 | 23 | 78 | 60 | 0 |
| CONIFER | 0 | 0 | 3 | 30 | 9 | 0 | 0 | 0 | 14 | 11 | 90 | 0 |
| LEAVES | 0 | 0 | 1 | 13 | 10 | 0 | 0 | 0 | 5 | 5 | 100 | 0 |
| CATKIN | 0 | 0 | 1 | 0 | 10 | 0 | 0 | 0 | 5 | 0 | 100 | 0 |
| MOSS | 0 | 0 | 4 | 87 | 8 | 0 | 0 | 0 | 18 | 31 | 80 | 0 |
| ROOTS | 0 | 10 | 18 | 72 | 0 | 0 | 0 | 16 | 82 | 26 | 0 | 0 |
| CHAFF | 0 | 0 | 0 | 2 | 10 | 0 | 0 | 0 | 0 | 1 | 100 | 0 |
| STRAW | 0 | 0 | 4 | 139 | 10 | 0 | 0 | 0 | 18 | 50 | 100 | 0 |
| CORN | 0 | 0 | 0 | 0 | 10 | 0 | 0 | 0 | 0 | 0 | 100 | 0 |
| GRASS | 0 | 0 | 0 | 201 | 8 | 0 | 0 | 0 | 0 | 72 | 80 | 0 |
| NUT | 0 | 18 | 9 | 1 | 8 | 0 | 0 | 29 | 41 | 0 | 80 | 0 |
| FRUIT STONE | 0 | 23 | 3 | 0 | 9 | 0 | 0 | 37 | 14 | 0 | 90 | 0 |
| SEED | 0 | 28 | 0 | 107 | 9 | 0 | 0 | 44 | 0 | 38 | 90 | 0 |
| OTHER BOT | 0 | 1 | 0 | 1 | 10 | 0 | 0 | 2 | 0 | 0 | 100 | 0 |
| TOTAL | 0 | 63 | 22 | 278 | 10 | 0 | Total number of classified layers: 373 | | | | | |

scale to the description. Although the share of components was noted as a percentage the archaeologists have used only a few values. One can conclude that it would be easier to use a scale of points as it was done for the purpose of computer analyses. Most of components on the card have been used. Verification of kinds of organical components, if necessary, will be possible only along with examination of the samples from the layers made by a biologist.

The archaeologists produced their own classifications, which we believe to be closer to the reality than those made by the computer. It means that we still do not know how to translate our observations to numbers.

The observed diversity of the descriptions produced by the two archaeologists could have many reasons: differences in mentality and experience, various conditions of work (weather, time of the year, etc). There is possibly also a more objective reason: they worked on the two different sets of layers. Only the description of the same layers by both archaeologists could help to exclude the influence of objective reasons.

We believe that the layer description card used in Oslo, albeit still needing corrections, gives as a starting point, a great opportunity of creating such a sheet which along with the computer expert system would be a convenient tool in the course of excavations on multi-layers sites. The statistical methods presented here can help the archaeologists to analyse their documentation before the final publication but cannot be treated as a way of presentation of the results of the excavations.

Table 3a. CLUSTERING ANALYSIS k-means method (15 variables used). Example 1.

| | Absolute values | | | | | | Percentage | | | | | | Chi-square test |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | E | F | A | B | C | D | E | F | |
| iteration 1 | | | | | | | | | | | | | |
| | 0 | 18 | 2 | 66 | 3 | 0 | 0 | 28.5 | 9 | 23.7 | 30 | 0 | 13054 |
| | 0 | 12 | 5 | 81 | 3 | 0 | 0 | 19 | 22.7 | 29.1 | 30 | 0 | |
| | 0 | 17 | 9 | 73 | 3 | 0 | 0 | 26.6 | 40.9 | 26.3 | 30 | 0 | |
| | 0 | 16 | 6 | 58 | 1 | 0 | 0 | 25.4 | 27.3 | 20.9 | 10 | 0 | |
| iteration 13 | | | | | | | | | | | | | |
| | 0 | 21 | 4 | 40 | 10 | 0 | 0 | 33.3 | 18.2 | 14.4 | 100 | 0 | 10887 |
| | 0 | 15 | 2 | 67 | 0 | 0 | 0 | 23.8 | 9 | 24.1 | 0 | 0 | |
| | 0 | 7 | 13 | 66 | 0 | 0 | 0 | 11.1 | 59 | 23.7 | 0 | 0 | |
| | 0 | 20 | 3 | 105 | 0 | 0 | 0 | 31.4 | 13.6 | 37.8 | 0 | 0 | |

**Table 3b.** CLUSTERING ANALYSIS k-means method (15 variables used). Example 2.

| | Absolute values | | | | | | Percentage | | | | | | Chi-square test |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A | B | C | D | E | F | A | B | C | D | E | F | |
| iteration 1 | | | | | | | | | | | | | |
| | 0 | 63 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 12686 |
| | 0 | 0 | 22 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | 0 | |
| | 0 | 0 | 0 | 278 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | 0 | |
| | 0 | 0 | 0 | 0 | 10 | 0 | 0 | 0 | 0 | 0 | 100 | 0 | |
| iteration 2 | | | | | | | | | | | | | |
| | 0 | 29 | 3 | 38 | 0 | 0 | 0 | 46 | 13.6 | 13.7 | 0 | 0 | 12166 |
| | 0 | 10 | 16 | 22 | 0 | 0 | 0 | 15.9 | 72.7 | 7.9 | 0 | 0 | |
| | 0 | 21 | 3 | 209 | 0 | 0 | 0 | 33.3 | 13.6 | 75.2 | 0 | 0 | |
| | 0 | 3 | 0 | 9 | 10 | 0 | 0 | 4.8 | 0 | 3.2 | 100 | 0 | |
| iteration 6 | | | | | | | | | | | | | |
| | 0 | 22 | 2 | 46 | 0 | 0 | 0 | 34.9 | 9 | 16.5 | 0 | 0 | 12166 |
| | 0 | 11 | 17 | 60 | 0 | 0 | 0 | 17.5 | 77.3 | 21.6 | 0 | 0 | |
| | 0 | 23 | 3 | 162 | 0 | 0 | 0 | 36.5 | 13.6 | 58.2 | 0 | 0 | |
| | 0 | 7 | 0 | 10 | 10 | 0 | 0 | 11.1 | 0 | 3.6 | 100 | 0 | |

## Appendix 1

$x_j$ (numerated by index $j=1,..,M$)parameters for every layer identified by index $i=1,..,N$) where $k=1,....,6$. Layers are numerated separately 2within every class so $i=1,..,N_k$ where $N_k$ is equal to the number of layers classified as $k$-class.

Basing on the established earlier fact that $x_j$ are statistically independent, we have computed for every $k$-class a function $L_k(x_1,..,x_M)$which gives (a priori) probability, that we have for a layer just $x_j(j=1,..,M)$ measurements. Having these six functions, we can classify (numerical classification) a layer for which we have $(x_1,..,x_M)$ measurements, assigning it to this $k$-class for which $L_k(x1,..,x_M)$is maximal.

Let us define $N_{jkl}$ as the number of layers in $k$-class which have $x_j$ value equal to $l$. Probability, that we have for a layer just $x_j$ measurement can be expressed as:

$$p_{jkxj} = \frac{N_{jkxj}}{N_k}$$

If we take into account that $(x_1,..,x_M)$variables are statistically independent we can express a probability that we have for a layer just $(x1,..,x_M)$ measurements as an appropriate product:

$$L_k(x_1,..,x_M) = \prod_{j=1}^{M} p_{jkxj}$$

Functions $L_k$ (in statistical terminology referred as likelihood) will be used for discriminant analysis as well as for clustering algorithm.

## Appendix 2

The database collecting all data was designed in the dBase III+. Small part of the computations was made by programs written in dBase IV (correlation analysis, chi-square tests, simple statistics), part using programs written in Turbo-Pascal (histograms) but the main discriminant analysis and the clustering analysis were made by means of programs written in FORTRAN.

### II.1. THE DESCRIPTIVE PART OF THE LAYERS INVENTORY (as a computer database)

| | | | | | |
|---|---|---|---|---|---|
| INVNO | XXXX | | | | |
| COLOUR | CCCCCCCCCCCCC | | | MOISTURE | X |
| SHADE | XX | | | ELASTICITY | X |
| | | HOMOGENEITY | X | COMPACTNESS | X |
| LEAKAGE | X | COHESION | X | HUMIFICAT | X |
| BURNING | X | LAMINATION | X | INCLINATION | X |
| SCORCHING | X | GRANULARITY | X | ACCUMULAT | X |

| | | | | | |
|---|---|---|---|---|---|
| STONES | XXXX XX X | INSECT | XXXX | CHIPS | XXXX XX X |
| GRAVEL | XXXX XX X | BEETLE | XXXX | TWIGS | XXXX XX X |
| SAND 1 | XXXX XX X | BONE | XXXX | BARK | XXXX XX X |
| SAND 2 | XXXX XX X | HORN | XXXX | CONIFER | XXXX XX X |
| SAND 3 | XXXX XX X | HAIR | XXXX | LEAVES | XXXX XX X |
| LOAM | XXXX | SHELL | XXXX | CATKIN | XXXX XX X |
| CLAY | XXXX XX X | OTHER ANIM | XXXX | MOSS | XXXX XX X |
| LIME | XXXX | | | ROOTS | XXXX XX X |
| ASH | XXXX | | | CHAFF | XXXX XX X |
| CHARCOAL | XXXX XX X | EXCREMENT H | XXXX XX X | STRAW | XXXX XX X |
| OTHER MIN | XXXX XX X | EXCREMENT A | XXXX XX X | CORN | XXXX XX X |
| | | HUMUS | XXXX XX X | GRASS | XXXX XX X |
| | | | | NUT | XXXX XX X |
| | | | | FRUIT STONE | XXXX XX X |
| | | | | SEED | XXXX XX X |
| | | | | OTHER BOT | XXXX XX X |

| | | | |
|---|---|---|---|
| NATURAL | CCC | KULTOLK | MEMO |
| LEVELLIING | CCC | | |
| BUILDING | CCC | | |
| HABITATION | CCC | | |
| DESTRUCTION | CCC | | |
| OTHER | CCC | | |