# Orientations of Bronze Age Houses: Statistical Analysis of Directional Data

Nick Fieller,
Department of Probability and Statistics,
University of Sheffield.

Sue O'Neill,
Department of Prehistory and Archaeology,
University of Sheffield.

## 1. Introduction.

Archaeological data frequently include measurements of variables which are inherently angular. Common examples are the compass orientations of graves in a cemetery or of doorways of houses and the angles of wear marks on teeth or cutting implements. All too commonly the statistical analysis performed on such data ignores the circularity of the variables and it is pretended that these angular measurements are just ordinary Normal data amenable to the standard techniques of t-tests, regression, principal components, clustering or whatever the situation demands. It is particularly tempting to do this when the angular variable is just one among a large number of variables measured on the same items. This is especially so when the analysis is performed by computer using a standard package. Most statistical analysis packages do not have special facilities for analysing directional data, notable among these are the most frequently used packages such as SPSS, SAS, GENSTAT, MINITAB, GLIM, CLUSTAN, MDSCAL and BMDP.

The two essential features of directional data are firstly that the origin of the measurement scale is arbitrary and secondly that the scale is inherently circular. The first property requires that any analysis should be invariant under changes of origin: conclusions should be the same whether directions are measured as deviations from north or as deviations from east or from any other arbitrary reference direction. An illustration of the second property is that a deviation of $1^0$ anticlockwise is the same as one of $359^0$ clockwise, and both are the same as a deviation of $361^0$ anticlockwise. In particular, two directions of $5^0$ and $355^0$ (measured, say, clockwise from north) are just as 'close together' as two directions of $30^0$ and $40^0$ : they are both 10 apart. Any form of analysis which ignores these features can produce misleading results.

We present here an analysis of a particular set of directional data. The measurements involved are the orientations of the doorways of Bronze Age houses in the Plym Valley on Dartmoor, taken from a large data set containing values of many variables recorded for each house. Our aim is to illustrate how quite simple techniques of data display and analysis, taking into account the inherent circularity of the measurement scale, can provide useful insight into the archaeological questions raised. We also shew how ignoring the circularity could have led to contradictory conclusions. Of course, the analysis of just the directional measurements forms only a part of the complete investigation of the available data. A full account of this more extensive investigation is given elsewhere (O'Neill, 1982).

The techniques of the statistical analysis we use are well-known in the statistical literature; many of the graphical displays of the data are our own. A fundamental reference for the would-be analyst of directional data is the book on the subject by Mardia (1972). A more recent work is the book by Batschelet (1982). The latter is designed for the non-mathematician and is consequently more orientated towards the applications to data analysis rather than to the distributional theory. However, in many cases Batschelet refers the reader to Mardia (1972) for the finer details and tables of critical values of the tests available.

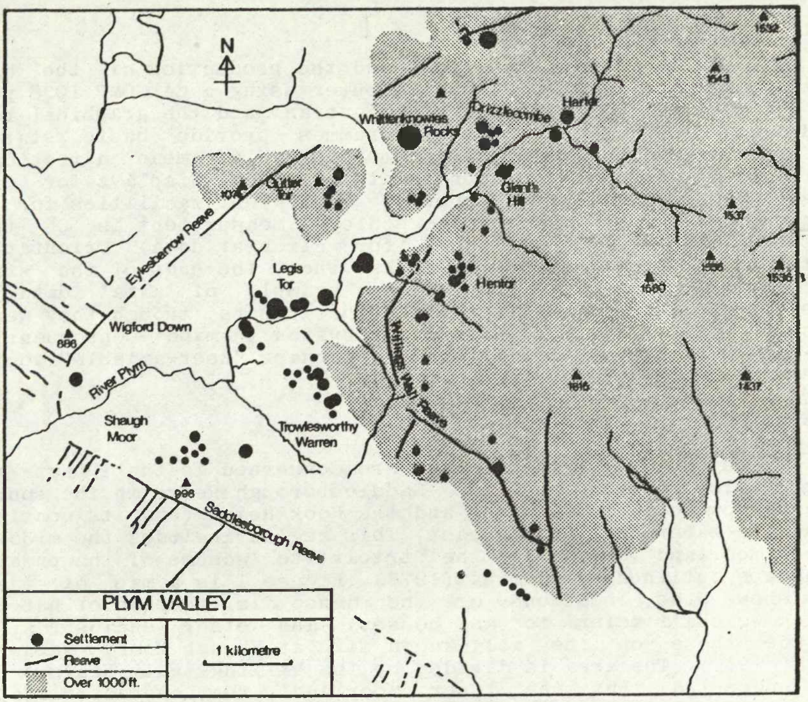



FIGURE 1

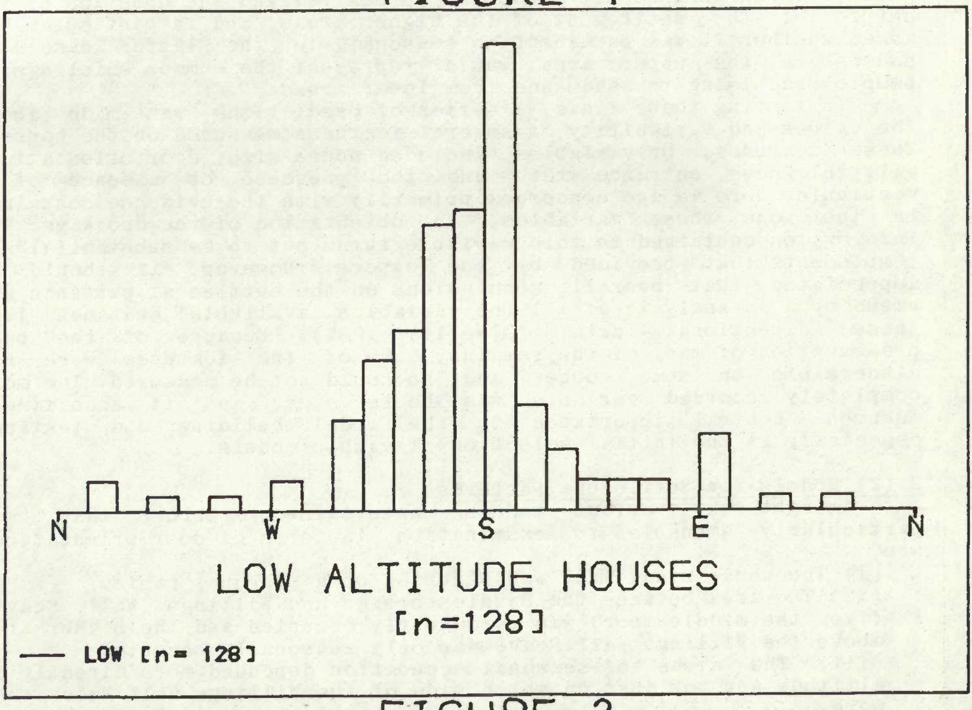



FIGURE 2

All the statistical analyses and the production of the diagrams were carried out on a PRIME 750 computer using a CALCOMP 1036 plotter. The programmes are written in Fortran and the graphical routines utilise the GINO library. The programmes provide basic statistical estimations and tests of homogeneity of mean directions and concentrations as well as appropriate graphical displays for circular data. Similar programmes provide equivalent facilities for 'axial' data, that is data for which the scale of measurement is 0° to 180° (rather than the 0° to 360° for circular data). Orientations of features such as striations on rocks, where the ends of the striation are indistinguishable, provide an example of 'axial' data. All of these programmes are available from the authors, though they were not written with the aim of wide distribution in mind. Consequently they may not be entirely "user-friendly"; perhaps "user-amiable" would be a more exact description.

## 2. Archaeological Background.

### 2.(1) Preliminaries

The study area with which we are concerned is that part of the Plym Valley bounded by the Saddlesborough Reave in the south, the Eylesbarrow Reave in the west and the Rook Reave (and its continuation along the watershed) in the east. This region includes the middle and upper moorland areas in the tripartite scheme of the prehistoric landscape outlined by Fleming (1978). Figure 1 is a map of the area and shews the locations of the dense distribution of hut circles (which we shall refer to as houses) and other enclosures. These include those on the well-known site at Shaugh Moor, (Wainwright & Smith, 1980). The area is divided by the Willings Wall Reave. Fleming has suggested that the lower moorland area, characterised by the 'parallel reave' systems and scattered settlements (based on river valleys), was permanently occupied. He has revived the question of the nature of the settlement of the higher areas, and in particular has asked whether it was permanent or seasonal. In the latter case the houses in the upper area would represent the summer shielings of people practising transhumance from lower areas.

In testing these ideas, a series of predictions was made about the values and variability of several features measured on the houses. These features, or variables, included house size, door orientation, wall thickness, entrance width and the presence or absence of a vestibule. Here we are concerned primarily with the evidence contained in just one these variables, the orientation of the doorways. The information contained in this variable turns out to be substantial and complements that provided by the others. However, it should be appreciated that overall conclusions on the settlement patterns are based upon an analysis of all the variables available and not just these directional data, (O'Neill, 1982). Because of the poor preservation of many of the remains, some of the features were not discernible on some houses and so could not be measured. The most completely recorded variable was house size, and it accordingly assumed central importance in the model building and testing, especially in the initial selection of viable models.

### 2.(2) Models for Settlement Patterns

Amongst the various models entertained, three that are particularly amenable to examination in terms of door orientations are,

(i) The whole study area was occupied on a seasonal basis.

(ii) The area between the Saddlesborough and Willings Wall Reaves (i.e. the middle moor) was permanently occupied and the higher area above the Willings Wall Reave was only seasonally occupied.

(iii) The areas of seasonal occupation depended more directly on altitude and not just on which side of the Willings Wall Reave they were.

Here, for the sake of illustration, we will concentrate on the validity of model (ii), and only brief attention will be given to the others. A full account of the investigation into competing models is given by O'Neill (1982).

Initial investigations led to the discarding of model (i), principally because the distributions of house sizes and doorway orientations did not shew the 'expected' consistency of small size and lack of consistency of preferred direction respectively.

Secondly, model (ii) was examined to see what implications it held for the data collected. In accordance with this model, the houses above Willings Wall Reave would be expected to represent summer shielings and would therefore be unlikely to be capable of permanently housing a nuclear family. Contrastingly, the lower area would be expected to have larger family-sized houses. This prediction was supported by straightforward statistical analysis, O'Neill (1982).

A second stage in testing this model was the investigation of the many other variables recorded, in particular the doorway orientations. Predictions about these these latter data were based upon a number of factors. Worth (1953: 106) says, in relation to the doorway orientations of houses on Dartmoor

> "The entrances are invariably placed facing south to south-west, the builders being evidently of the opinion that even the rains and gales from that quarter were preferable to the storms from the north and nor'-nor'west which, on Dartmoor, are the greatest enemies of all living things."

Fox (1957: 22) noted, following excavations at Dean Moor, Dartmoor, that houses were protected from the south-west weather, and entrances avoided the north-west. Fleming (1979: 125) suggested,

> "Insight into the seasonality of the moorland occupation might be derived from a closer study of the orientations of house entrances......"

At Holme Moor he noted that houses face east or south-east even if it meant that they faced uphill and said,

> "......it might be argued that these houses were occupied in winter and that the bitter north-westerlies were regarded as the most fearsome winds."

He pointed out that the Dean Moor houses were orientated in different directions although Fox had noted that the north-west was avoided. This would support an argument for summer occupation.

It was suggested, therefore, that a 'tighter' grouping, about a preferred direction, of orientations of houses might be expected in those areas occupied during the winter (so as to preserve comfort). Contrastingly, houses occupied just in the summer would be less consistently orientated since the climatic factors would no longer be the overiding consideration determining orientation.

## 3. Statistical Analysis.

### 3.1 Graphical Displays

An essential preliminary to a formal statistical analysis of the data is an exploratory and informal examination of its structure. The purpose is to reveal any unsuspected features in the data (outliers, multimodality etc.) and to convey informally whether the probabilistic models fitted in the course of the succeeding formal analysis are appropriate. For data with a simple structure a graphical display is ideal. For univariate linear data, construction of a histogram is an obvious first step.

Figure 2 shews a histogram of the orientations (measured anti-clockwise from north) of the houses below Willings Wall Reave. (Here and later these will be referred to as the 'Low Altitude' houses, and those above the Reave as the 'High Altitude' houses). This histogram shews the predominance of directions between south-west and south-east. Apart from an apparent slight skewness in the distribution, this histogram might lead one into a false sense of security and a belief

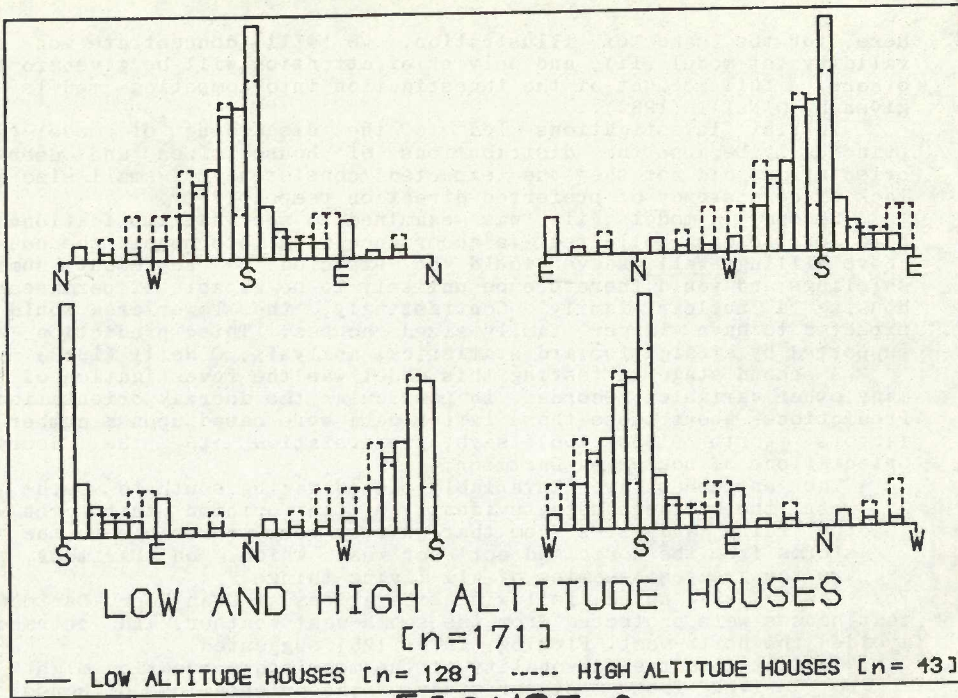



FIGURE 3

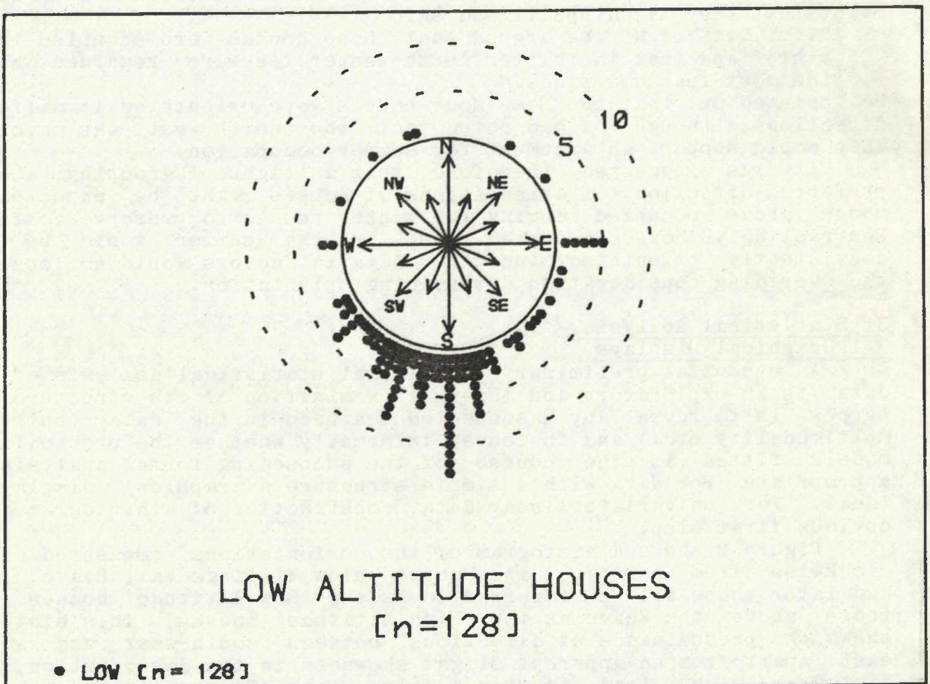



FIGURE 4

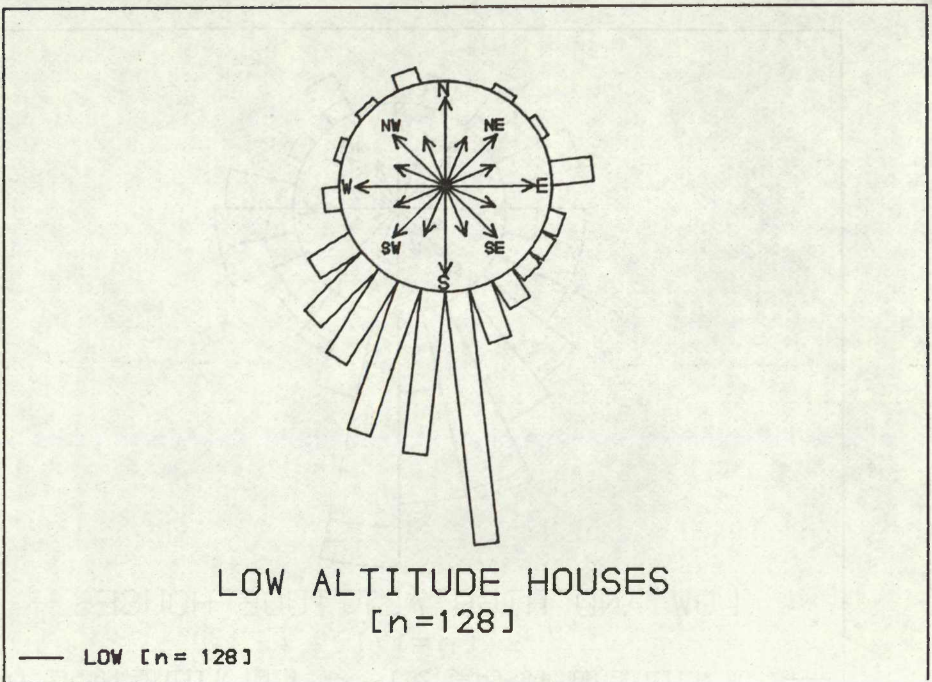



FIGURE 5

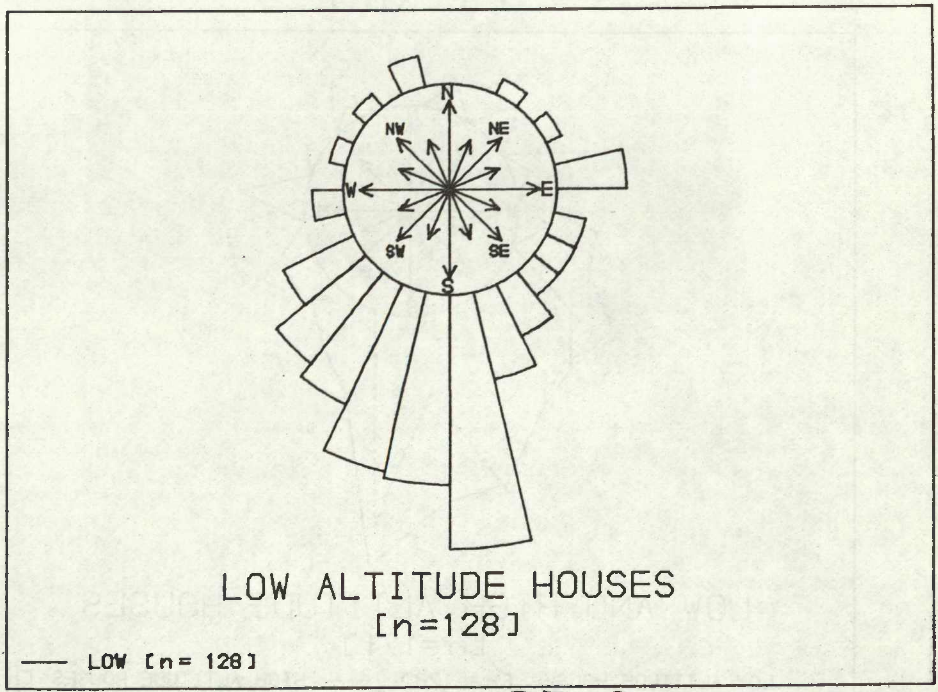
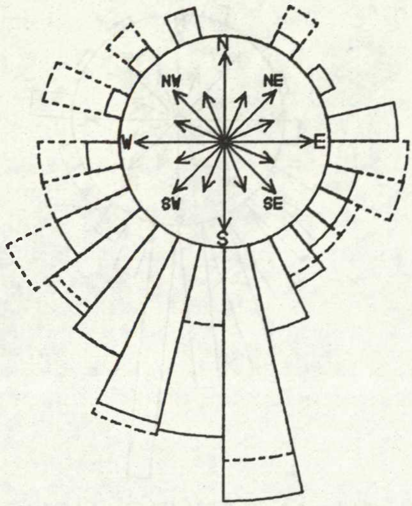


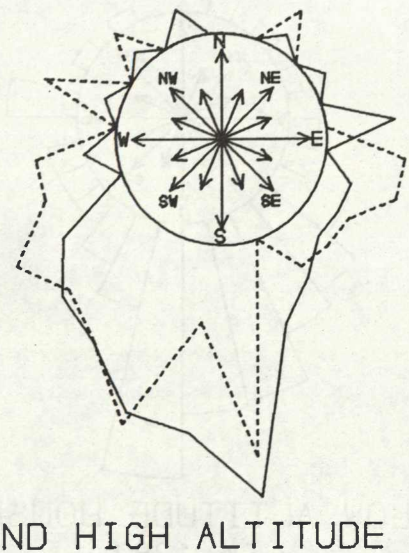

FIGURE 6




LOW AND HIGH ALTITUDE HOUSES
[n=171]

—— LOW ALTITUDE HOUSES [n= 128] ----- HIGH ALTITUDE HOUSES [n= 43]

FIGURE 7




LOW AND HIGH ALTITUDE HOUSES
[n=171]

—— LOW ALTITUDE HOUSES [n= 128] ----- HIGH ALTITUDE HOUSES [n= 43]

FIGURE 8

that the orientations of these houses could be modelled as arising from Normal distributions. However, this display takes no account of the circularity of the data and the fact that north was taken as the base line for measurement only because magnets happen to point in that direction.

Figure 3 shews histograms of the data taking each of the four cardinal directions in turn as base line. Illustrated here are superimposed histograms for both sets of data; the area in each 'column' is taken proportional to the relative frequency in the corresponding interval, (rather than the actual frequency. This ensures that the total area under each histogram (for low and high altitude houses) is the same, thus allowing comparison of distributional shape without dependence on total frequency. This means that the vertical scale is not (and indeed must not) be annotated with frequencies. Similar remarks apply to the circular histograms illustrated below. Figure 3 emphasises just how dependent an ordinary histogram display is on the choice of origin. Had east or west been chosen as base direction, then a very skew distribution would have resulted, while measuring orientations from south results in an alarming bimodal picture. What is needed is a picture which preserves the inherent circularity and which shews that the 'two' peaks in the third histogram of figure 3 are in fact adjacent.

Figure 4 shews a frequency plot, or 'circular blob-plot', for the low altitude houses. Each blob corresponds to one house and is placed in the corresponding position relative to the compass rose. A numerical radial scale is indicated by the dotted circles. Such a display is useful for shewing isolated outliers and unsuspected clusters worthy of further investigation (e.g. the five houses pointing due east). However, to compare two different sets of data an analogue of a histogram is required; one which illustrates relative frequencies. Figure 5 shews the effect of wrapping the histogram of figure 2 around a circle. It might be termed a 'Wrapped Histogram'. However, the 'spikyness' of this display detracts not only from its aesthetic appeal but also from its usefulness in suggesting an appropriate smooth probability distribution underlying the data. A smoother diagram results by constructing wedges rather than rectangles to represent relative frequencies, figure 6. Here the area of each wedge is proportional to the relative frequency. Thus, the radius of the outer circumference of the wedge is

$$\sqrt{r^2 + (R^2 - r^2)f/h}$$

where r is the radius of the inner circle containing the compass rose, R is the maximum radius allowed, f the relative frequency in the corresponding angular interval, and h is the maximum relative frequency in any interval. This choice is preferable to making the radius proportional to relative frequency, which would result in the areas being proportional to the squares of the relative frequencies and a consequent over emphasis to the human eye of the larger ones. A display such as figure 6 is properly termed a 'Circular Histogram', though similar diagrams are also known as rose or polar-wedge diagrams. Figure 7 shews superimposed circular histograms for the low and high altitude groups of houses, and figure 8 shews superimposed relative frequency polygons, obtained by connecting the mid-points of the outer arcs. The latter diagram is a particularly useful guide when the number of groups to be compared is large.

## 3.2 Formal Analysis

The considerations of section 2(ii) shew that to test the tenability of model (ii) for the settlement pattern (permanent occupation in the lower group, seasonal in the upper) it is appropriate to investigate the variability of the orientations of the houses in the two groups. The model predicts that the variability of the second group will be greater than that of the first.

Before looking at appropriate circular models, it is illuminating to see what would happen if the circularity were ignored and the data analysed as if they were Normal samples. In this case the appropriate procedure is to compare the ratio of the two sample variances with the F-distribution on, in this case 127 and 42 degrees of freedom. Doing this with the deviations measured from north would allow one to conclude, at about the 1.5% level of significance, that the variation of the upper group was indeed greater than that of the lower. However, if either east or south were taken as the reference direction, then the opposite conclusion would be reached, (i.e. that the lower group were more variable than the upper), though only at about the 10% significance level. Taking west as the reference direction would restore the conclusion of greater variability in the upper region (at the very impressive significance level of 0.05%). Examination of figure 3 explains these paradoxical results. Of course, in this simple case it is clear from these four histograms that north is the preferable base direction since there are few data points in that area. In more complex cases this would not be obvious, and it is not often that the data analyst has the luxury of such extensive examination of just such a small fraction of his data.

Examination of figure 7 shews that in an intuitive sense the variation of the upper group is greater than that of the lower; there is less of a peak towards south and more of a spread towards east and west. The orientations of the high altitude houses seem more evenly spread around the points of the compass. To test this formally requires specification of a distributional model for the data. For linear data a standard F-test presupposes the data arise from Normal distributions with density function depending upon two parameters, $\mu$ and $\sigma$ describing the mean and variance. Similarly, for circular data of the type encountered here, one can presuppose that the data arise from von Mises distributions which again depend upon two parameters, $\theta$ and $\kappa$ describing in this case the mean direction and the concentration (i.e. the reciprocal of the variance) about this mean direction. Examination of the circular histograms in figures 6 & 7 indicate that von Mises distributions at least approximate reality in this case. (If this were not so then the equivalent, though weaker, non-parametric procedures would be appropriate). A test of equality of concentration parameters is available, Mardia (1972: 161), and applied to these data it allows one to conclude, at about the 3% level of significance, that indeed the 'concentrations' of the two parent populations for the upper and lower groups of houses do indeed differ, the upper group being the more variable. Mathematical details of the tests and estimations of parameters are not given here.

In passing, it should be explained that a test of equality of mean directions of the two groups is not appropriate in this case for two reasons. The first is that it is not demanded by a test of the archaeological model and has no relevance to it. The second is that a 'mean' or 'preferred' direction is less well defined if the dispersion is large, and the analysis above indicates that this is so in the case of the upper group of houses.

## 4. Conclusions.

One cannot be certain about the climate during the Bronze Age, but in the British Isles it seems to have been dominated by westerly and northwesterly winds, Lamb (1977:385). Today, although south-westerlies prevail (Brunsden & Gerrard, 1971) the north-westerlies are noted for their ferocity. Evidence from the stunting of trees (Worth, 1953: 94) corroborates this.

The statistical examination of the relative consistency of orientations in the two regions divided by the Willings Wall Reave indicates that the houses in the lower area were more consistently directed away from these unpleasant winds, while in the upper area this was not so. A similar examination, designed to test model (iii), where the houses were grouped solely on the basis of altitude, failed to reveal such differences, O'Neill (1982). The statistical procedure for this examination is the circular analogue of Bartlett's homogeneity of variance test.

Thus, careful analysis of the circular data on orientations of doorways lends credible support to the theory of permanent occupation below the Willings Wall Reave, and only seasonal occupation above it. Ignoring the circular structure could have led to erroneous conclusions. This is an illustration of a more general point; it can be grossly misleading to ignore the circularity of such data. In particular, mixing angular and linear data (e.g. when calculating similarity coefficients based on Euclidean distances) could produce bizarre results.

Examination of figure 6 might suggest that the 'preferred' orientation of the houses in the lower altitude group is not directed exactly away from the inclement winds. Recent excavations at Shaugh Moor (Wainwright & Smith, 1980) have revealed three examples of 'porches' or vestibules, which follwed the arc of the house to the east of the entrance. Porches are difficult to distinguish by fieldwork alone. If house doors and porch exits in the valley were maintained in house plan consistency in relation to each other, then the position of the porches and the direction of their exits help explain this apparent discrepency.

## 5. References.


Batschelet, E. 1981 — CIRCULAR STATISTICS IN BIOLOGY. Academic Press, London and New York.

Brunsden, D. & Gerrard, J. 1971 — 'The Physical Environment of Dartmoor' DARTMOOR: A NEW STUDY, ed Gill, C., David & Charles, Newton Abbot

Fleming, A. 1978 — 'The Prehistoric Landscape of Dartmoor. Part 1: South Dartmoor'. PROCEEDINGS OF THE PREHISTORIC SOCIETY **44**:97-123.

Fleming, A. 1979 — 'The Dartmoor Reaves: Boundary Patterns and Behaviour Patterns in the Second Millenium B.C.'. DEVON ARCHAEOLOGICAL SOCIETY **37**:115-135.

Fox, A. 1957 — 'Excavations on Dean Moor, in the Avon Valley, 1954-1956. The Late Bronze Age Settlement'. TRANSACTIONS OF THE DEVONSHIRE ASSOCIATION **89**:18-77.

Lamb, H.H. 1977 — CLIMATIC HISTORY AND THE FUTURE, vol 2 of CLIMATE PRESENT, PAST AND FUTURE. Methuen, London.

Mardia, K.V. 1972 — STATISTICS OF DIRECTIONAL DATA. Academic Press, London and New York.

O'Neill, S.H. 1982 — (in preparation)

Wainwright, G.J.& Smith, K. 1980 — 'The Shaugh Moor Project: Second Report- The Enclosure'. PROCEEDINGS OF THE PREHISTORIC SOCIETY **46**:65-127.

Worth, R.H. 1953 — DARTMOOR. David and Charles, Newton Abbot.