METEOROLOGICAL OFFICE

FORECASTING TECHNIQUES MEMORANDUM

N° 16

A DATA SORTING PROGRAMME FOR USE IN LOCAL FORECAST STUDIES

by

D.S. GILL

Met. O.18

FEBRUARY 1968

FTM 16

METEOROLOGICAL OFFICE

FORECASTING TECHNIQUES MEMORANDUM

No. 16

A DATA SORTING PROGRAMME FOR USE IN LOCAL FORECAST STUDIES

by

D.S. GILL

FEBRUARY 1968

P.15515/JMC/2/68/350

A data sorting programme for use in local forecast studies

by

D.S. Gill

P.15515/JMC/2/68/350

A data sorting programme for use in local forecast studies

by D.S. Gill

1. Introduction

Many outstation investigations which involve the use of synoptic data require the information in a readily appreciated, tabular form instead of scattered through a number of Daily Registers. One of the simplest ways of presenting data in such a form is the frequency table - either a table of frequencies of occurrence or a table of percentage frequencies. For example, if a study of the variation of cloud base with wind direction were to be made, the obvious starting point would be a table showing, for each wind direction, the number of occasions on which the cloud base lay within various height ranges.

Before the advent of punched-card sorting methods such tables had to be compiled laboriously by hand from the Daily Registers and so a considerable effort was needed before an analysis could be completed. Since 1959, punched-card sorting methods have been used to compile such tables and to eliminate some of the hand work involved. Such data processing, carried out at Headquarters, has saved much tedious work at outstations but has not been completely satisfactory for the investigators. Results have not been available quickly because of the relative slowness of this type of data processing and because of other commitments of the Punched Card Installation; this has meant a considerable delay between the request for an analysis and its completion. The result has been that analyses by punched-card sorting have often been more detailed than was finally required because the best form of frequency tables could not be judged beforehand and it was better to have too many categories, which could be combined by hand, than too few, which would need a further machine analysis. This in turn has meant that a considerable amount of hand work has still had to be done at the outstation even after the machine analysis was completed.

The data recorded on punched cards are now being transcribed to magnetic tape - the most suitable medium for the COMET computer, which can now be used for the type of analysis previously produced by punched-card sorting. Several advantages have resulted from this. Firstly the time interval between an analysis being requested by an outstation and the completed analysis arriving at the outstation has been cut from months to days. Secondly the ranges of the parameters used in the analysis can be varied easily so that it is no longer necessary to produce an over-elaborate analysis which would need to be simplified by hand at a later stage. An assessment can be made, in consultation with Met.O. 18c if necessary, of the best form of the frequency table. If the results prove unsatisfactory, the ranges of the parameters can be changed and a second analysis carried out fairly quickly. Thirdly, more complicated sorting can be carried out and percentage frequency tables, which previously had to be worked out by hand, can now be computed and printed out as easily as ordinary frequency tables.

This Memorandum is intended as a guide to the outstation investigator of the sort of analysis which can now be obtained by computer through Met. O. 18c and the details which need to be specified when asking for an analysis. Details of the layout of the resulting analyses are given and additional features of the present programme are discussed briefly.

2. Data sorting programme

The computer programme which produces the type of analysis mentioned in the Introduction has been called the data sorting programme and will be referred to as the DSP.

Data processing by the DSP is possible only if the data are recorded on magnetic tape. Generally the surface observations so recorded are those

given in climatological returns since 1 January 1957. A detailed list of the observations which are, or will become, available is given in Annex A.

The simplest form of analysis produced by the DSP is a single frequency table for two parameters. An example of this would be a table of the form mentioned in the Introduction, namely a frequency table of wind direction against height of cloud base. Often a more complicated analysis is required. In the example quoted it might be useful to compile separate tables for various ranges of wind speed or for different seasons of the year; for either of these analyses three parameters would be needed. If separate tables were required for various ranges of wind speed and different seasons of the year then four parameters become necessary, and this is the maximum number of parameters the DSP can deal with at one time, although the month can, exceptionally, be regarded as a fifth parameter. An example of the type of table produced by a four-parameter analysis is given in Table I. In this table the basic analysis is of visibility against wind direction but separate tables have been produced for different wind speed ranges and for the winter and summer half-year. The particular table shown is for a wind speed of 7-10 kt in the winter half-year.

TABLE I - FREQUENCY TABLE* FOR VISIBILITY AND DIRECTION OF SURFACE WIND AT FINNINGLEY ON OCCASIONS DURING THE MONTHS OCTOBER TO MARCH WHEN THE WIND SPEED WAS 7-10 KT

Wind direction ranges	Visibility ranges (metres)				
	<200	200-399	400-999	1000-1999	≥2000
350°-010°	3	2	7	23	249
020°-040°	0	0	3	10	143
050°-070°	0	1	2	12	124
080°-100°	1	1	4	9	139
110°-130°	2	2	5	23	110
140°-160°	5	1	19	26	136
170°-190°	4	5	17	52	167
200°-220°	0	1	4	14	150
230°-250°	0	0	0	12	146
260°-280°	0	1	9	20	168
290°-310°	1	2	12	39	221
320°-340°	2	5	13	31	210

* Derived from three-hourly observations during the years 1962-66

So far only wind speed, direction, visibility, cloud base and month or season have been mentioned as parameters. A list of all the parameters which have been incorporated in the programme is given in Table II, together with the units in which the parameters are measured. The present-weather code figure is used as a parameter when separate analyses are required for different weather states, e.g. for precipitation and non-precipitation cases. The code figures are split into two groups, one labelled 0 (non-occurrences) and the other labelled 1 (occurrences).

TABLE II

PARAMETERS USED IN THE DATA SORTING PROGRAMME

Parameter	Units	Notes
Hour	1 hour	midnight = 0, 0100 = 1, etc.
Month	1 month	January = 1, February = 2, etc.
Visibility	hundreds of metres	Not rounded up, e.g. 199 metres counts as 1
Wind direction	tens of degrees	As in original observation
Wind speed	knots	As in original observation
Relative humidity	per cent	As in original observation
Total amount of cloud	oktas	As in original observation
Total amount of low cloud	oktas	For dates after 31.12.58.
Base of lowest cloud	hundreds of feet	For dates after 31.12.58.
Base of lowest layer of cloud 5 oktas or more	hundreds of feet	Cases where no such layer exists are allocated a base of 50,000 feet
Combined visibility and cloud base		Allocates observations to 4 ranges 0 = cloud 5 oktas or more base below 300 ft and visibility <1000m 1 = cloud 5 oktas or more base below 300 ft, visibility >1000m 2 = visibility <1000m, no cloud 5 oktas or more base below 300 ft 3 = visibility >1000m, no cloud 5 oktas or more base below 300 ft
Temperature	tenths of degrees C	
Present weather		Present weather code figures are allocated to 2 ranges 0 = Code figures not specified for the analysis 1 = Code figures specified for the analysis

There are two ways in which the ranges of a parameter can be specified - non-cumulative and cumulative. Table I is an example of the non-cumulative type, in which the lower limit of visibility in a given range follows on from the upper limit of the preceding range. It is sometimes useful, however, to study the total number of occurrences below various limits, and the ranges are then cumulative. In this way the ranges of visibility in Table I could be specified as <200 m, <400 m, <1000 m, etc., so that, for example, the column

headed <1000 m would include all occasions of fog. The DSP allows the ranges of any one parameter to be cumulative; details of which parameters can be used for this purpose are given in the section dealing with the procedure for requesting analyses.

The DSP also produces percentage frequency tables from the original tables and the user has a choice of three different types. The occurrences in the original table can be expressed as:

(i) percentages of the total for the whole table (overall percentage),

(ii) percentages of the totals for the individual rows of the table (percentage by rows),

or (iii) percentage of the totals for the individual columns of the table (percentage by columns).

Following the summary of the DSP, a more detailed description will be given of the layout of results and the procedure for requesting an analysis from Met. O. 18c.

3. Layout of results

The first page of each printout starts with some information for the computer user which can be ignored. The first line of interest to the person using the analysis is the title 'statistics compiled for (station name)'. This is followed by lines containing information about the period used for the analysis, and the months of the year and hours of the day used within this period. If present-weather code figures have been used as a parameter a list of the code figures labelled 1 will be printed out.

The frequency tables are printed out on the second and subsequent pages in sets of two, the two tables being printed out on one page where possible. The first table is an ordinary frequency table for the ranges specified with the addition of row totals and column totals and is labelled 'Frequency Table'. If more than two parameters have been used the ranges of the extra parameters to which the table refers are printed out above the heading; in the example of Table I the wind speed range and the range of months used would be printed out in this position. The ranges of the parameter used in compiling the columns of the table are printed across the paper and the ranges of the parameter used in compiling the rows of the table are printed down the paper.

The second table on each page is a percentage frequency table compiled from the first table. The type of percentage frequency is shown by the labelling of the table, the meanings of the various labels being:

(i) 'Overall percent'. The frequencies in the first table are expressed as percentages of the total number of cases in the first table.

(ii) 'Percent by rows'. The frequencies in each row of the first table are expressed as percentages of the total number of cases in that row.

(iii) 'Percent by cols'. The frequencies in each column of the first table are expressed as percentages of the total number of cases in that column.

The units of the various parameters are as shown in Table II.

Table III is an example of part of an analysis using the DSP. The analysis was for Finningley, using data from 1 January 1962 to 31 December 1966. The months used were January, February, March, October, November and December. Observations at 0000, 0300, 0600, 0900, 1200, 1500, 1800 and 2100 GMT were used. Three parameters were chosen: visibility for the

columns of the table, wind direction for the rows and wind speed as an extra parameter. The tables shown are for the wind speed range of 7-10 kt, the first table being the frequency table and the second a percentage frequency table in which the individual frequencies in a row are expressed as percentages of the total for the row.

TABLE III - AN EXAMPLE OF PART OF AN ANALYSIS PRODUCED BY THE DSP

Wind Speed 007-010

Frequency Table

	Visibility					
	LT002	LT004	LT010	LT020	LT999	
Wind Direction						
LT002	3	5	12	35	284	284
002-004	0	0	3	13	156	156
005-007	0	1	3	15	139	139
008-010	1	2	6	15	154	154
011-013	2	4	9	32	142	142
014-016	5	6	25	51	187	187
017-019	4	9	26	78	245	245
020-022	0	1	5	19	169	169
023-025	0	0	0	12	158	158
026-028	0	1	10	30	198	198
029-031	1	3	15	54	275	275
032-034	2	7	20	51	261	261
	18	39	134	405	2368	2368

Percent by Rows

	Visibility					
	LT002	LT004	LT010	LT020	LT999	
Wind Direction						
LT002	11	18	42	123	1000	1000
002-004	0	0	19	83	1000	1000
005-007	0	7	22	108	1000	1000
008-010	6	13	39	97	1000	1000
011-013	14	28	63	225	1000	1000
014-016	27	32	134	273	1000	1000
017-019	16	37	106	318	1000	1000
020-022	0	6	30	112	1000	1000
023-025	0	0	0	76	1000	1000
026-028	0	5	51	152	1000	1000
029-031	4	11	55	196	1000	1000
032-034	8	27	77	195	1000	1000
	8	16	57	171	1000	2368

The labelling of the ranges is always in a standard form regardless of the parameter being used. The first range is labelled 'LT (less than) upper limit of first range +1'. Thus in the example given in Table III the upper limit of the first range of visibility is 1(00) m. Succeeding ranges are labelled 'LT upper limit of range +1', if cumulative, as the ranges of visibility are in the example; or 'upper limit of previous range +1 (dash) upper limit of range', if non-cumulative, as are the ranges of wind direction in the example.

The last range of visibility in the example is <999(00) m and therefore includes all visibilities as its upper limit is higher than visibilities are normally reported. Any range where the upper limit is 999 is normally of this type, but in the case of non-cumulative ranges only those cases above the upper limit of the penultimate range would be included.

The second parameter's ranges, shown in the first column, are non-cumulative 30-degree sectors of wind direction. In the cases of wind direction, hour and month where the values of the parameter are discontinuous the first range of the parameter may include values above the limit of the last range. Thus, in this example, although the first range of wind direction is labelled 'LT002' it includes winds with reported directions of 35(0) and 36(0) degrees as well as 01(0) degrees because the upper limit of the last direction range is 34(0) degrees.

The percentages in the second table are row percentages, i.e. the first row gives the percentage probability of the visibility being in the specified ranges when the surface wind is 35(0)-01(0) degrees, 7-10 kt. Percentages are always printed out to the nearest tenth percent with the decimal point omitted, a fraction of more than 0.05 being rounded up. If a percentage is less than 0.05 but greater than zero 11111 is printed. If percentages are computed out of a total of zero the result is printed as 9999.

4. Procedure for requesting analyses

In order to produce frequency tables for individual requirements the DSP has to be run with a specially prepared tape to 'guide' the programme through the routines required. This guide tape is prepared in Met. O. 18c using information supplied by the person requiring the analysis. A proforma has been devised for this purpose and a copy is included at Annex B. Brief notes of points to be noted when completing the proforma are included and a more detailed description of these and other points follows.

The period for which the analysis is required consists of the starting and finishing dates for the analysis. The fact that only some months within the period may be used should be ignored at this point. The period must start at the beginning of a month and finish at the end of a month. It is possible to carry out the same analysis for two different periods if this is required.

Any number of hours of the day can be used provided the data required are recorded on magnetic tape.

The parameters available at present and the units used are given in Table II. The ranges of such parameters should be specified in terms of the units shown in Table II. The present-weather parameter is a special case as was mentioned in the section dealing with the DSP. In this case a list of present-weather code figures to be counted as 1 should be entered in the range space.

If required the ranges of the parameter used for the columns of the frequency tables or one of the extra parameters can be made cumulative, e.g. instead of the ranges of cloud base being 0-3(00) ft and 4(00)-6(00) ft they could be specified as 0-3(00) ft and 0-6(00) ft. It should be noted that the range of cloud base 0-3(00) ft would include all cloud whose base was reported as code figure 00 to 03 in a synoptic observation, i.e. cloud bases between 0 and 399 ft. There is no need for the ranges to be of uniform size, e.g. ranges of visibility could be specified as <2(00)m, 2(00)-3(00)m, 4(00)-9(00)m and 10(00)m or more. The values of wind direction, hour and month are discontinuous but a range specification can overlap such a discontinuity, e.g. a wind direction range of 35(0)-01(0) degrees is acceptable.

The ranges of the parameter used for the columns of the frequency table are printed across the paper and the width of paper available limits the number of ranges which can be specified for this parameter to 13. The number of ranges for the parameter used for the rows of the frequency table is restricted to 99. Normally a maximum of four parameters can be used but if one of the extra parameters chosen is either hour or month the number of parameters can be increased to five. This option increases the computing time considerably and should be used as sparingly as possible.

The type of percentage frequency table required is specified as one of the three types mentioned in the section on the layout of results.

5. Additional features

A version of the programme using data not normally available from punched cards can be used but special paper tapes then have to be prepared and this takes a considerable time. At present the following data are available on paper tape in Met. O. 18c;

(i) Leuchars/Shanwell: 900-metre winds, 4 per day, 1957-66 inclusive.

(ii) London (Heathrow) Airport: computed geostrophic winds, 8 per day, for October - March inclusive for the period October 1949 - March 1962.

(iii) Watnall: measured geostrophic winds, 8 per day, for the period April 1962 - March 1967 inclusive.

Another version of the programme which will use data from upper air stations is being developed. The parameters available for use by this programme will probably be as follows:

(i) At any one pressure level. Temperature, dew-point, relative humidity, humidity mixing ratio, dew-point depression, height, wind direction, wind speed.

(ii) Between any two pressure levels. Lapse rate, thickness, wind shear.

Further additions to the parameters available for use by the DSP can be made. For example the pressure and dew-point could be included among the surface parameters. In general any data entered on Met form 3256B or 3257B could be made available for use. Eventually data entered on Met form 3259A will also be available.

Any requests for further information about any versions of the DSP mentioned in the Memorandum should be addressed to Met. O. 18c through the normal channels.

ANNEX A

STATIONS FOR WHICH DATA ARE OR WILL BE AVAILABLE
FOR PROCESSING BY THE DSP

STATION NAME	NUMBER OF OBSERVATIONS PER DAY	FIRST OBSERVATION	LAST OBSERVATION
Lerwick	24	1/1/57	
Kirkwall	8	1/1/57	
Benbecula	8	1/1/57	
Stornoway	24	1/1/57	
Kinloss	8	1/1/57	31/12/58
	24	1/1/59	
Wick	24	1/1/57	
Usan	8	1/1/65	
Aberdeen (Dyce)	8	1/1/57	
Tiree	24	1/1/57	
Machrihanish	8	1/1/65	
Mull of Galloway	8	1/11/57	
Prestwick	24	1/1/57	
Abbotsinch	24	1/5/66	
Glasgow (Renfrew)	24	1/1/49	30/4/66
Edinburgh (Turnhouse)	24	1/1/57	
Eskdalemuir	24	1/1/57	
Leuchars	8	1/1/57	
Seahouses	8	1/11/63	
Isle of Man (Ronaldsway)	8	1/1/57	
Point of Ayre	8	1/1/57	
Silloth	8	1/1/57	31/12/60
Carlisle	8	1/1/61	
Acklington	8	1/1/57	
Leeming	24	1/10/65	
Dishforth	24	1/1/57	30/9/65
Tynemouth	8	1/1/57	
Middleton St George	8	1/1/57	31/3/64
Whitby	8	1/9/61	
Valley	24	1/1/57	
Blackpool (Squires Gate)	6	1/1/57	31/8/57
	7	1/9/57	30/11/57
	8	1/12/57	
Liverpool (Speke)	8	1/1/57	
Manchester (Ringway)	24	1/1/49	

STATION NAME	NUMBER OF OBSERVATIONS PER DAY	FIRST OBSERVATION	LAST OBSERVATION
Watnall	8	1/1/57	
Finningley	8	1/1/58	
Lindholme	8	1/1/57	31/12/57
Waddington	24	1/1/57	
Driffield	8	1/1/57	30/6/59
Leconfield	8	1/7/59	
Manby	8	1/1/57	
Spurn Head	8	1/1/57	31/3/64
Filnsea	8	1/4/64	
Bardsey Island	8	1/1/62	
Shawbury	8	1/1/57	
Wittering	8	1/1/57	
Marham	8	1/1/57	
Coltishall	8	1/11/62	
Goreleston	8	1/1/57	
Aberporth	24	1/1/57	
Defford	8	1/1/57	31/10/57
Pershore	8	1/11/57	
Birmingham (Elmdon)	24	1/6/49	
Mildenhall	24	1/1/49	
Wattisham	8 8	1/4/61 1/11/61	31/7/61
Aldeburgh	8	1/9/63	
Pembroke Dock	8	1/1/57	30/9/57
Milford Haven	8	1/6/64	
Mumbles	8	1/8/58	
Bristol (Filton)	8	1/1/57	
Gloucester	8	1/12/60	
Abingdon	8	1/1/57	
Stansted	8	1/1/59	
Walton	8	1/10/62	
Felixstowe	8	1/1/57	31/10/61
Chivenor	8	1/1/57	
Cardiff (Rhoose)	24	1/1/57	
Boscombe Down	24	1/1/57	
Blackbushe	8	1/1/57	31/5/60
South Farnborough	8	1/1/57	
London (Heathrow) Airport	24	1/12/48	
Kew	24	1/1/57	
London (Gatwick) Airport	8	1/6/58	

STATION NAME	NUMBER OF OBSERVATIONS PER DAY	FIRST OBSERVATION	LAST OBSERVATION
West Malling	8 8	1/1/57 1/7/64	31/8/60
Manston	24	1/1/61	
Scilly (St Mary's)	8	1/1/58	
St. Mawgan	8	1/1/57	
Plymouth (Mount Batten)	24	1/1/49	
Exeter	8	1/1/57	
Portland Bill (Coastguard)	8	1/1/67	
Portland Bill (Lighthouse)	8	1/1/57	31/12/66
Bournemouth (Hurn)	8	1/1/57	
Thorney Island	24	1/8/58	
Dungeness	8	1/1/57	
Ballykelly	8	1/1/57	
Aldergrove	24	1/1/49	
Belfast (Nutts Corner)	8	1/1/57	

ANNEX B

REQUEST FOR STATISTICAL ANALYSIS BY DSP

1. Statistics required using observations for (station name) .....
.....

2. Period required[1] from .....
to .....

3. Months of the year to be used .....
.....

4. Hours of the day to be used[2] .....
.....

5. Parameter to be used for columns of frequency table/s[3] .....
.....
Ranges of this parameter[5] .....

6. Parameter to be used for rows of frequency table/s[4] .....
.....
Ranges of this parameter[5] .....

7. Extra parameters required[6]

a. Type .....
Ranges[5] .....

b. Type .....
Ranges[5] .....

c. Type[7] .....
Ranges[5] .....

8. Type of percentage frequency required[8] .....

9. Station requiring statistics (if different from 1).....
.....

NOTES

1. The period must start at the beginning of a month and finish at the end of a month. More than one period can be specified if required.

2. Any number of hours can be used provided the data are on magnetic tape.

3. The ranges of this parameter must not exceed 13 in number.

4. The ranges of this parameter must not exceed 99 in number.

5. Ranges need not all be of the same size. If required, the ranges of the parameter in section 5 or of an extra parameter can be made cumulative. Ranges can overlap a discontinuity in values of the parameter.

6. These extra parameters are optional.

7. A third extra parameter can be specified only if the first or second extra parameter is either hour or month.

8. Specify 'overall', by 'by rows' or 'by columns'.

Meteorological Office

Forecasting Techniques Memoranda

Amendment No 1 to FTM 16

p 7. Add after "(iii) Watnall ... inclusive":

(iv) Topcliffe: measured geostrophic wind, surface wind, surface and estimated 900 mb temperatures in whole degrees, sea level pressure in whole millibars, all 8 per day from March 1968 to February 1970 inclusive, with some hours missing.

pp 8–10, Annex A. Delete the following stations, whose observations are not on magnetic tape:

Usan	Lindholme	Aldeburgh	Blackbushe
Mull of Galloway	Bardsey I	Pembroke Dock	West Malling
Seahouses	Gorleston	Mumbles	Portland Bill (Coastguard)
Point of Ayre	Defford	Walton	Belfast (Nutts Corner)

Amend the entries of the following stations to read as shown:

Blackpool (Squires Gate)	8	1.12.57	
*Waddington	24	1. 8.50	

*Scampton observations used Aug 1953–May 1955 and Jul–Aug 1968

Leconfield	8	1. 7.59	31.12.67
Pershore	8	1.11.57	30. 9.67
Wattisham	8	1.11.61	
South Farnborough	8	1. 1.57	30. 6.66
	7	1. 7.66	
London (Heathrow) Airport	24	1. 1.49	

Met O 18c

March 1971

G. 13083/MD/3/71/180

FORECASTING TECHNIQUES MEMORANDA

1. **Forecasting precipitation - methods and techniques in use in the Meteorological Office.** By W.D.S. McCaffery and D.S. Gill. 1964.
2. **Surface and 900 mb wind relationships.** By G.A. Howkins, T.N.S. Harrower and D.S. Gill, 1965.
3. **Wet bulb temperatures - a comparison between readings from clean and dirty instruments.** By W.D.S. McCaffery. 1965.
4. **Operational numerical forecasts.** By P. Graystone. 1965.
5. **Temperature and humidity in the lowest 3000 ft - effectiveness of current radio-sonde reporting procedures.** By W.D.S. McCaffery and D.S. Gill. 1965.
6. **Comparison of equivalent headwinds from 300 mb objective numerical forecasts and subjective forecasts.** By G.A. Howkins and I.H. Chuter. 1965.
7. **Forecasting methods and techniques in use at Meteorological Office outstations.** By W.D.S. McCaffery and T.N.S. Harrower. 1965.
8. **Tests of thunderstorm forecasting methods.** By W.E. Saunders. 1965.
9. **Three-parameter atmospheric model used for numerical weather prediction.** By G.A. Bull. 1965.
10. **Objective analysis in the numerical weather forecasting system.** By G.A. Bull. 1966.
11. **Some results from standard-programme tabulations of visibility and height of cloud base.** By J.E. Atkins. 1966.
12. **On sea-breeze forecasting techniques.** Edited by W.D.S. McCaffery. 1966.
13. **Techniques in use at Meteorological Office outstations for forecasting local cooling at night.** By J.E. Atkins. 1966.
14. **Further tests of thunderstorm forecasting methods.** By W.E. Saunders. 1967.
15. **Work on problems in local forecasting.** By W.D.S. McCaffery. 1967.
16. **A data sorting programme for use in local forecast studies.** By D.S. Gill. 1968.