M. Wansleeben L.B.M. Verhart

# The Meuse Valley Project: GIS and site location statistics

The application of Geographical Information Systems in Archaeology is growing fast. With this a more critical attitude towards methodological issues arises. Here the statistical ways to relate site locations to landscape attributes are evaluated and some alternatives for the commonly used Chi-square test are presented.

# 1. Introduction

Of late, the use of Geographical Information Systems has increased dramatically in archaeological research. In the last 4 to 5 years GIS have caught on in a big way. The number of articles about GIS, in the annual proceedings of "Computer Applications and Quantitative Methods in Archaeology" (CAA), has grown from 1 in 1986 (Harris) and 1 in 1988 (Wansleeben) through 6 in 1991 and to 16 in 1995. A few years ago no one could have imagined that this technique from the fields of physical geography, geology and remote sensing would be taken up in such a big way. With hindsight there were of course indications to explain the GIS-"boom". On the one hand there are developments in automation technology. Rapid advances in capacity and processing speed of personal computers and reductions in price have provided many people with the opportunity to process extensive geographical data. On the other hand archaeology displays several traits of a typical spatial science. The source material of archaeology consists of course of the (mobile and immobile) artefacts themselves, on the other hand the spatial distribution (context) of those sites is of equal importance.

Many European archaeologists have discovered the ease of using GIS for analysing and presenting spatial information. A growing number of specialised articles (e.g. CAA), monographs (e.g. Allen *et al.* 1990, *Interpreting space: GIS and archaeology*) and meetings (e.g. *Impact of Geographic Information Systems on Archaeology*, fall 1993, Italy) has been the result. As with most new methods, at first everybody is overjoyed. After a while, however, the drawbacks become clearer and this is happening to GIS as well. In geography, people have been critical of GIS for many years and fundamental research is done into some crucial elements of GIS. For example, to what degree do the choice of grid size and errors in the basic maps influence the end result of the analyses. In archaeology Kvamme is one of the more critical users of GIS (Kvamme 1990, 1993). In this context, Allen, Green and Zubrow were right too to conclude:

"In either case, we caution against the use of GIS as an end in itself. Good research and management is based on asking good [archaeological] questions – something GIS does not do for us." (Allen et al. 1990, 383, [text] added).

In future, more if these criticisms will probably be forthcoming. This article is an expression of the more critical attitude towards GIS.

## 2. GIS

The strength of GIS is to a great extent the ability to manipulate spatial information in a quick and simple way. One of the best known examples is the conversion of a map of rivers into one showing the distances to those rivers. Something almost impossible to do by hand. In archaeology, other widely used options are computing the gradient and the exposure from a digital elevation model (DEM). The origins of GIS, in geography and geology, are however still very obvious. Many options refer to the representation and analysis of areas (e.g. soils) and lines (e.g. rivers). A soil map can be compared with a geological map and a new overlay containing all combinations of legend units is easy to compile. Archaeological information, however, has a character of its own. In regional research all archaeological sites are represented as specific points. Some GIS are not really suited to represent and analyse point data. In a grid cell the frequency of sites can be recorded in an overlay, but the archaeological overlay is treated like a soil map, when trying to correlate sites with a geographical variable. The unit of observation is not the archaeological site, but the grid cell. A cell containing 3 sites would count for only one in a statistical analysis. It is typical that a separate routine had to be written, when using the popular, DOS-based GIS program Idrisi, to allow the Chi-square calculation of the relationships between archaeological sites and a geographical variable, a common procedure in archaeology. It is great that a GIS allows easy and quick

manipulation of geographical data, but in archaeological research the point of the exercise is to access the relationships between sites and geography. In this article we shall examine the statistical techniques available in GIS and their usefulness for answering archaeological questions.

### 3. Meuse Valley Project

In the regional archaeological research into the transition from Mesolithic to Neolithic in the southeast of The Netherlands, GIS has been used almost from the start in 1986. This so-called Meuse Valley Project, attempts to trace the economic changes during this transition with the aid of changes in the settlement system (Wansleeben/Verhart 1990). We defined the settlement system as the combination of the nature, distribution and geographical location of sites from a specific archaeological period. A more or less random distribution pattern of almost equal Mesolithic base camps along brooks and rivers represents an economic system entirely different from clustered villages with wooden buildings, inhabited for 400 years, lying in the fertile loess soils in the first phase of the Neolithic. Next to the nature and distribution, the geographical location is an important source of information, to gain insight into the neolithisation process in the southeast of The Netherlands. Traditionally, the geographical location is investigated in archaeology using site location and site catchment analyses.

In the Meuse Valley Project research takes place on four different spatial levels. The data presented here refer to the highest, so-called macroregional level. Over an area of more than 4400 square kilometres data about almost 4000 Stone Age sites were compiled from literature and from the archaeological data bank of the State Service for Archaeological Investigations in the Netherlands (ROB). The quality of the data obtained about the sites differs widely. Still, by considering the presence or absence of guide artefacts to be a major factor in dating, many sites can be ascribed to one or several archaeological periods. In all, 8 archaeological periods could be distinguished. On the basis of the maps available (scale 1:25.000), geographical data were compiled about 7 different geographical characteristics. These data were stored in a raster-based GIS grid with unit cells of 1 square kilometre. In the past, the GIS was only used for the site location analysis of these sites on the macroregional level. At present, the spatial information on each spatial levels is stored and edited using GIS technology, and the applications are no longer confined to location analysis. However, this intensive use had also led to a more critical attitude towards GIS.

#### 4. Site location analysis and GIS

Site location analysis is a technique describing the geographical position of sites, to detect locational

preferences, if any, of a given society. This revolves around two questions. Taking for example the soil type, these are:

- Is the soil something that people took into account when deciding on the location for a settlement?

## And if so:

– Which types of soils were selected?

When it is clear which geographical units were preferred, these can be correlated with certain economic activities. The economic interpretation is certainly not exclusively based on this location, but as mentioned also on the nature of the site and the presence of other sites nearby. That is why in the Meuse Valley Project the *settlement system* is the true archaeological correlate of the economy. Methodologically, it is useful to have a closer look at site location analysis and see how the two questions can be answered with the aid of GIS.

The 'standard' method to investigate the distribution of sites in relation to a geographical variable is the Chi-square test. The distribution of sites is compared to the distribution of the geographical units (tab. 1).

In this example the distribution of the sites of the first Neolithic society in the southeast of The Netherlands, the Linear Bandceramic Culture (LBK) is compared to the soil texture type (fig. 1). The Chi-square test compares the frequency of sites observed in each geographical legend unit to an expected frequency. The expected frequency is based on a random distribution pattern. Let's assume that the LBK-people were not at all interested in the soil type when choosing a location for their settlements. In that case, the 39 sites would be distributed proportionaly over the legend units. In the sandy area, comprising 57.84% of the research area, 57.84% of the 39 sites would be located (= 22.6). The Chi-square value of 91.082 indicates a significant deviation from a random distribution.

Before continuing in this vein, some points must be made in order not to complicate matters needlessly. This discussion is based on the hypothesis, not very realistic in everyday archaeological research, that the observed distribution is not caused by other factors, like postdepositional processes or investigative influences, but is solely the result of behaviour in the past. Furthermore it should be noted that the Chi-square test in this situation does not meet all statistical requirements. To recall a famous rule of dumb: When the number of classes (k) is larger than 2, the Chi-square test may be used if fewer than 20 per cent of the cells have an expected frequency of less than 5 and if no cells have an expected frequency of less than 1 (Siegel 1956, 110). In the example given above two classes have an expected value lower than 5 and one class lower than 1. The result of the Chi-square may therefore be suspect.

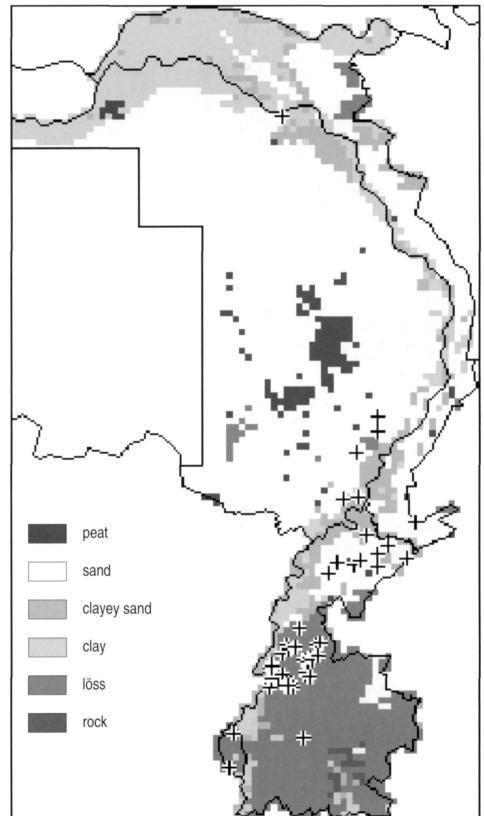


Figure 1. The spatial distribution of settlements of the Linear Bandceramic Culture in relation to the soil texture types in the southeast of The Netherlands.

#### ANALECTA PRAEHISTORICA LEIDENSIA 25

Table 1. Chi-Square Goodness-of-Fit Test.

	population (texture)		sample		
			number	1	
class	number of cells	proportion	observed	expected	Chi-square
peat	139	.0315	0	1.2	1.23
sand	2553	.5784	8	22.6	9.39
clayey sand	388	.0879	3	3.4	.05
clay	653	.1479	1	5.8	3.94
löss	663	.1502	27	5.9	76.30
rock	18	.0041	0	.2	.16
totals	4414	1.0000	39	39.0	91.08

Chi-Square = 91.082 d.f. = 5 p < 0.000

Table 2.	Density	and	proportion	of	LBK sites.	
----------	---------	-----	------------	----	------------	--

class	number of cells	number of sites observed	proportion	density
peat	139	0	.0000	.000
sand	2553	8	.2051	.003
clayey sand	388	3	.0769	.008
clay	653	1	.0256	.001
löss	663	27	.6923	.041
rock	18	0	.0000	.000
totals	4414	39	1.0000	

A possible solution to this problem is applying Yates' correction of continuity (Thomas 1986). Application of this correction yields a new Chi-square of 84.771, significant in itself (p < 0.001). If the number of cases becomes very small, even the correction of continuity is no longer applicable. *If the number of cases is less than 20, or if the number of cases is between 20 and 40 and the smallest expected frequency is less than 5, the Fisher exact probability test should be used* (Siegel 1956, 110). Our example meets this last condition, but even then the Chi-square is significant (p < 0.001).

Keeping in mind these limitations, we can draw the conclusion that there clearly is a significant deviation from a random site location pattern. The LBK-people made, directly or indirectly, a whole-hearted decision in favour of certain soil types. This makes it clear that the Chi-square test can answer the first question.

To gain insight into the choice of soil types, we may look for high Chi-square values per class. These occur in this case for sand and loess and refer to too few sites on sand and too many on loess. This should lead to the archaeological conclusion that the loess was preferred and sand was avoided. However to us this conclusion poses some questions. There are still 8 sites (20.5%) located in the sand. Did these have no economic significance for the LBK- people at all? Likewise, the mean density of sites in the loess area is in an absolute sense low (0.041 sites/square km), but relatively the highest (tab. 2). Second in order is the clayey sand area with 0.008 sites/square km. In spite of the fact that the frequency observed hardly deviates from the expectation, should the conclusion be drawn that the clayey sand area was important as well? So the Chi-square test seems to present difficulties in choosing the relevant legend units and as such, to the archaeological interpretation. Still, this is almost the only way in which point locations can be analysed in GIS.

An other way to choose between legend units might be attributing relative weights to the various units. An easy to use weight is the proportion of the number of sites in a unit: 69% of the LBK-sites is located in the loess area (tab. 2).

It can be assumed that the greater the proportion, the greater the economic importance of a unit. Loess is economically the most important soil type, followed by sand, clayey sand and clay. Peat and rock did not have sufficient economic importance to be settled. In itself, this approach seems archaeologically meaningful, but it does not lead to any kind of generalisation or choice of legend units. Furthermore the information that there are far less sites in the sand area than expected is not taken into account. An alternative is provided by Atwell and Fletcher (1985, 1987). They try to define for each class a so-called weight-factor, that can be considered as an estimate of relative importance of a geographical unit.

"... to access the relative importance of eight environmental characteristics in the choice of cairn locations..." (Atwell/Fletcher 1987, 2).

Assuming there are three geographical units  $\alpha$ ,  $\beta$  and  $\gamma$ , this leads to the following formulas for the weight-factors (A, B and C):

```
A = a'bc / (a'bc + ab'c + abc')

B = ab'c / (a'bc + ab'c + abc')

C = abc' / (a'bc + ab'c + abc')

where:
```

a, b, c = proportion of the geographical units  $\alpha$ ,  $\beta$  and  $\gamma$ . a', b', c' = proportion of the number of sites in geographical units  $\alpha$ ,  $\beta$  and  $\gamma$ .

The values of A, B and C range from 0 to 1 and are 0.33 for each class in a random site distribution. In general, the expected value is 1 divided by the number of classes. The higher the weight-factor, the greater the relative importance. Referring to the LBK-example, the weight-factors are shown in table 3.

The weight-factors seems to reflect the relative importance of the legend units well. In the calculations the size of the legend unit is taken into account, comparable to the effect of that size in the Chi-square test. The large sand area receives a smaller weight-factor than the small clayey sand area, in spite of the fact that in an absolute sense, there are more sites there. The Atwell and Fletcher procedure is an improvement over the proportions in themselves, however there is still no statistically based criterion for choosing legend units.

To this end, Atwell and Fletcher (1985) propose a test: is there a significant deviation from the expected value for one or more observed weight-factors? Atwell and Fletcher use a simulation to determine the expected value. In this procedure, a random site distribution is simulated one hundred times and all weight-factors are calculated. For each simulation, the highest weight-factor is determined. In this way a distribution is obtained of the maximum weight-factor (fig. 2). With these data a threshold expected value can be determined, for example for a 5% confidence level. All observed weight-factors exceeding that threshold value are considered significant by Atwell and Fletcher (1985). These geographical units clearly were preferred for settlements. In the same way a distribution can be calculated for the minimum weight-factor. All geographical units with an observed weight-factor lower than that 5% threshold were significantly avoided.

We feel that Atwell and Fletcher are wrong to use the highest and lowest weight-factors per simulation in calculating the distributions. For each simulation, a different legend unit may show this minimum or maximum. In our opinion, a theoretical distribution should be calculated for each legend unit separately. To test this hypothesis, the LBK-example was used in a computer simulation. The experiment proved that each geographical unit did have its own theoretical distribution, very much different from the maximum or minimum (fig. 3). The shape of the distribution depends on the size of the legend unit; for small units erratic fluctuations may occur. At the same time, the shape of the distribution proved to be dependent on the total number of sites.

The procedure suggested by Atwell & Fletcher results in very conservative tests. Both preferred and avoided legend units are considered <u>not</u> significant too easily. Only the loess would be preferred significantly (p = 0.003). If our criticism of Atwell and Fletcher is valid, the judgement of the weight-factors should be adjusted. In our approach, the loess is significantly preferred (p < 0.001) and sand (p = 0.013) and clay (p = 0.005) are significantly avoided.

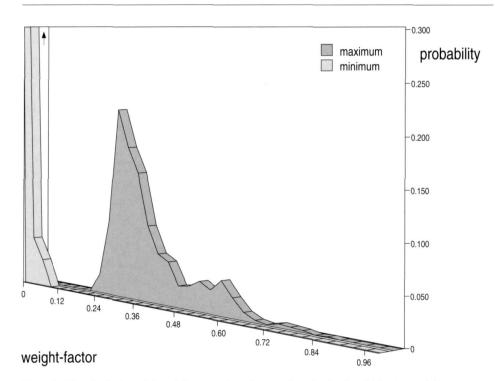
Using this procedure, a statistically based choice of relevant legend units seems feasible, answering the second question. The Atwell-Fletcher test or its modification, does not take into account the simultaneous differences for all legend units. The Chi-square test is therefore still valuable. Both procedures, Chi-square and Atwell-Fletcher, might be combined, or is there still another approach?

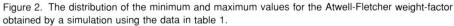
Especially in Cultural Resource Management it is attempted to model the distribution of archaeological sites. For a geographical variable it is decided which legend units are likely and which are unlikely to contain archaeological sites. This in order to select legend units and allow the management of the archaeological soil archive to be as efficient as possible. The decision to drop certain units almost always means that a number of sites are not included in the model. Many archaeologists feel more or less justified in doing so when most sites turn out to be included in the model and the model covers only a relatively small part of the research area. Some examples are:

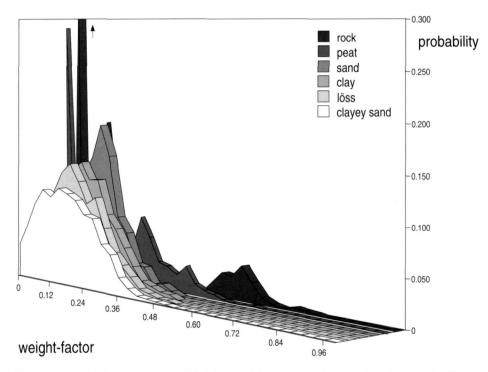
"By applying Bayes' Theorem to the model results it can be suggested that about 72 per cent of the prehistoric sites in the region should occur in 45 per cent of the available land area" (Carmichael 1990, 222).

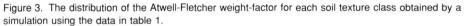
"Although more than 95 percent of the known sites fall within the favourable area, this region covers only about 50 percent of the total study area, pointing to the predictive gain of this model" (Kvamme 1989, 181).

We used this principle to define a kind of site location parameter  $(K_i)$ . We feel two points should be stressed. First









104

class	population (texture)		sample (L		
	number of cells	proportion	number of sites observed	proportion	weight-factor
peat	139	.0315	0	.0000	.000
sand	2553	.5784	8	.2051	.059
clayey sand	388	.0879	3	.0769	.146
clay	653	.1479	1	.0256	.029
löss	663	.1502	27	.6923	.767
rock	18	.0041	0	.0000	.000
totals	4414	1.0000	39	1.0000	1.000

Table 3. Atwell-Fletcher Weight-factors.

Expected value of Weight-factor: 0.167

Table 4. Site Location Parameter Ki

	population (texture)		sample (LBK)		highest	
class	number of cells	proportion	number of sites observed proportion		K <sub>j</sub>	(after step)
peat	139	.0315	0	.0000	.622	(4)
sand	2553	.5784	8	.2051	.393	(6)
clayey sand	388	.0879	3	.0769	.641	(2) ←max
clay	653	.1479	1	.0256	.572	(5)
löss	663	.1502	27	.6923	.615	(1)
rock	18	.0041	0	.0000	.639	(3)
totals	4414	1.0000	39	1.0000		

of all the proportion of sites included in the model  $(p_s)$  is incorporated into the parameter. A model including all sites  $(p_s=1.00)$  is better than one where only a small percentage of the sites is represented. Secondly, the difference between the proportion of sites  $(p_s)$  and the proportion of the research area  $(p_a)$  is an important factor in this parameter. This difference  $(p_s-p_a)$  indicates the relative gain of the model. In a valuable model this difference is large. Kvamme's example yields a relative gain of 0.45 (0.95 minus 0.50). A small difference on the other hand indicates a low predictive value of the model.

Both factors,  $p_s$  and  $p_s$ - $p_a$ , should be as high as possible. In the case where an entire research area is included in the model, the value of  $p_s$  equals 1.00, but since  $p_a$ =1.00 as well, there is no relative gain at all ( $p_s$ - $p_a$ =0.00). Therefore the model has hardly any predictive value, and the parameter  $K_j$  should have a low value. Arithmetically this can be achived by using the product of the two factors. Multiplication of 2 ratios however yields a distribution with an underrepresentation of the higher values, hence the decision to use the root of the product. A formula for the location choice parameter  $K_j$  could be as follows:

# $\mathbf{K}_{\mathbf{j}} = \sqrt{[\mathbf{p}_{\mathbf{s}} * (\mathbf{p}_{\mathbf{s}} - \mathbf{p}_{\mathbf{a}})]}$

One more refinement is necessary: theoretically, the maximum value of p<sub>s</sub> is 1.00. However, the maximum difference between p<sub>s</sub> and p<sub>a</sub> does not equal 1.00, therefore both factors do not contribute equally to K<sub>i</sub>. The maximum difference depends on the degree to which the sites are clustered in the research area. Let us assume that 5 sites are located in 4 cells of a research area of 16 square kilometres (fig. 4). A perfect model comprises the 4 cells with sites  $(p_s = 1.00)$ . The differences  $p_s - p_a$  is however only 1.00 -0.25 = 0.75. This value equals the proportion of the area without sites. The difference is corrected for the degree of clustering of the sites. To do so,  $p_s-p_a$  is divided by the proportion of the research area containing no sites  $(p_w)$ . In this example, 12 of the 16 cells contain no sites ( $p_w=0.75$ ). The corrected value  $(p_s-p_a)/p_w$  rises to 1.00. Both factors now have the same maximum and the maximum of K<sub>i</sub> equals 1.00 as well. In formula, K<sub>i</sub> is now defined as:

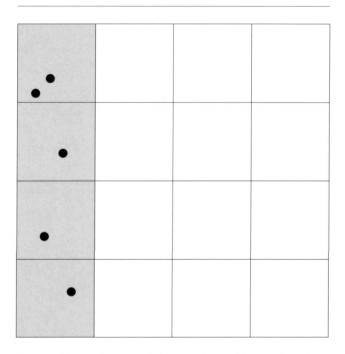
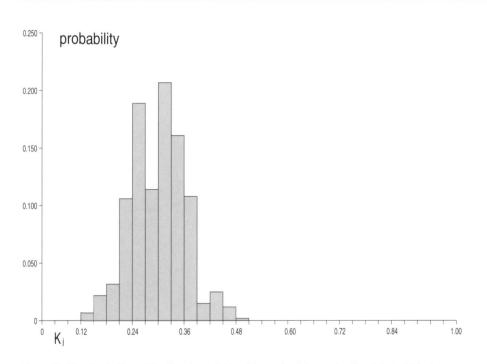


Figure 4. The maximum predictive gain of a model depends on the degree to which the sites are clustered. For the best possible model (shaded area) the value of  $p_s$ - $p_a$  is only 0.75.





$$\mathbf{K}_{j} = \sqrt{[\mathbf{p}_{s} * \frac{\mathbf{p}_{s} - \mathbf{p}_{a}}{\mathbf{p}_{w}}]}$$

where:

The average site density is less then 1 per unit cell  $p_s =$  the proportion of the sites incorporated in the model  $p_a =$  the proportion of de area incorporated in the model  $p_w =$  the proportion of the area without archaeological sites

To calculate  $K_j$  let us once again return to the LBKpeople and the soil types. First of all (step 1), for each legend unit the  $K_j$  value is calculated. If the proportion of sites is lower than the proportion of the area (cells),  $K_j$  is not applicable. In that case, the value of  $p_s$ - $p_a$  is negative and the root can not be extracted. This situation will however only occur in the calculations of the first step.

The highest value at this first step, as marked with (1)in table 4, is found with loess ( $K_i=0.615$ ). By including only all loess cells in the model and leaving out all other units, 69.2% of the sites are already included in a model covering only 15.0% of the research area (relative gain: 54.2%). The logical next question is: can the  $K_i$  value be improved by adding one or more legend units? New calculations (step 2) are made for each combination of loess with the other geographical units. This makes it clear that the model can at best be improved by adding the clayey sand soils (highest value of K<sub>i</sub> after step 2: 0.641). The relative gain of the model is somewhat smaller (53.1%), but the proportion of sites added to the model (7.7%) makes amply up for that loss. To our minds, this is a better model then the previous one. Adding another legend unit results in a deterioration of the model (highest value of K<sub>i</sub> after step 3, adding rock: 0.639). The best model has already been obtained after the second step. When only the loess and clayey sand soils are included (23.8% of the research area), 76.9 % of the sites are represented anyway.

Calculating  $K_j$  is a continuous addition of a single unit to an existing model. Strictly, the process could be stop when there is no longer any improvement in  $K_j$ . The geographical units included in the model may be considered the areas that the LBK-people preferred. Calculation of the  $K_j$  value does not shed any light on the avoided geographical units.

It is possible to test whether the value found for  $K_j$  deviates significantly from the value obtained when there is absolutely no preference for a geographical variable. In that case the value of  $K_j$  should be 0.00, but there are always accidental variations. One thousand times a random site distribution over the given legend units has been simulated. Every time the  $K_j$  value has been calculated. The distribution obtained in this way (fig. 5) gives an idea how reliable the observed result is. The shape of this distribution proved to vary with the number of legend units and the total number of sites. There is a significance of less than 0.001

for the observed  $K_j$  value of the LBK-sites in relation to the soil type. So it seems that the LBK-people did take the soil into account when deciding on the location of a settlement. Not only was the loess of major importance, but beyond that the clayey sand soils seem to have been attractive as well.

So, the site location parameter  $K_j$  can be used both for testing the importance of a geographic variable (first question) as for highlighting the most important legend units (second question). However, it does not uncover which units were significantly avoided.

#### 5. Conclusions

In almost all Geographic Information Systems the emphasis is on ways to manipulate and present geographical information. This is easily explained by the geographical origin of GIS. There is only a very limited supply of the kind of statistical procedures that are of paramount importance to archaeologists: those that test the relationship between point data (sites) and the landscape. Some GIS do still not offer the possibility to treat the sites as a unit of observation, but results are always based on cells.

Fortunately in most GIS a Chi-square test is available. The Chi-square test is the most common way to investigate the relationship between sites and terrain characteristics, but it has its limitations both methodologically and theoretically. Alternatives are almost never available in GIS. Both the (modified) Atwell-Fletcher test and the newly proposed site location parameter  $K_j$  may be useful in site location analyses. Probably, both will not perform well under all circumstances, but they do offer the opportunity to look at the same data in a way different from the Chi-square test.

The limitations due to the article size prevent us from a discussion of other problems connected with site location analysis. To name but three. First of all, how can the mentioned tests be applied to non-nominal variables, as e.g. altitude, gradient and other geographical interval or ratio data? Is the regularly used Kolmogorov-Smirnov onesample test (D) an archaeologically useful alternative? Secondly the choice between a univariate and multivariate approach is an important methodological decision, deserving elaboration. And in the third place in the examples provided here the entire research area has been included in the analysis. This implies the assumption that there is not a single archaeological site in the empty cells. In practice, however, there may be lots of reasons for the lack of sites in no way connected to the habitation in the past. How to take this into account in a site location analysis is a problem on its own. In the Meuse Valley Project a pilot study is run to look at these and other methodological problems and decisions. The choice of statistical test as discussed here is one of the problems being investigated.

# references

Allen, K.M.S. et al (eds)	1990	Interpreting space: GIS and Archaeology, Taylor/Francis, London.
Atwell, M.R., M. Fletcher	1985	A new technique for investigating spatial relationships: Significance testing. In: A. Voorrips/S.H. Loving (eds), <i>To pattern the past</i> , Pact, volume II, Strassbourg, 181-189.
	1987	An Analytical Technique for investigating Spatial Relationships, <i>Journal of Archaeological Science</i> 14, 1-11.
Carmichael, D.L.	1990	GIS predictive modelling of prehistoric site distributions in central Montana. In: K.M.S. Allen <i>et al.</i> (eds), <i>Interpreting space: GIS and Archaeology</i> , Taylor/Francis, London, 216-225.
Harris, T.M.	1986	Geographic Information System design for archaeological site information retrieval. In: S. Laflin (ed.), <i>Computer Applications and Quantitative Methods in Archaeology</i> , University of Birmingham, Birmingham, 148-161.
Kvamme, K.L.	1989	Geographical Information Systems in Regional Archaeological Research and data management. In: M.B. Schiffer (ed.), <i>Archaeological Method and Theory</i> I, 139-203.
	1990	GIS algorithms and their effect on regional archaeological analysis. In: K.M.S. Allen <i>et al.</i> (eds), <i>Interpreting space: GIS and Archaeology</i> , Taylor/Francis, London, 112-125.
	1993	Spatial Statistics and GIS: an integrated approach. In: J. Andresen/T. Madsen/I. Scolllar (eds), <i>Computing the Past. Computer Applications and Quantitative Methods in Archaeoly</i> 1992. Aarhus University Press, Aarhus, 91-103.
Siegel, S.	1956	Nonparametric statistics in behavioral sciences, McGraw-Hill, New York.
Thomas, D.H.	1986	Refiguring Anthropology, Waveland Press, Inc., Prospect Heights, Illinois.
Wansleeben M.	1988	Applications of geographical information systems in archaeological research. In: S.P.Q. Rathz (ed.), <i>Computer Applications and Quantitative Methods in Archaeology</i> 1988, BAR international series 446 (ii). 435-451.
Wansleeben, M., L.B.M. Verhart	1990	Meuse Valley Project: the Transition from the Mesolithic to the Neolithic in the Dutch Meuse Valley. In: P.M. Vermeersch/P. Van Peer (eds), <i>Contributions to the Mesolithic in Europe</i> , Leuven University Press, Leuven, 389-402.

M. Wansleeben Instituut voor Prehistorie postbus 9515 NL-2300 RA Leiden

L.B.M. Verhart Rijksmuseum van Oudheden postbus 11114 NL-2301 EC Leiden