8

How many tombs make a site?

Mike Fletcher
Department of Mathematics, North Staffordshire Polytechnic

Mike Attwell
Department of Computing, North Staffordshire Polytechnic

8.1 Introduction

In this paper we propose an objective method for investigating a point pattern distribution (whether of artifacts or sites) to determine whether clustering in 2-dimensional space is present, and to identify the distance at which such clustering occurs. The proposed method is aimed primarily at detecting clusters whose dimensions are small by comparison with the overall size of the study area. Two techniques are commonly applied to this type of problem: 2-dimensional cluster analysis and nearest neighbour analysis. However, both of these are of limited use since the former is primarily descriptive, whilst the latter provides only minimal information and in many applications is subject to distortion produced by the influence of topographical factors.

The method was developed in response to theoretical considerations arising from a locational analysis of the factors which may have influenced the choice of siting of the chambered tombs of County Leitrim, Ireland. It was decided to determine whether tightly packed groups of tombs occurred, with the aim of reducing them to single locations within the subsequent analysis. The implications and interpretations of this work are beyond the scope of this paper and will we hope be the subject of a future publication. Thus, we will be concerned here solely with the new method and no attempt will be made to place archaeological interpretations upon any of the results.

Two data sets will be considered: the chambered tombs of County Leitrim and the chambered tombs of County Donegal, Ireland (Fig. 8.1).

8.2 Leitrim

The application of nearest neighbour analysis to the distribution of the 52 chambered tombs in County Leitrim (Fig. 8.2; deValera & Ó Nualláin 1972) produces a mean nearest neighbour distance of 1.52 km and a ratio statistic of 0.57. This suggests an aggregated distribution but provides little indication of the possible existence of tightly packed clusters. An analysis of the distribution using the SPSSX cluster analysis package (centroid squared Euclidean distance method) produces the results shown in Fig. 8.3. The steep negative gradient in the early section of the curve (*i.e.* for low distances) suggests that there could be a tendency for sites to cluster below 500 metres. However, this is far from objective and the shape of the curve could simply be a product of relatively large scale topographic influences.

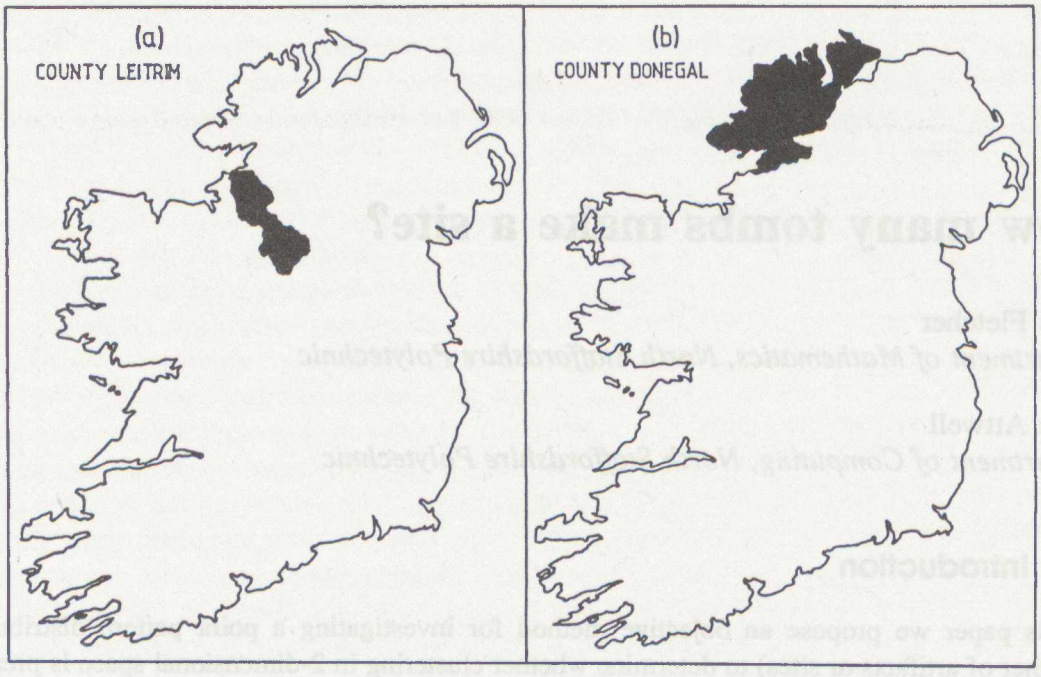

Fig. 8.1: County Leitrim and County Donegal

It should be noted that the cluster analysis package produces tabular output consisting of a unitary scale of clusters formed, and the cluster coefficient (squared Euclidean distance). For the purposes of clarity and ease of interpretation we have written an algorithm to transform the SPSSX output in order to provide graphs of a unitary scale of distance v. the number of clusters remaining. The effect of this can be seen in Table 8.1. For example, at a distance of 4 units (400 metres) there is a reduction of twelve in the number of clusters remaining since there are twelve stages with coefficients of 16 or less.

At this stage it would seem appropriate to generate random data sets which could be subjected to cluster analyses and the results then compared with those produced by the real data. However, this process would prove extremely time-consuming since, amongst other things, it would involve digitising the boundaries of the study area. It was decided, therefore, to adopt an alternative strategy and to perturb the real data by a series of relatively small distances, and to perform cluster analyses on the data sets thus produced. The results of these analyses can then be compared with those produced from the real data. This approach is similar to that used in Lock & Fletcher 1986 in a recent investigation into methods of identifying patterns within post-hole distributions.

It was expected that the difference between the results produced by cluster analyses on the perturbed and on the real data would become more apparent as the perturbation distance was increased. For example, clustering at a distance of less than 400 metres should be destroyed by perturbing by approximately 500 metres. It should be noted that this may also slightly increase the number of clusters at larger distances.

In the analysis of the Leitrim tombs the real data was initially perturbed by 500 metres. This was achieved by employing a rectangular (or uniform) perturbation algorithm with the x and y

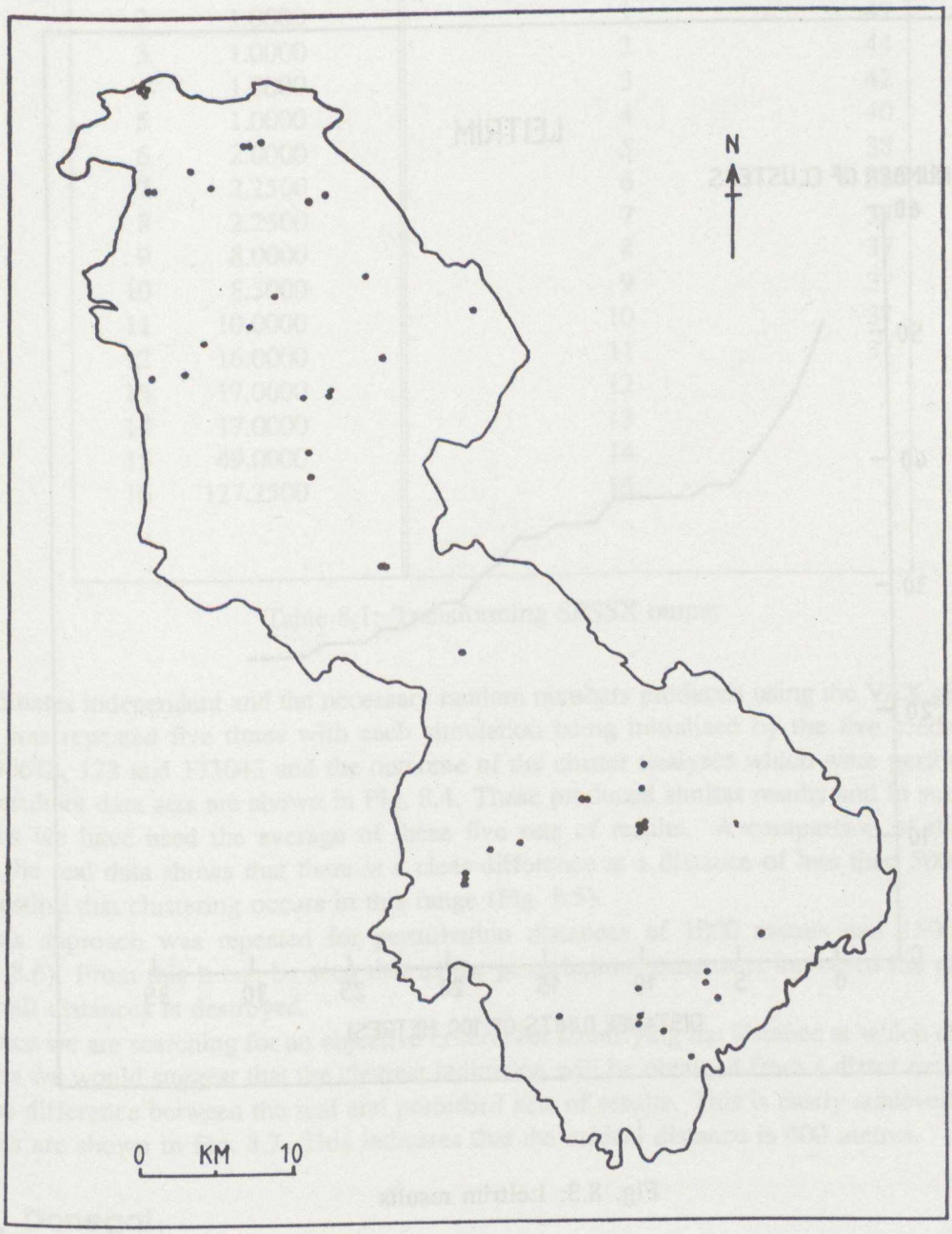

Fig. 8.2: Distribution of chambered tombs in Co. Leitrim

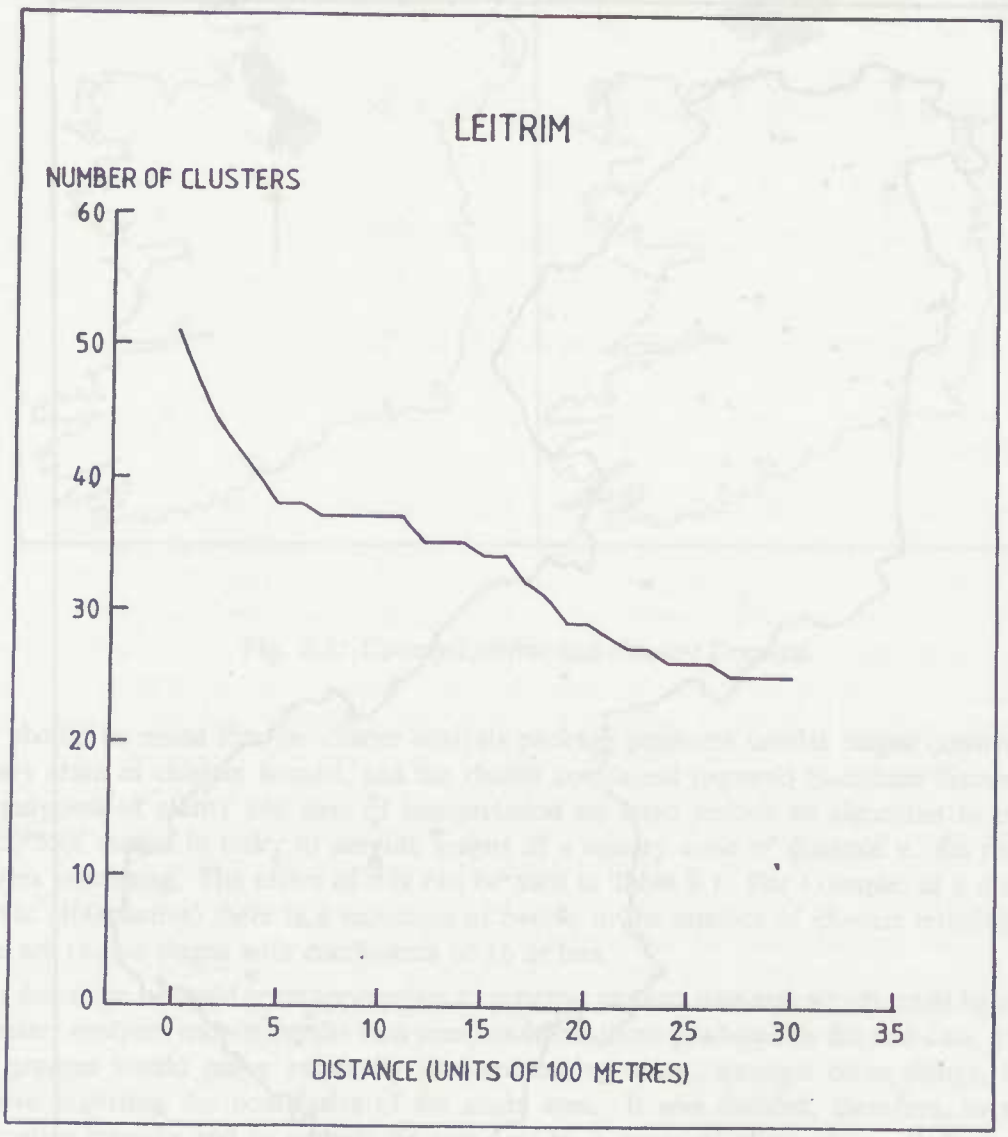

Fig. 8.3: Leitrim results

Standard SPSSX output		Transformed results	
Stage	Coefficient	Distance (100m)	Number of Clusters
1	0.0000	0	51
2	1.0000	1	47
3	1.0000	2	44
4	1.0000	3	42
5	1.0000	4	40
6	2.0000	5	38
7	2.2500	6	38
8	2.2500	7	37
9	8.0000	8	37
10	8.5000	9	37
11	10.0000	10	37
12	16.0000	11	37
13	17.0000	12	.
14	17.0000	13	.
15	49.0000	14	.
16	127.2500	15	.
.	.	.	.
.	.	.	.

Table 8.1: Transforming SPSSX output

coordinates independent and the necessary random numbers produced using the VAX generator. This was repeated five times with each simulation being initialised by the five seeds 14630, 17, 30618, 123 and 171043 and the outcome of the cluster analyses which were performed on the resultant data sets are shown in Fig. 8.4. These produced similar results and in subsequent graphs we have used the average of these five sets of results. A comparison of this curve with the real data shows that there is a clear difference at a distance of less than 500 metres, suggesting that clustering occurs in this range (Fig. 8.5).

This approach was repeated for perturbation distances of 1000 metres and 1500 metres (Fig. 8.6). From this it can be seen that as the perturbation distance is increased the clustering at small distances is destroyed.

Since we are searching for an objective criteria for identifying the distance at which clustering occurs we would suggest that the clearest indication will be obtained from a direct examination of the difference between the real and perturbed sets of results. This is easily achieved and the results are shown in Fig. 8.7. This indicates that the critical distance is 400 metres.

8.3 Donegal

The preceding analysis was repeated on the distribution of 137 chambered tombs in County Donegal (Ó Nualláin 1983) using the same perturbation distances—500 metres, 1000 metres and 1500 metres. The outcome of the cluster analyses is shown in Fig. 8.8. Again, a discrepancy exists between the real and perturbed curves at low distances. When we examine the difference between the curves directly it is apparent that, unlike Leitrim, there are three key distances (Fig. 8.9). These are approximately 200 metres, 1 km, and 2.5 km.

In order to investigate the two later peaks in greater detail each group of tombs with

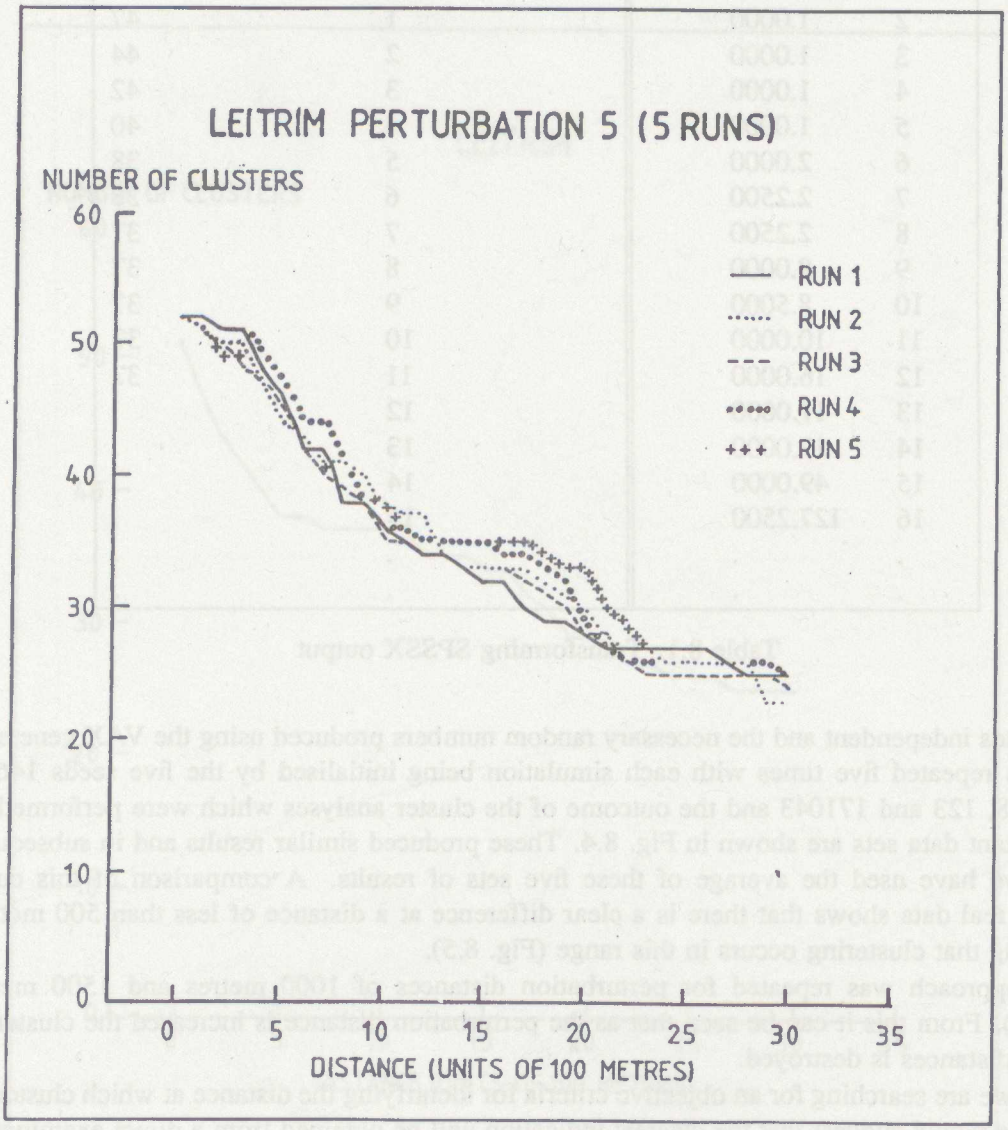

Fig. 8.4: Leitrim: data perturbed by 500 metres

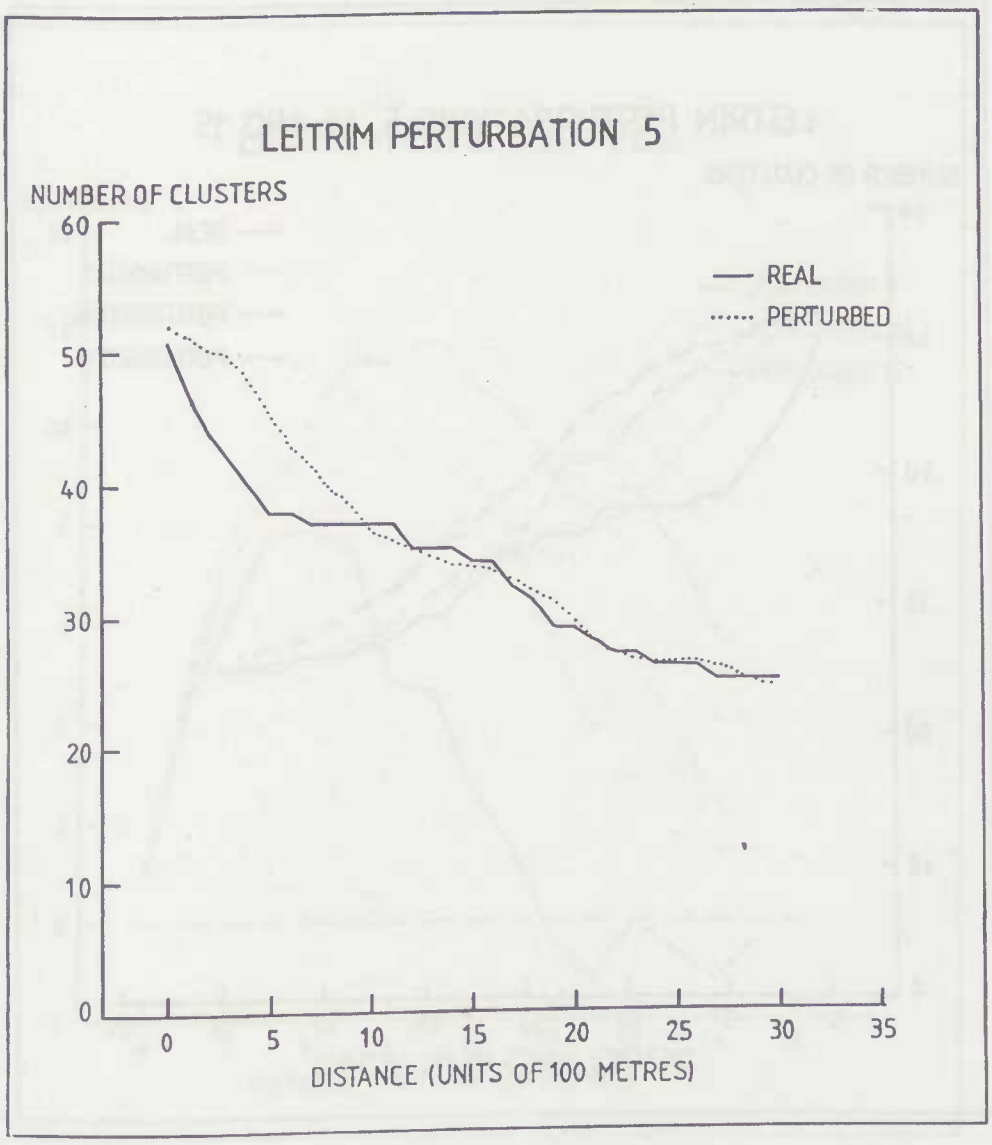

Fig. 8.5: Leitrim: comparison of real data and data perturbed by 500 metres

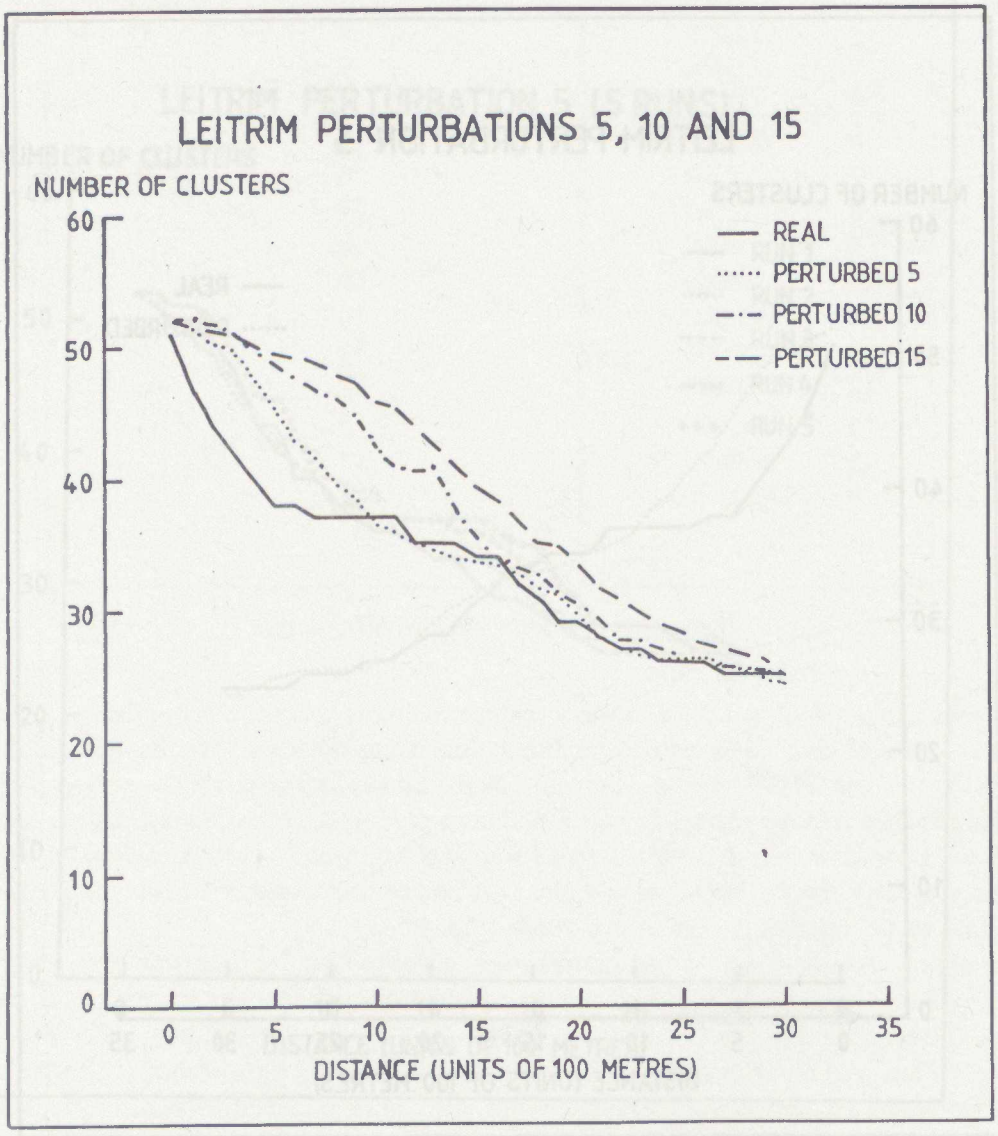

Fig. 8.6: Leitrim: comparison of real data and data perturbed by 500, 1000 and 1500 metres

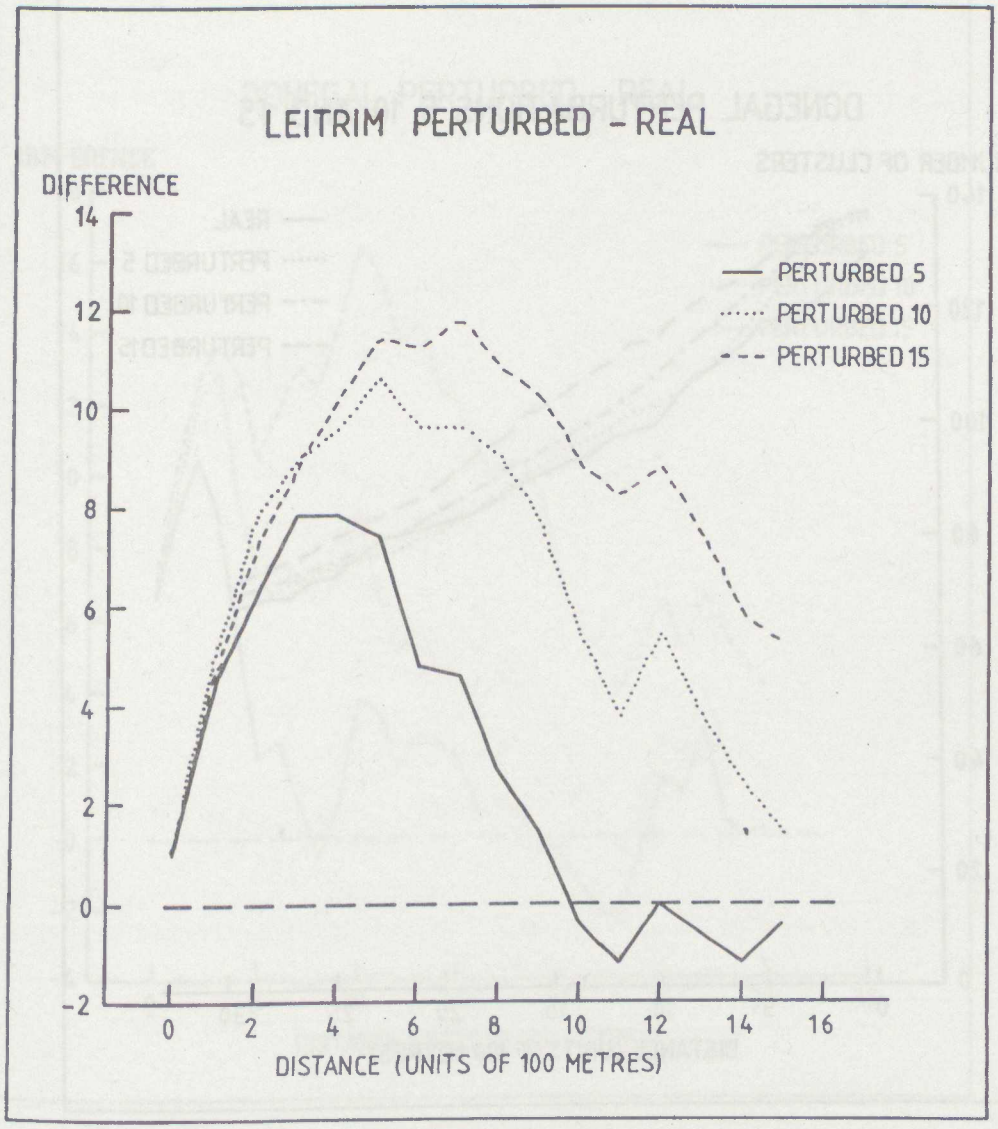

Fig. 8.7: Leitrim: differences between real and perturbed data

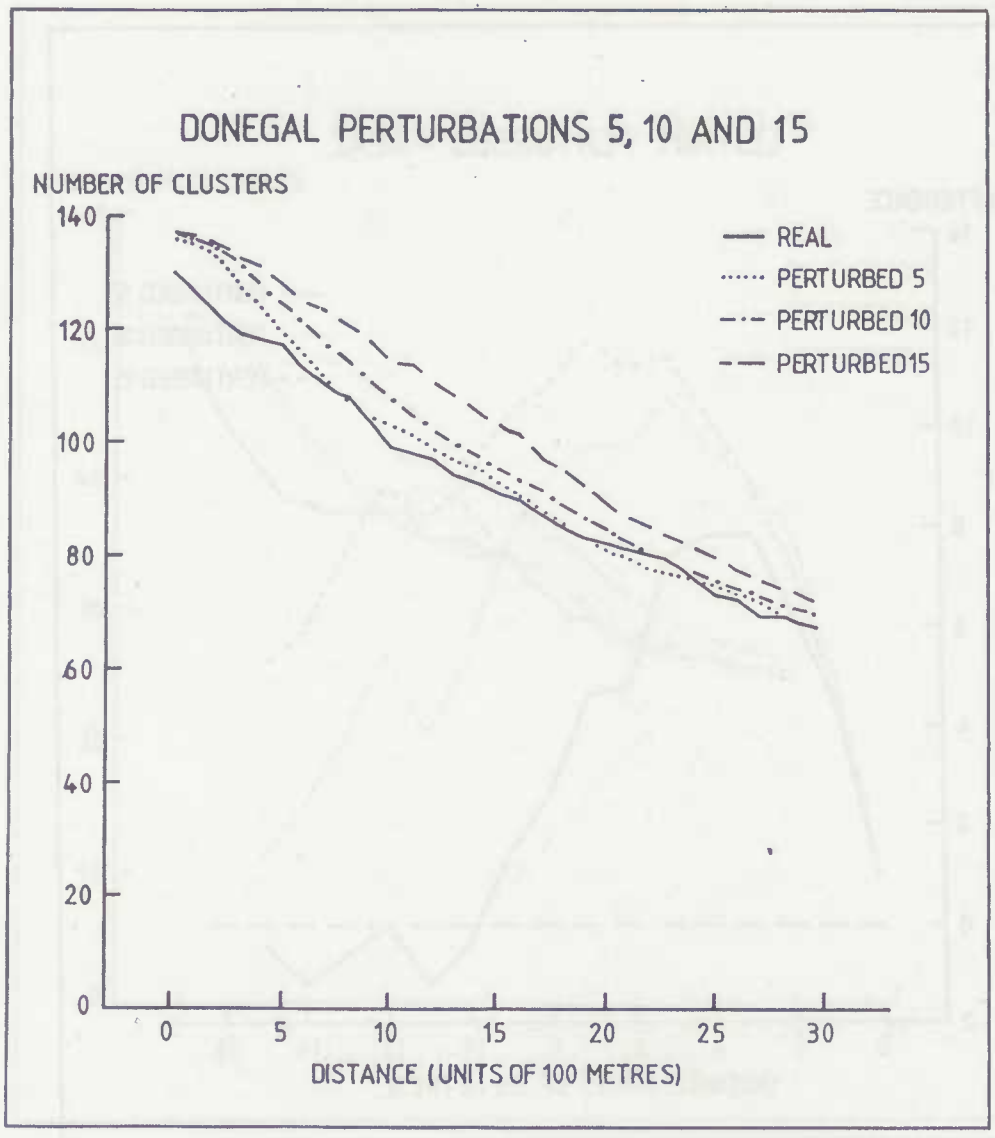

Fig. 8.8: Donegal: comparison of real data and data perturbed by 500, 1000 and 1500 metres

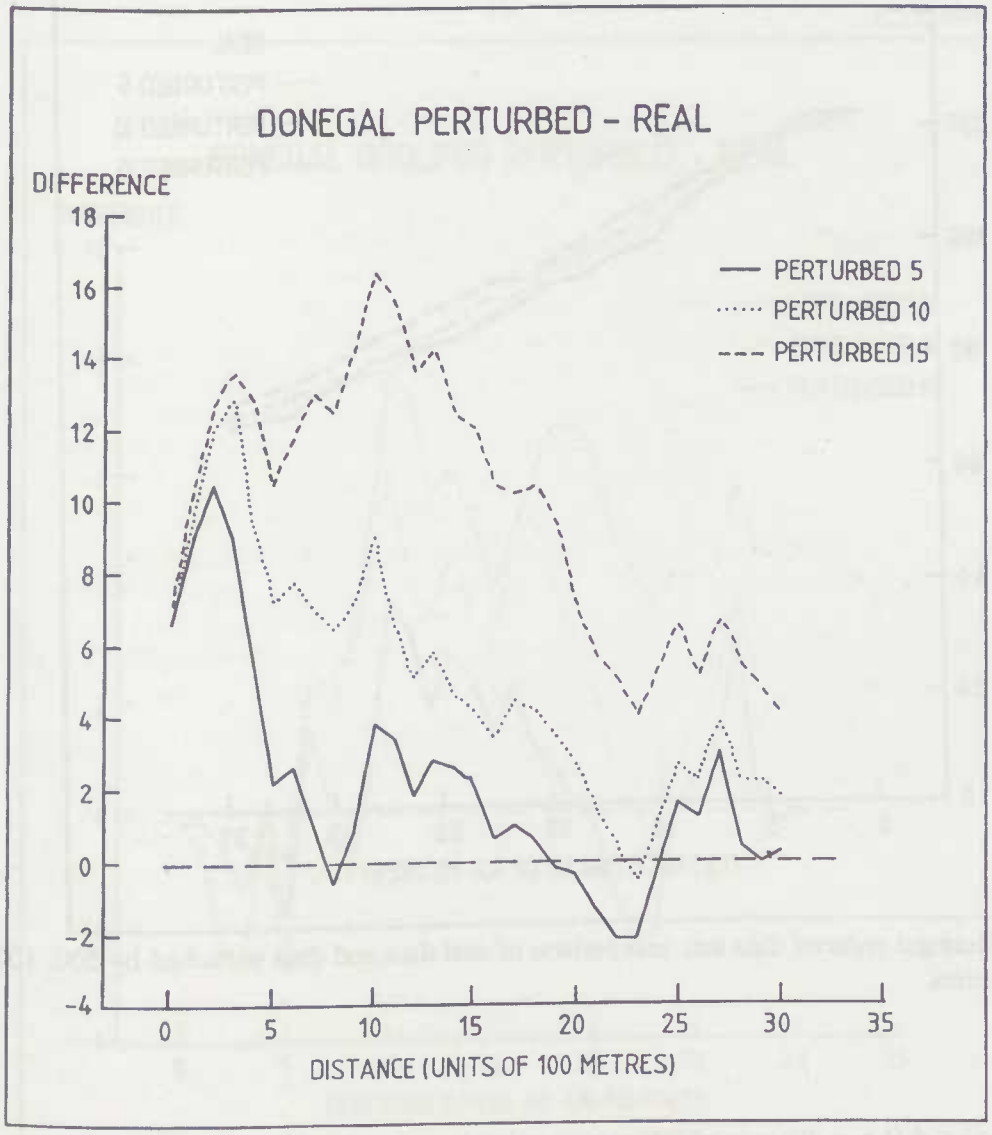

Fig. 8.9: Donegal: differences between real and perturbed data

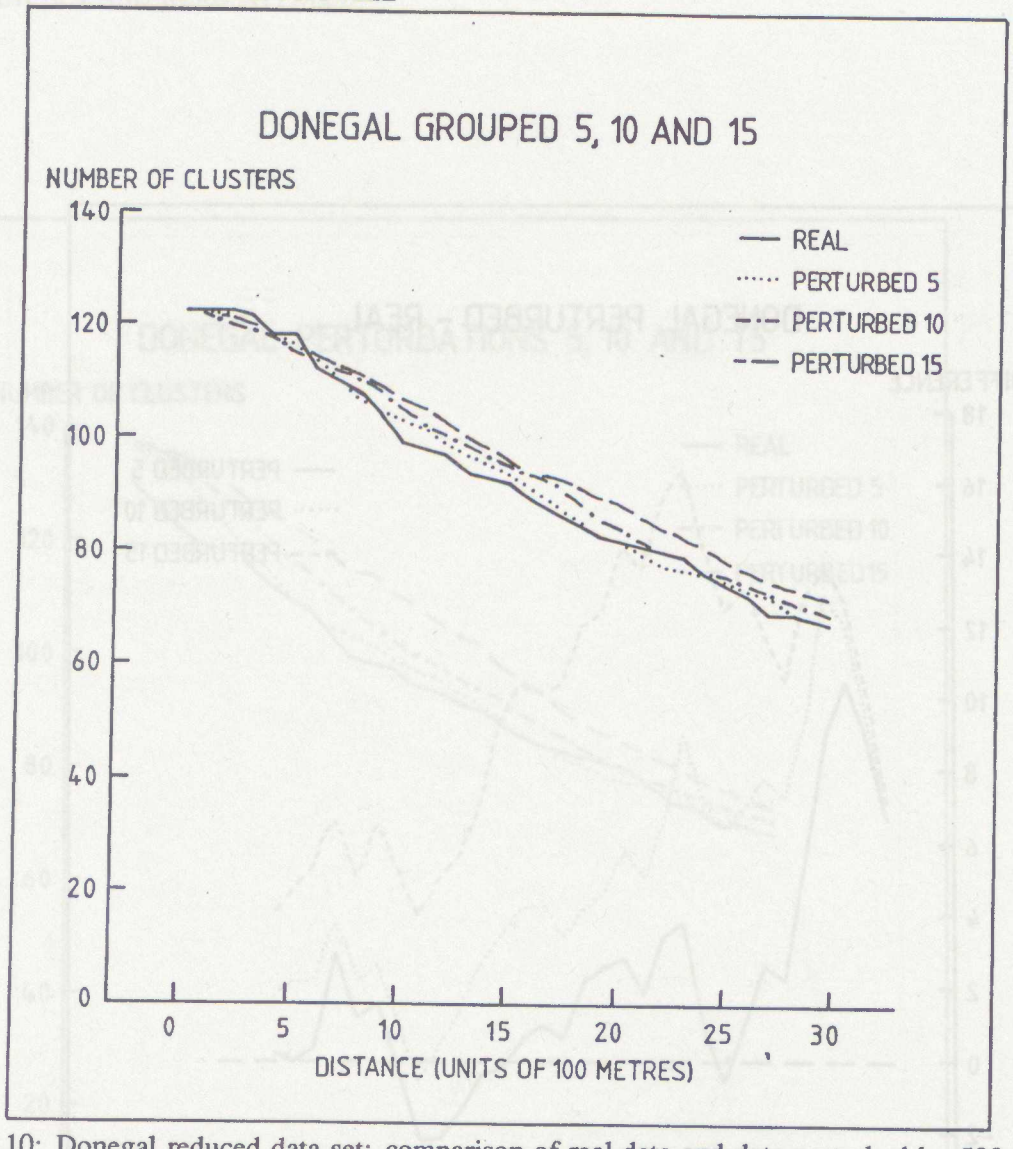

Fig. 8.10: Donegal reduced data set: comparison of real data and data perturbed by 500, 1000 and 1500 metres

coefficients of ≤ 4 (*i.e.* a distance of 200 metres or less) was replaced by a single location. This reduced the data set by fifteen from 137 tombs to 122 locations. As would be expected, the cluster analyses of the reduced real and perturbed data sets show no differences at low distances (Fig. 8.10). However, there is an indication of a divergence at 1 km where the gradient of the real curve decreases.

A direct examination of the differences between the real and perturbed results confirms the existence of the clustering at 1 km (Fig. 8.11). In the case of the 1500-metre perturbation, the two peaks have become vague because the distance between them is comparable to the size of the perturbation. A proper interpretation of these two peaks should take into account the probable variation from the generation of the perturbations. Further statistical analysis will be the subject of future work.

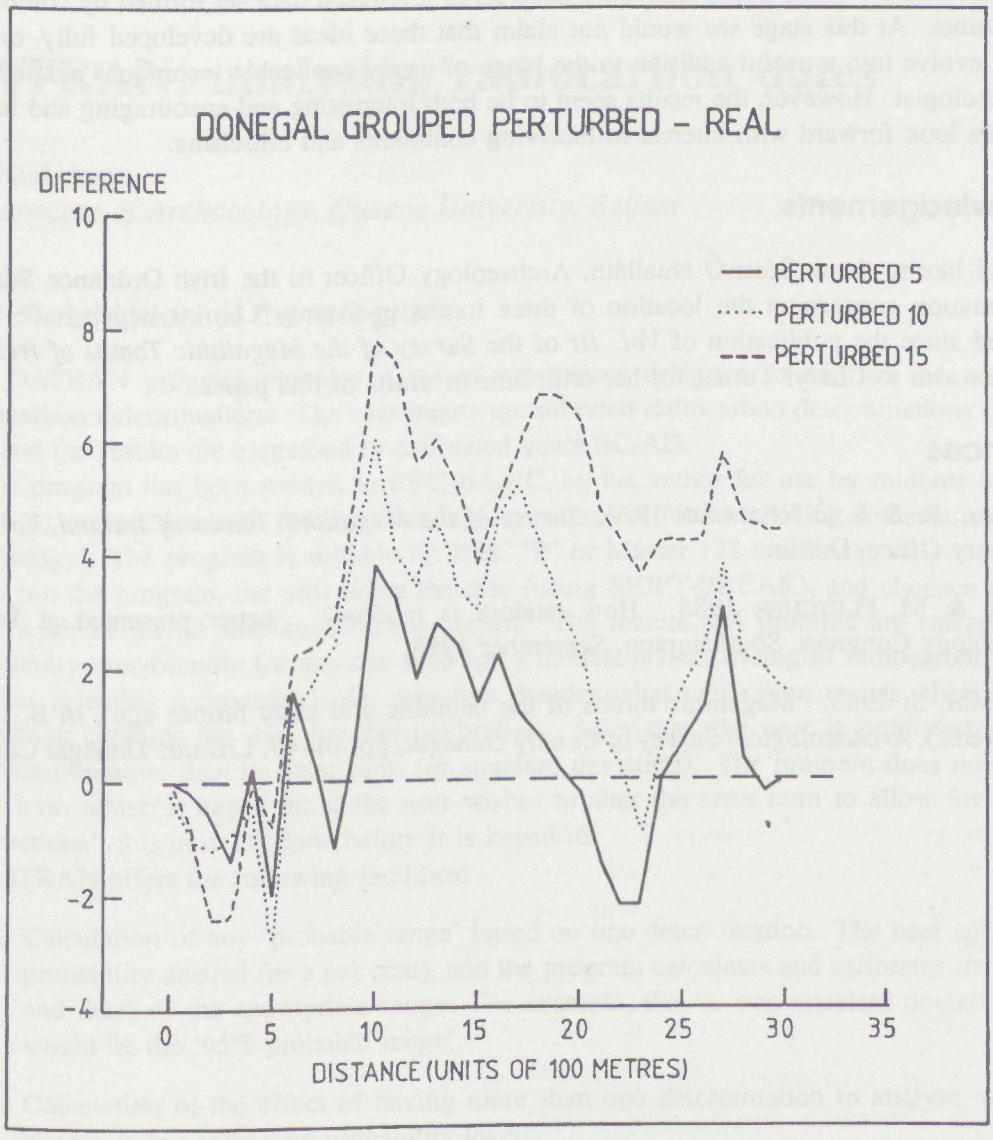

Fig. 8.11: Donegal reduced data set: differences between real and perturbed data

8.4 Conclusion

In conclusion, underlying the preceding analysis are two main elements: the application of a standard technique (in this case cluster analysis) to real and perturbed data sets followed by a comparison of the results; and a secondary analysis of a reduced data set formed by combining 'close' points. At this stage we would not claim that these ideas are developed fully, or that they will evolve into a useful addition to the range of easily applicable techniques available to the archaeologist. However, the results seem to be both interesting and encouraging and in this context we look forward with interest to receiving comments and criticisms.

Acknowledgements

We would like to thank Séan Ó Nualláin, Archaeology Officer to the Irish Ordnance Survey, for information concerning the location of three tombs in County Leitrim which have been discovered since the publication of *Vol. III* of the *Survey of the Megalithic Tombs of Ireland.* Our thanks also to Cheryl Turner for her criticisms of drafts of this paper.

References

DE VALERA, R. & S. Ó NUALLÁIN 1972. *Survey of the Megalithic Tombs of Ireland, Vol. III,* Stationery Office, Dublin.

LOCK, G. & M. FLETCHER 1986. How random is random? paper presented at World Archaeology Congress, Southampton, September 1986.

Ó NUALLÁIN, S. 1983. 'Megalithic tombs of the neolithic and early bronze age', *in* B. Lacy *et al.* , (eds.), *Archaeological Survey of County Donegal,* pp. 14–49, Lifford: Donegal County Council.