# COMPUTER MAPPING OF SPECIES DISTRIBUTION IN A COUNTY FLORA

By J. G. HAWKES
*Department of Botany, University of Birmingham*

B. L. KERSHAW
*Computer Services, University of Birmingham*

and R. C. READETT
*Birmingham Natural History Society*

ABSTRACT

Methods of computer mapping of distribution, habitat and frequency data of vascular plants in Warwickshire are described. A simplified mapping scheme for bryophytes is also outlined.

The primary data for vascular plants are transferred in the form of coded information to punched tape by means of a teleprinter and checked for errors. The data are then processed in the Birmingham University KDF9 computer by means of three separate programs, each taking the data forward to the next stage by means of magnetic tape.

Intermediate print-outs give a preliminary statistical treatment of the data, as well as a simple sketch map, which can be printed on the teleprinter or line printer.

The end product of the final program is a tape designed to give the necessary instructions to an incremental graph plotter. This draws symbols in the correct positions indicative of the habitat or habitats in which the species has been recorded. Two grades of frequency are shown by means of different line thicknesses. The plotted map will be reproduced with an overprinting in another colour, indicating the county boundary and various other features, as well as grid lines and border.

A series of bryophyte maps without habitat and frequency information but with data on fruiting will be printed on the teleprinter.

The wider applications of these mapping techniques are briefly discussed.

## INTRODUCTION

The hand plotting of distribution data, whether of plants or of any other class of organisms, can be both tedious and expensive. If the data to be mapped are at all complex, errors may easily occur, which cannot entirely be eliminated, even after the expenditure of even more time and effort.

Within the last decade high speed electronic computers and associated equipment have become generally available and have made it possible for data to be stored with accuracy, and sorted and retrieved in a wide variety of ways. It would seem desirable, therefore, to use such equipment to sort and store distributional data, provided they can be presented to the computer in an appropriate way, and to devise systems by which they can be printed out in map form.

An excellent example of the use of electronic equipment in the production of distribution maps in this country is the B.S.B.I. Distribution Atlas (Perring & Walters 1962). The original maps for this publication were printed on a tabulator from data supplied to it from I.C.T. punched cards. The symbol on the map, which was a filled-in circle (except for some other symbols put in by hand afterwards), was positioned by making use of numerical co-ordinates taken from the National Grid at 10 km intervals.

Soper (1964, 1966) has discussed several methods of mapping the distribution of the vascular plants of Ontario by tabulating machines and digital plotters. The primary data, consisting of a code reference for the species together with geographical co-ordinates and some other information, were put on to punched tape and later transferred to 80-column I.B.M. punched cards. Soper pointed out that both tabulators and digital plotters with a printing head were capable of printing a limited number of symbols in a reasonably satisfactory manner. On the other hand, a digital plotter with a drawing pen could be made to draw symbols of any shape and size, though it would, of course, need to be 'programmed' for each symbol. He pointed out that the method was not very satisfactory for drawing completely blacked-in symbols, and that smooth curves could not be formed very well because the pen could move only in successive rectilinear steps.

## The Flora of Warwickshire project

A survey of plant distribution in vice-county 38, Warwickshire, was begun in 1950 as a joint project of the Birmingham Natural History and Philosophical Society and the Birmingham University Department of Botany. This has been a co-operative study, involving a team of amateur and professional botanists in many hours of field work over some 16 years. It was begun by Mr. R. C. Readett and Mr. Peter Green. After the latter's move from Birmingham one of the present authors (J.G.H.) took his place, and the organizing panel now also includes Miss D. A. Cadbury and Mr. M. C. Clark.

The scheme was modified in 1954, when the 'basic square' technique of area recording was introduced (Hawkes & Readett 1954, 1963). This involved the detailed survey of 1 km squares as the unit of recording, one square at random being selected from each block of four, or 'tetrad'. An even survey over the whole county was thus feasible with only one quarter of the effort that would have been involved in recording from every square. There are thus 25 tetrads in each 'major square' of 10 × 10 km, which was the unit of recording for the B.S.B.I. scheme referred to above. In addition to this, a more conventional method of recording rare plants was carried out, since they would have been likely to have been missed altogether on the random square method.

The records obtained from the basic square surveys were written on to species cards representing a 10 × 10 km block of 1 km squares, printed at the scale of 1 in. to the mile and hence directly comparable with the Ordnance Survey map of that scale (Hawkes, Readett & Skelding 1955). This method also helped to give an idea of the distribution of the species at any time during the survey, since the records were displayed on a grid system in their correct spatial relationship.

Habitat data and frequency were also recorded on a pre-determined scheme using a system of major habitat groupings (e.g. woodland, waterside, water, cultivated land, etc.); each of these was divided into a number of minor habitats for greater precision (e.g. mixed woodland, oak woodland, conifer woodland, etc.) and suitable lexicographic symbols were devised for each (e.g. WO–m; WO–o; WO–c; for the habitats just given). Frequencies were recorded on the usual simplified system: abundant (a), locally abundant (1a), frequent (f), locally frequent (1f), occasional (o) and rare (r). It was emphasized that the frequencies should refer to each habitat for which the species was recorded, and not to the square as a whole.

From the inception of the scheme it had been hoped that it would be possible to transfer the habitat data in some way on to species distribution maps by means of symbols that would give a general impression of the community or communities in which a species was to be found throughout the county. A number of problems had to be solved here; in the first place, none of the symbols could be of the blacked-in type, since they would need to be presented in various combinations in the same square. Each must therefore be clearly visible, even when combined with several of the others. It was also necessary to devise symbols which would give easily recognizable patterns over the map as a whole. Thus, symbols which relied for their interpretation on their *position within a square* would clearly be valueless, since such a scheme would not enable the reader to obtain a general picture at a glance; each square would have to be examined individually. It was further decided to represent the record from each random square in the complete tetrad in which the random square occurred. This would prevent a false impression of distribution being produced as a result of 'clustering' of the random squares. A reasonably satisfactory series of habitat symbols was devised which conformed to the needs outlined above (see Cadbury, Hawkes & Readett 1957). These have been incorporated into the present scheme with some modifications. At that stage (prior to 1964) it seemed impossible to devise a scheme for habitat recording which at the same time showed frequencies within habitats, though a general idea of frequencies could of course be obtained by looking at the spread of the species over the county as a whole.

The major habitat categories (woodland, waterside, water, heath, marginal, grassland, cultivated land and ruderal), eight in number, were, as we have already mentioned, divided into minor ones; there were 34 of these minor habitat categories for Warwickshire, and it was clearly impossible to devise a system of symbols to denote all these when each symbol needed at the same time to be distinguishable in a large number of different combinations. It was therefore considered that symbols showing the eight major categories would be all that could be reasonably expected for maps of this sort.

## Electronic data processing

By 1964 a number of problems urgently needed attention. Since it was hoped to draw maps of some 700 species of vascular plants and 200 species of bryophytes a long and expensive period of map drawing seemed to be ahead of us. Although the bryophyte data were fairly simple, the prospect of drawing by hand the habitat symbols on 700 vascular plant distribution maps was one to which we did not look forward with much pleasure. There were somewhere in the region of 1,000 records a species for the more common species, each of which would need to be mentally interpreted into a symbol by the draughtsman and placed accurately on to the base map.

A further difficulty became apparent about that time. The habitat coding scheme, which was devised in 1950 and came into general use in its present form in 1954, was later felt to contain some errors of arrangement of minor into major habitat categories. Thus, the marginal category was found to be rather artificial from an ecological point of view. Whilst retaining the system for recording purposes, it was decided to keep 'hedgerow' in the marginal category and add 'scrub woodland' to it, at the same time removing all the others. 'Roadside' and 'railway' would fit best with the ruderal category (enlarged slightly to include 'viatical' elements); 'walls' and 'quarries' were removed from marginal and ruderal, respectively, and a ninth major category of 'stones, etc.' was created for them. When this was done, clearly an even more complex mental process would be needed to convert the original code letters on the species card into the correct map symbols.

A final point might be considered here. It was planned at a fairly early stage (see Cadbury, Hawkes & Readett 1957) to include habitat analyses in the Flora; these would show for each of the commoner species the range of habitats in which it had been recorded, expressed as the number of records for each habitat and the percentage figure, as well as the apparent 'preference' of the species for a particular habitat. For the second time in a county flora numerical data would be available for statistical analysis, the first being Good's *Geographical Handbook of the Dorset Flora, 1948*. The information available was sorted in one or two cases by counting occurrences in the different habitats and calculating percentages, as the first stage of the process. The data were extremely interesting, but the labour involved was quite considerable and carried with it as many possibilities of error as the drawing of the maps themselves.

Fortunately, at about this time Birmingham University bought an English Electric KDF9 computer, and it was therefore decided to process the whole of our data on this machine and at the same time to explore the possibilities of computer mapping of the distribution data. Clearly, some kind of graph plotter seemed to be the answer to the mapping problem, where X and Y co-ordinates could be used to position the symbols in the correct places on a previously printed map on continuous stationery. Unfortunately, most graph plotters incorporate a printing head with very small standard symbols. These completely obscure each other if overprinted, only one symbol can generally be printed in each position, and the symbols can often only be printed in a fixed sequence. Furthermore, the symbols available seemed to be far too small for our purposes.

The machine we finally adopted, after many exploratory discussions with a variety of firms, is an incremental plotter with a drawing pen, marketed by Benson-Lehner Ltd., of Southampton. Since these instruments are relatively expensive and we should only be needing one for this particular project, we agreed in April, 1965 to buy time on one of Benson-Lehner's own machines. We should like at this point to record our thanks to the representatives of this firm for their valuable help and co-operation in the project.

## Preparing the data for the computer

Apart from the very necessary programming work,* the first stage in preparing the data for plotting is to put it into a form which can be used by the computer. We had no punched card facilities at Birmingham when this phase of the work started and we therefore decided to use 5-hole punched tape. To this end we bought a Westrex teleprinter, which is in effect an electric typewriter with tape punching and reading facilities.

Punching began in December 1965, and we were fortunate in having the services of a skilled teleprinter operator (Mrs. M. Kershaw) for the whole period, working on a part-time basis, until Easter, 1967. This included not only the punching of the tapes, but also editing them; it must be admitted, in all fairness to the operator, that a lot of the 'editing' was nothing more than adding recently received records to a tape which was in fact a perfect copy of the original data.

The data were taken directly from the species cards, species by species, and in sequence of major square cards, the data being copied exactly and a 'print-out' produced as a by-product of the punching operation (Fig. 1). The print-out was used for checking against the original cards, a tedious exercise carried out by Miss D. Cadbury, Mrs. E. Pickvance, Mrs. Kershaw and others. Thanks are due to them for the dogged way in which they carried out this monotonous but all-important stage of the work.

After punching and checking, the data are ready to be passed to the computer. Three programs are used, but only the first reads the paper tape data. Each in turn operates on the data, taking the work a stage further and passing on partly processed intermediate results to the next stage by means of magnetic tape. The last program produces punched paper tape which is sent to Benson-Lehner and used to control the plotter to draw the final map.

## Entering the data into the computer

The computer first checks the format of the data for four types of error:

(1) Major square number unacceptable.
(2) Minor square number unacceptable.
(3) The same 1 km square mentioned twice.
(4) Observation code unacceptable.

If an error is detected, the computer prints a message to guide the punch operator in correcting the mistake, and switches into a mode of operation in which it does not prepare intermediate results for further processing, but continues to read the data and print messages for any errors detected. In this way all errors in a particular data tape are found in one computer run and can all be corrected before the tape is resubmitted.

If, however, no errors are detected, the computer assembles within its memory an 'electronic image' of the required distribution map. Once this image is obtained, the rest of the computer process is directed to translating it into suitable control signals for the plotter. The first level map image is placed on magnetic tape ready to be further treated by the second program.

The first program, however, does rather more than this. Observations which are accepted and entered on the map image are counted, giving 204 totals, one for each minor habitat/frequency combination. These totals are printed to form the first table in the

* See Appendix, p. 360.

```
3

acer campestre l.    3

04

69.mhrf.
/

05

41.wooo.
43.mhro.
44.mror.wosca.
50.mhro.
54.wosco.mhro.
58.mhro.womo.
60.mhro.
62.mhrr.
65.womlf.
66.mhrf.
68.mhro.
75.womo.mhrf.
76.woscr.mhro.
82.mhro.
89.cfr.
92.mhro.
94.mhrlf.
96.womo.
98.mhrr.
/

06

43.mhro.
44.wosco.
51.mhrf.
54.womo.mhrf.
56.mhrr.
60.mhro.
64.mhrf.
65.mhrf.
67.mhro.
70.mhro.
73.mhrf.
75.mhro.
76.mhro.
77.mhrf.
80.mhro.
82.mhro.
84.womf.
87.mhro.
92.mhro.
93.mhro.
94.mhro.woscr.
98.mhro.
/

07

72.mhrr.
73.wosco.
82.wsdr.
90.mhro.
/
```

Fig. 1. Part of the data for *Acer campestre*, Field Maple, as typed on the teleprinter, showing major square numbers heading blocks of data, minor square numbers at the start of lines, and letter strings representing individual observations (e.g.: mhrf = hedgerow, frequent; wooo = oak wood, occasional; wosca = scrub wood, abundant). The figure 3 at the top represents the B.S.B.I. code number for this species.

statistics (Fig. 2). The second table is simply the first recast in terms of nine major habitats and two frequencies, corresponding to the symbols drawn on the maps.

It should be noted that this first program also incorporates a 'random square sieve' which allows habitat and frequency data through for the simple statistics only if they have been recorded for random squares. Data for non-random squares are therefore suppressed, thus rendering what remains more statistically reliable. Many records have been submitted by collectors for non-random squares as well as for random ones and all these have been used for mapping. However, for showing habitat spread and preference of the species to be mapped it was considered essential not to bias the data in any way towards a preference of a collector for recording in a particular habitat, rather than a preference of a plant for growing in it! This preliminary statistical treatment of the data will not only help us to make useful and meaningful statements on habitat spread and preference in the general text of the Flora; it is also intended to provide a comparative table of habitat records for each of the 700 species that are mapped, giving both actual and percentage records in each of the nine major habitat categories.

*Acer campestre* L. 3

| | A | LA | F | LF | O | R | TOTAL |
|---|---|---|---|---|---|---|---|
| MHR | 2 | 3 | 134 | 24 | 255 | 71 | 489 |
| WOM | | 1 | 7 | 1 | 43 | 14 | 66 |
| MSC | | | 3 | 1 | 15 | 6 | 25 |
| RURO | | | 5 | 1 | 8 | 1 | 15 |
| RURY | | | 1 | | 7 | 3 | 11 |
| WSR | | | 1 | | 8 | 1 | 10 |
| WSC | | | | | 3 | 1 | 4 |
| WSD | | | 1 | 1 | 1 | 1 | 4 |
| WOO | | | | | 3 | 1 | 4 |
| GP | | | 1 | | 1 | 1 | 3 |
| GM | | | 1 | | | 1 | 2 |
| GR | | | | | | 2 | 2 |
| WSP | | | | | | 2 | 2 |
| RUW | | | | | | 1 | 1 |
| STQ | | | | | | 1 | 1 |
| TOTAL | 2 | 4 | 154 | 28 | 344 | 107 | 639 |

| HABITAT | ABUNDANT | | OCCASIONAL | | TOTAL | |
|---|---|---|---|---|---|---|
| | OBS. | PERCENT | OBS. | PERCENT | OBS. | PERCENT |
| MARGINAL | 167 | 26.1 | 347 | 54.3 | 514 | 80.4 |
| WOODLAND | 9 | 1.4 | 61 | 9.5 | 70 | 11.0 |
| RUDERAL | 7 | 1.1 | 20 | 3.1 | 27 | 4.2 |
| WATERSIDE | 3 | 0.5 | 17 | 2.7 | 20 | 3.1 |
| GRASSLAND | 2 | 0.3 | 5 | 0.8 | 7 | 1.1 |
| STONES ETC. | | | 1 | 0.2 | 1 | 0.2 |
| TOTAL | 188 | 29.4 | 451 | 70.6 | 639 | 100.0 |

Fig. 2. Data format check and simple statistics for *Acer campestre*. The input tape for this species contained no errors, so the computer has printed no error messages. In the first block of data the records are listed by minor habitats in one direction and frequencies in the other; general totals are given in the right-hand column and bottom line.

In the second block of data the records are listed in major habitat groups on two grades of frequency (abundant and occasional). Totals and percentages are also shown.

*Note:* these figures are based on random square data only.

The 'sketch map' (Fig. 3) uses each print position to represent one tetrad. If the species has not been observed in that tetrad, 'space' (blank) is printed; if the only observation(s) are occasional or rare, '+'; and if at least one abundant to frequent observation has been recorded, 'X'. There are 10 print positions to the inch across the page, but only 6 lines to the inch vertically; by inserting 'space' in alternate positions we have 5 symbols to the inch horizontally and 6 to the inch vertically, which is as near as it is possible to get to equalizing the horizontal and vertical scales.

The maps are therefore slightly wider than they should be, but are nevertheless most useful in giving a general impression of the distribution of each species throughout the county. They will not appear in the Flora (but see bryophyte maps, mentioned below) and will be used merely to provide a general indication of species distribution when the textual part is being drafted.

Those tetrads through which the county boundary passes are 'marked' in the computer's memory. When such a tetrad is empty, full stop is printed for it instead of space. These full stops build up a suggestion of the county boundary, assisting the interpretation of the sketch map.

## Further treatment of the map images

The map images are taken from magnetic tape by the second program and converted into suitable control signals for the plotter, which are placed on a further magnetic tape. The third program transfers from this magnetic tape to paper tape; this is done for the convenience of the computer room staff and need not be discussed further here.

The plotter consists, effectively, of a pen moving over a sheet of paper, capable of drawing any shape required by incremental movements of 0·1 mm (or 0·005 in. on the non-metric model) in any one of the four cardinal directions. Left to right movements are

*Acer campestre* L. 3

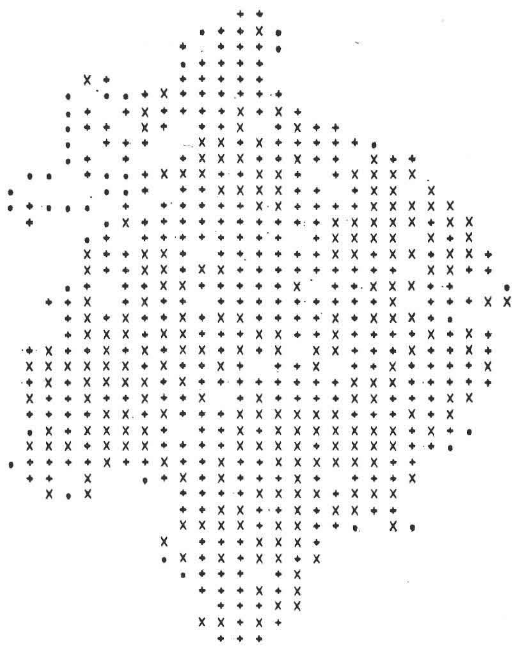


Fig. 3. Interim 'Sketch Map' for *Acer campestre* L.
Printed on the computer's fast output line printer, this map took only three seconds to produce.
+ Occasional to rare in tetrad.
× Frequent to abundant in tetrad.
· Indicates position of county boundary when no record exists in this tetrad.

accomplished by the actual movement of the pen, whilst up and down movements are accomplished by corresponding movements of the paper. Diagonal lines are drawn by combined pen and paper movements, whilst further control signals instruct the pen to begin or to cease marking.

A reasonable solution to the problem of showing frequencies, as well as habitats, on the same map has also been devised, though admittedly not all grades of frequency can be given. This was done by indicating 'rare' to 'occasional' by means of a thin line and 'frequent' to 'abundant' (including 'locally frequent' and 'locally abundant') by a thick one. If these ideas were to be translated into a completely automated system it would clearly be impossible to change pens every time a change of line thickness was required. We therefore decided to use a thin pen throughout and to program it to re-draw the symbol after displacing the pen by two increments if a thick line was required. Since the pen thickness (0·2 mm) is the same as this displacement, the two single thin lines will ideally produce a line twice the thickness of the thin one. In theory there should not be a perfect coincidence of the two lines when diagonals and circles are drawn, but in practice they seem to coincide quite well.

All the symbols (except the circle, which is a special case) are made up of vertical, horizontal or diagonal straight lines (Fig. 4).

Fig. 5 shows how the circles, thick and thin, are built up from short horizontal, vertical and diagonal straight lines. The exact arrangement of these lines was decided by a preliminary computer run to design the best possible circle within the limitations of the plotter.

Two problems arise in preparing the plotter control tape: moving the pen to the required map location for a symbol, and drawing the symbol once the pen is correctly located. The second is easier of solution, since the drawing of a symbol is identical, regardless of its position on the map: 18 standard sets of drawing instructions, one for each symbol, are held in the computer and the correct one is selected after the pen has been positioned.

Positioning the pen to draw a symbol is complicated by the desire to keep pen movement, and therefore plotter time and cost, to a minimum. As each symbol is drawn it is erased from the computer's map image, and the machine then chooses the nearest remaining symbol to the plotter pen's current position to be drawn next.

Besides drawing the symbols, the plotter writes a reference number on each map so that it can be identified and given the correct title in printing; two register or 'fiducial' marks, one at the top left of the map and the other at the bottom right, are made so that the printers can correctly superimpose the plotted symbols on the standard outline of the county which is overprinted on every map (Fig. 6).

## Publication of the Maps

In this paper we cannot go into the plans for printing the Flora as a whole, but shall confine ourselves to the mapping process. Pergamon Press, Ltd., have agreed to publish the Flora, and have, of course, been consulted over the presentation of the maps. They plan to make a map of two printings. One of these, printed in black, will be reproduced from the plot of the symbols for each species, and will have keys and species names added at the time of photographing. The other will be printed from a standard block, in some other colour (probably blue, red or brown) and will comprise the county boundary, the grid lines, some physical features, such as conurbations, roads, railways, etc., and the border and panels. These will be combined, using the fiducial points for exact registration of the two printings. Transparent overlays will also be provided, giving geological and other information of possible interest for an interpretation of distribution patterns.

## Bryophyte distribution maps

The bryophyte survey, which has been planned and largely carried out by Mr. T. Laflin, is conceived on a slightly different basis from the vascular plant survey. In the bryophyte study, species lists without habitats but with data on 'fruiting' or 'sterile' will be punched

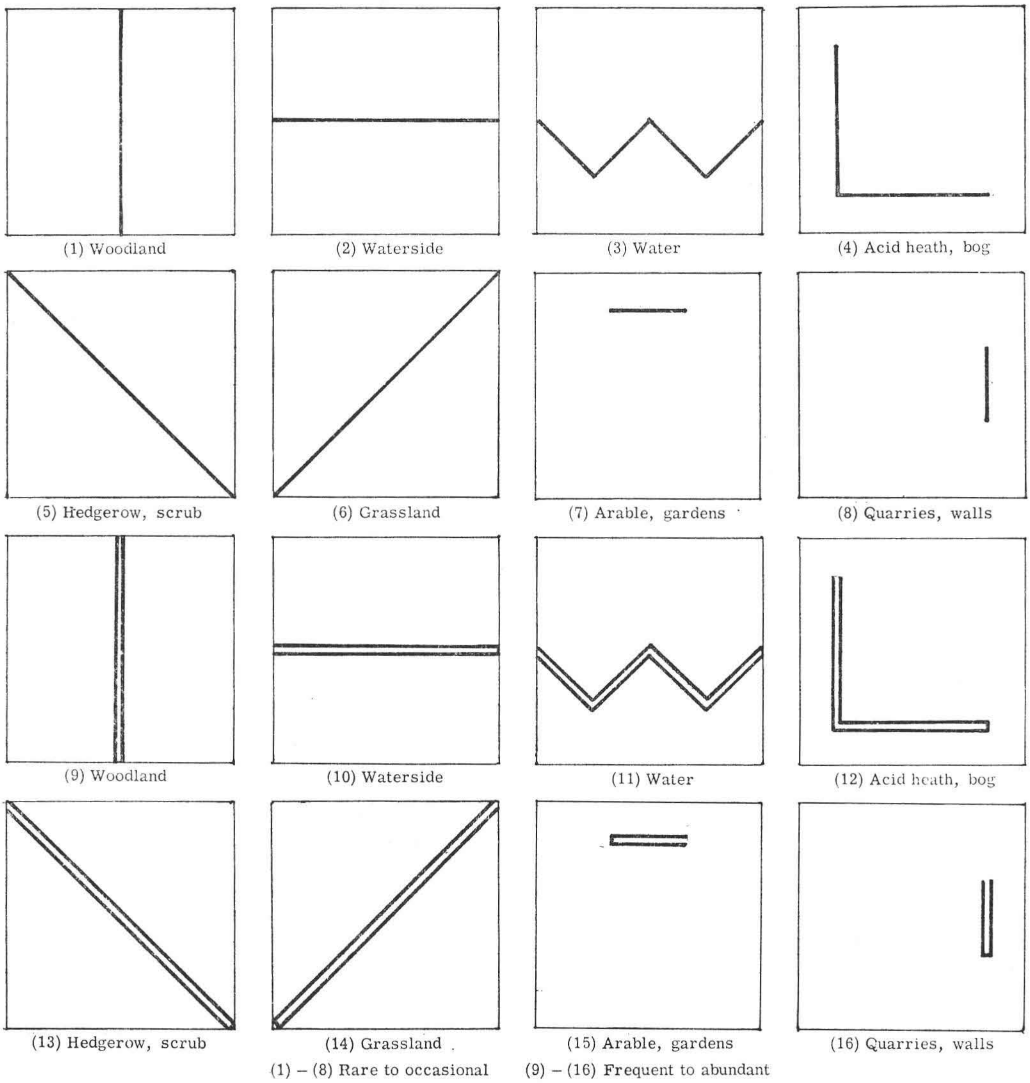



Fig. 4. Enlarged drawings of habitat symbols (other than 'ruderal') showing the track of the pen centre. At the actual size at which these are plotted (in a square of 6 mm sides) the line drawn (0·2 mm) is wide enough to fill completely the space between the lines of a thick symbol.

on the teleprinter. Thus, each record will consist of a species code reference, based on the bryophyte census catalogue number, together with a grid reference and a symbol denoting fruiting or sterile. The tapes will be passed through the computer with a different program, and an output tape produced which will give a map on the teleprinter. This device is used rather than the fast printer, since it is proposed to publish these maps direct, and the print quality of the fast printer is not really good enough for publication. Pilot maps have been produced already but at the moment of writing the primary bryophyte tapes have yet to be punched.

The teleprinter maps will bear two distinct kinds of symbol, for fruiting or sterile, and the records for these will be added up and percentaged as a by-product of the computer program (Fig. 7). These maps will also be overprinted with a standard map in some other colour, as with the flowering plant maps, though they will be reduced in size so that

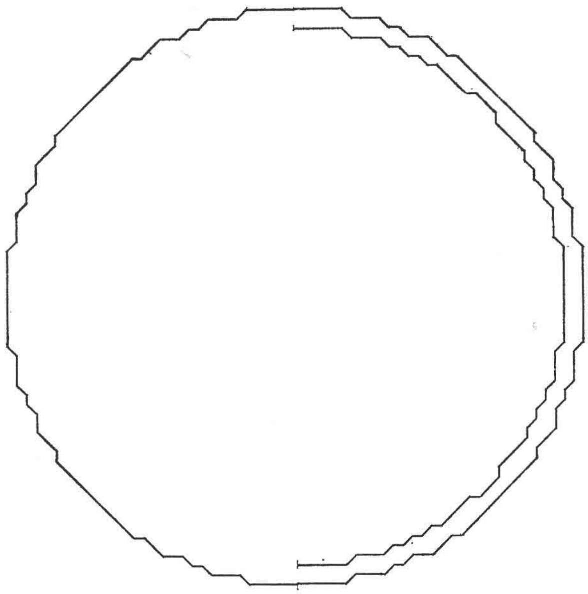


Fig. 5. The circle (Ruderal symbol), not shown in Fig. 4. On the left is the pen track to draw a thin circle. The two tracks on the right, separated by about two steps throughout, combine to form a thick circle.

four maps, instead of two, can be printed on each page. These maps, like those produced on the fast printer, suffer from inequality of vertical and horizontal scales, so the standard overprint map will have to be correspondingly distorted.

## Further applications

At the time the Warwickshire Flora survey was begun there was no prospect of a computer becoming available at Birmingham and for this reason the storage of data on punched cards or tape was not envisaged. Our data were therefore rather laboriously written on to species cards and only much later transferred to tape. It would almost certainly have been easier to have put our records on to punched cards in the first instance, as was done for the B.S.B.I. Atlas of the British Isles.

A Fungus Survey of Warwickshire has now been started and a flowering plant survey of the Wyre Forest, near Bewdley, is just beginning. With the card punching, sorting and reading facilities now available to the biology departments at Birmingham it is hoped that the data can be stored directly on punched cards or on tape. In both of these schemes mapping of some kind is envisaged, though the exact details have not yet been decided.

In County Flora work or in recording distribution of plants for even smaller areas one has limited and often specific objectives. For surveys of a wider nature on a national scale, on the other hand, the perspectives are remarkably wide. Such mapping techniques could be used to portray the distribution of any kind of organism, of fungal or other diseases, of genetic variability in cultivated plants or the results of demographic surveys, to name only a few applications. Simple and inexpensive maps of the 'sketch map' type can be produced with great speed, once the basic programs are available, and provided the data are collected in a suitable form. The more sophisticated type of map which is drawn on a graph plotter has an even wider application, limited perhaps only by the capacity of the data collectors to obtain the information desired. What seems quite clear is that computer mapping has an exciting future whose outlines at present we can only indistinctly perceive.

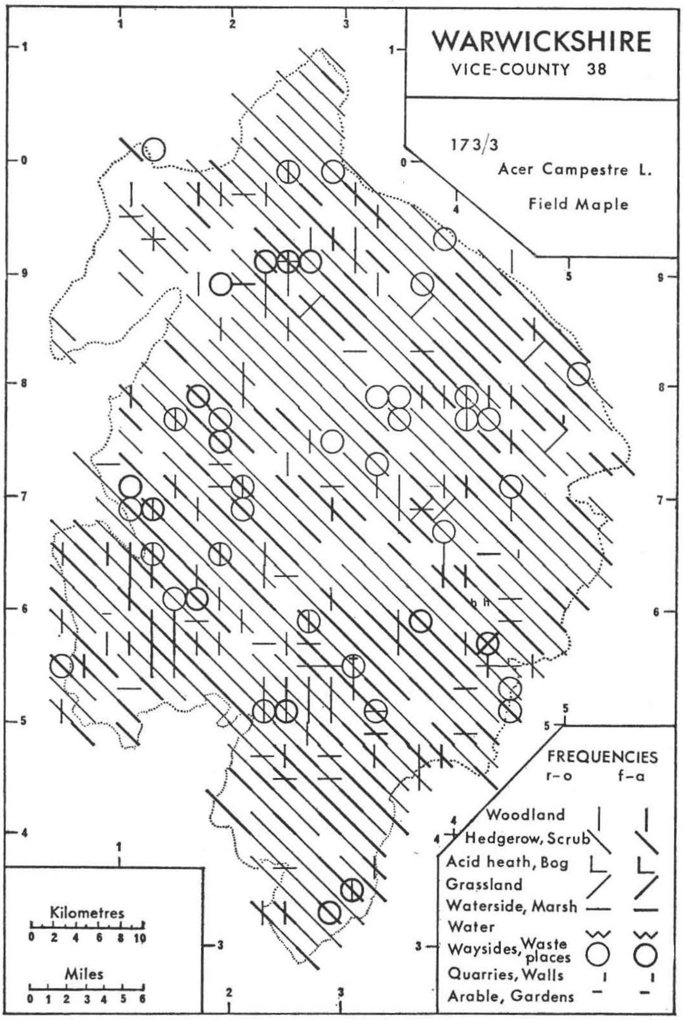



Fig. 6. Plotted map for *Acer campestre*. The county boundary, title, grid references, etc., have been added by hand. In the final publication these and other data will be overprinted in another colour. (Note: This is only a pilot map; symbol connectivity will be improved in the Flora itself.)

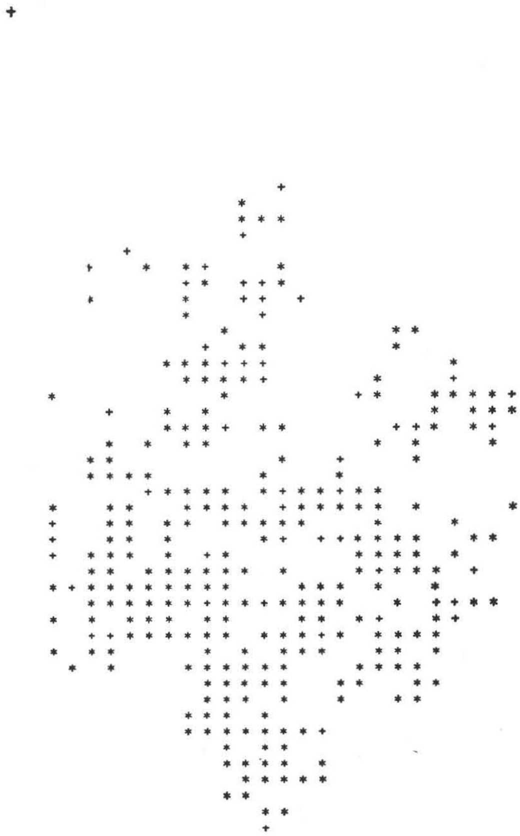
## Appendix

### Further details of the computer programs

Some readers may be interested in a rather more detailed description of the computer programs than was thought admissible in the body of the paper The language used was K-Code, the KDF9 dialect of Mercury Autocode, except for the final program which was written directly in symbolic assembler. K-Code possesses many sophisticated and useful bit handling and logical operations and allows the easy introduction of patches of assembler in areas of high-level language.

The unit of recording, the minor square, is equated with the computer word. The KDF9 word has 48 bits, of which the least significant 18 are used to represent the presence or absence of the species under consideration in the 18 possible habitats, 9 'abundant' and 9 'occasional'. One other bit indicates whether or not the square is random; this is the sign bit, giving fast discrimination of random/non-random squares by the sign detection hardware.

2129011
amblystegium serpens (hedw.) b., s. and g.



471 observations, 375 (79.62 percent) fruiting.

Fig. 7. A simple moss distribution map (*Amblystegium serpens*), produced on the teleprinter. Borders, grid references, etc., will be printed in another colour in the final publication.
+ Not fruiting, sterile.
* Fruiting.
Records of sterile and fruiting are totalled and percentaged.

If K-Code had facilities for 2-dimensional arrays, one would have been used to represent the map plotted by squares. This facility, however, is missing, so it is necessary to use a 1-dimensional array and calculate unique subscripts from the X- and Y-positions of the square under consideration, themselves calculated from the major and minor square numbers. If, when first mentioned, a square has any bit other than the sign bit non-zero, an input error, minor square mentioned twice, is signalled. Whenever an error is detected, an appropriate message is output and an error marker is set.

The input letter strings are packed into one word each. The input symbols read by K-Code are of 8 bits each, so the longest possible string, 6 letters (e.g. WOSCLA) just fits into one 48-bit word. Had there been 7-letter input strings the programming would have been considerably more difficult.

An assembled input string is compared word by word with a permanent list of 204 permissible strings. The subscript of the list word which matches the string gives a simple encoding, a number from 0 to 203 inclusive, of the input word. If no match is found the 'code unacceptable' error occurs. After a successful search of the list, the encoded form of the input string is used in three ways:

1. Placed directly onto magnetic tape, so that the input data, encoded, will be available for future work without the necessity of reading the paper tapes again.
2. As subscript to control the adding of 1 to one of 204 counters, thus accumulating the raw material for the statistics.
3. Also as subscript, to select a shift length. A 1 bit, shifted by this amount, becomes a marker for an occurrence in that major habitat/frequency and is combined with the previous contents of the correct work in the map image by the logical 'OR' operation, thus avoiding disturbing any markers already present in that word.

When the map is complete, known by the detection of the character '>' on the input (an arbitrary choice from the unused symbols on the teleprinter keyboard) the error marker is consulted. As messages are output as soon as an error is detected, if the marker is set no further action for that species is needed.

Assuming the error marker clear, however, the next action is to convert the map plotted by squares into the map plotted by tetrads. Suitable indexing selects the four squares forming one tetrad, and they are combined into one word. A slight complication arises here. A species may be both abundant and occasional in the same major habitat—e.g. abundant in oak woodland, occasional in mixed woodland. An abundant symbol blots out an occasional symbol in the plot, but to draw the thin symbol would be a waste of time and money. The routine which combines squares into tetrads therefore erases occasional observations wherever there is an abundant observation in the same tetrad and major habitat. This is achieved by a combination of shifting and logical operations. The tetrad record is placed into the appropriate location of a tetrad array, which is filled up by a suitable scanning process and becomes the final map image. This image is placed on magnetic tape ready to be handed on to program 2.

The statistics are assembled from the 204 counts, suitable totals being taken and the information re-ordered with the most favoured habitat being printed first, then the next most favoured and so on. Zero counts are suppressed.

The sketch map is formed by printing 'space' if a zero word is found during scanning of the tetrad array, 'X' if the value is greater than 511 and '+' if the value is non-zero but not greater than 511. Details of the county boundary are held in preset ('V') stores in which each bit represents a tetrad, being 1 if the boundary passes through that tetrad, zero if not. If the tetrad is empty, the appropriate bit is consulted to decide whether or not to print full stop.

The second program takes the first program's map images from magnetic tape and prepares a new magnetic tape which the final program converts into paper tape for the plotter. Having obtained an image from magnetic tape, program 2 scans it row by row. Each row is 'pre-scanned' from left to right and the positions of the left-most and right-most occupied tetrads compared with the current pen position. The latter is held in two words, one for X and the other for Y, and is assumed to start at the top left corner of every map. If the pen is nearer the left-most symbol the row is plotted with the pen moving generally from left to right, and *vice versa*. This is not a perfect way of choosing the shortest path to visit every symbol, but it is a reasonable compromise between waste pen movement and program complexity.

Having selected an occupied tetrad, the computer recalculates the current pen position in terms of an origin at the bottom left corner of that tetrad. All plotting uses this 'floating origin' technique; its advantage is that the symbol-drawing routines need not keep track of the varying positions of the symbols relative to a fixed origin. Instead they work relative to an origin which is always in the current tetrad, and a short and fairly simple routine moves the origin from tetrad to tetrad.

The computer now examines all entry points to all symbols in the selected tetrad and picks the nearest one to the current pen position. An entry point is a point at which drawing of a symbol may start; for example, the 'woodland occasional' symbol, a vertical line, may be drawn either by taking the pen to the top centre of the tetrad and moving it down, or by taking it to the bottom centre and moving it up. Top and bottom centre are the two entry points for this symbol. For the thick symbol, a narrow vertical rectangle, there are four entry points, one at each corner. Similar entries exist for all symbols, and having chosen the handiest the computer raises the pen and takes it to that point. No movement may be necessary, as with 'waterside' and 'water' symbols in the same or horizontally adjacent tetrads, in which case the pen is not raised.

Each symbol, including the circle, is built up from straight lines, and each straight line is defined by two items. It is rather easy to split KDF9 words into three 16-bit parts, known as the C, I and M parts. (Exactly why they are given these names would take us too far from our present subject.) Having selected a symbol entry point, the computer refers to a word whose C part specifies how many steps are needed to draw the first line of the symbol. The M part holds a bit pattern which can be interpreted as a step instruction by the plotter, so that between them the C and M parts completely specify the line. The I part is zero if that line completes the symbol, non-zero if at least one more line is needed to complete it. In the latter case, similar details for the next line are in the next word on the list.

This can best be explained by an example. Three lines are needed to draw the 'woodland abundant' symbol starting at its top left corner. The symbol is a rectangle 60 steps high and 2 steps wide, and the relevant control words are:

| | | |
|---|---|---|
| Word 1 C part | 60 | (Steps down) |
| Word 1 I part | Non-zero | (Another line follows) |
| Word 1 M part | 'Down' | (Expressed as a plotter control) |
| Word 2 C part | 2 | (Steps across) |
| Word 2 I part | Non-zero | (Another line follows) |
| Word 2 M part | 'Right' | (As a plotter control) |
| Word 3 C part | 60 | (Steps up) |
| Word 3 I part | Zero | (Symbol complete) |
| Word 3 M part | 'Up' | (As a plotter control) |

Compare with this the single word used to draw the 'woodland occasional' symbol (vertical line) from the top:

| | | |
|---|---|---|
| C part | 60 | (Steps down) |
| I part | Zero | (Symbol complete) |
| M part | 'Down' | (As a plotter control) |

Similar sets of words are provided for every entry point of every symbol, together with a 'symbol directory' giving details of where they begin in the control word array and of the pen positions before and after drawing them.

Having plotted one symbol the computer removes its marker from the map image and checks if the current tetrad is now empty, *i.e.* all symbols in it have been plotted and their markers removed. If not, the process of finding the nearest entry point to the (new) pen position, going to it, plotting the symbol and removing its marker is repeated until the tetrad is clear, when the next occupied tetrad is found, dealt with and cleared in the same way. After one row has been plotted the next is examined, the decision to plot from left to right or right to left made afresh and the whole process repeated until the map is complete.

The map area is extended by four rows above and below the county, and the fiducial marks are created by putting dummy 'woodland occasional' and 'waterside occasional' observations into the top left and bottom right tetrads. These plot as vertical and horizontal lines, which combine to form '+' crosses whose centres, always in fixed positions relative to the map grid, are the fiducial points.

On magnetic tape a line is represented by two or four 6-bit characters. In the two character case, the first is the actual control symbol for the plotter, which occupies 5 of the 6 bits. The next 6 bits are a binary number giving the length of the line in terms of plotter steps required to draw it. If the line is longer than 63 steps, the sixth (32) bit of the first character is made 1 and the second character is the quotient on dividing the number of steps required by 64. Any remainder is dealt with by another two characters, the first with its 32 bit zero and the second directly specifying the remainder. This system cannot draw lines longer than

$$(63 \times 64) + 63 = 4{,}095$$

steps. The longest possible line on a map, right from top to bottom, is 3,180 steps, so this limitation is not important.

The final program simply transfers the information from magnetic tape to paper tape. It takes two 6-bit characters from magnetic tape, examines the 32 bit of the first and if it is 1, clears it and multiplies the second number by 64. The first character is then output to paper tape the specified number of times. This system, which effects a considerable economy of magnetic tape, is to be adapted as part of the on-line plotting system when Birmingham University's plotter is delivered later this year.

## References

Cadbury, Miss D. A., Hawkes, J. G. & Readett, R. C. (1957). Flora of Warwickshire: Species distribution maps and habitat analyses. *Proc. Birmingham Nat. Hist. Phil. Soc.*, **18**, 135–145.

Hawkes, J. G. & Readett, R. C. (1954). The Warwickshire County Flora revision: a new method of recording. *Proc. Birmingham Nat. Hist. Phil. Soc.*, **18**, 61–74.

Hawkes, J. G. & Readett, R. C. (1963). Collecting the data. A description of the methods used in the current revision of the Flora of Warwickshire. *Local Floras* (Ed. P. J. Wanstall), B.S.B.I., London.

Hawkes, J. G., Readett, R. C. & Skelding, A. D. (1955). The Warwickshire County Flora Revision. Mapping distribution. *Proc. Birmingham Nat. Hist. Phil. Soc.*, **18**, 109–112.

Perring, F. H., & Walters, S. M. (1962). *Atlas of the British Flora*. B.S.B.I. and Nelson, London.

Soper, J. H. (1964). Mapping the distribution of plants by machine. *Canad. J. Bot.*, **42**, 1087–1100.

Soper, J. H. (1966). Machine-plotting of phytogeographical data. *Canad. Geogr.*, **10**, 1, 15–26.