Objective Analysis of Climatological Fields of Temperature, Sunshine and Rainfall

>2000 mm
1600 - 2000 mm
1200 - 1600 mm
1000 - 1200 mm
800 - 1000 mm
<800 mm

0⊏▅▅▅▅⊐350 Km

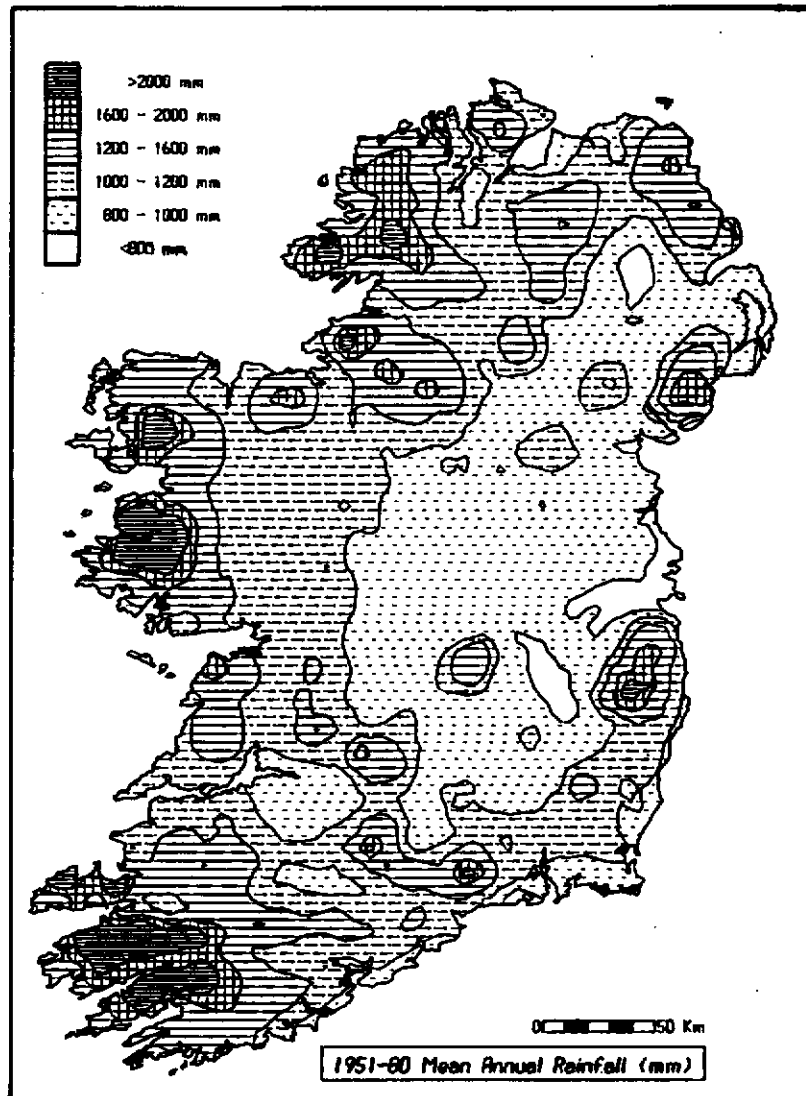*1951-60 Mean Annual Rainfall (mm)*

by

J.Hamilton, P.Lennon, B.O'Donnell

Objective Analysis of Climatological Fields of Temperature, Sunshine and Rainfall

by

J. Hamilton, P. Lennon, B. O'Donnell

May 1987

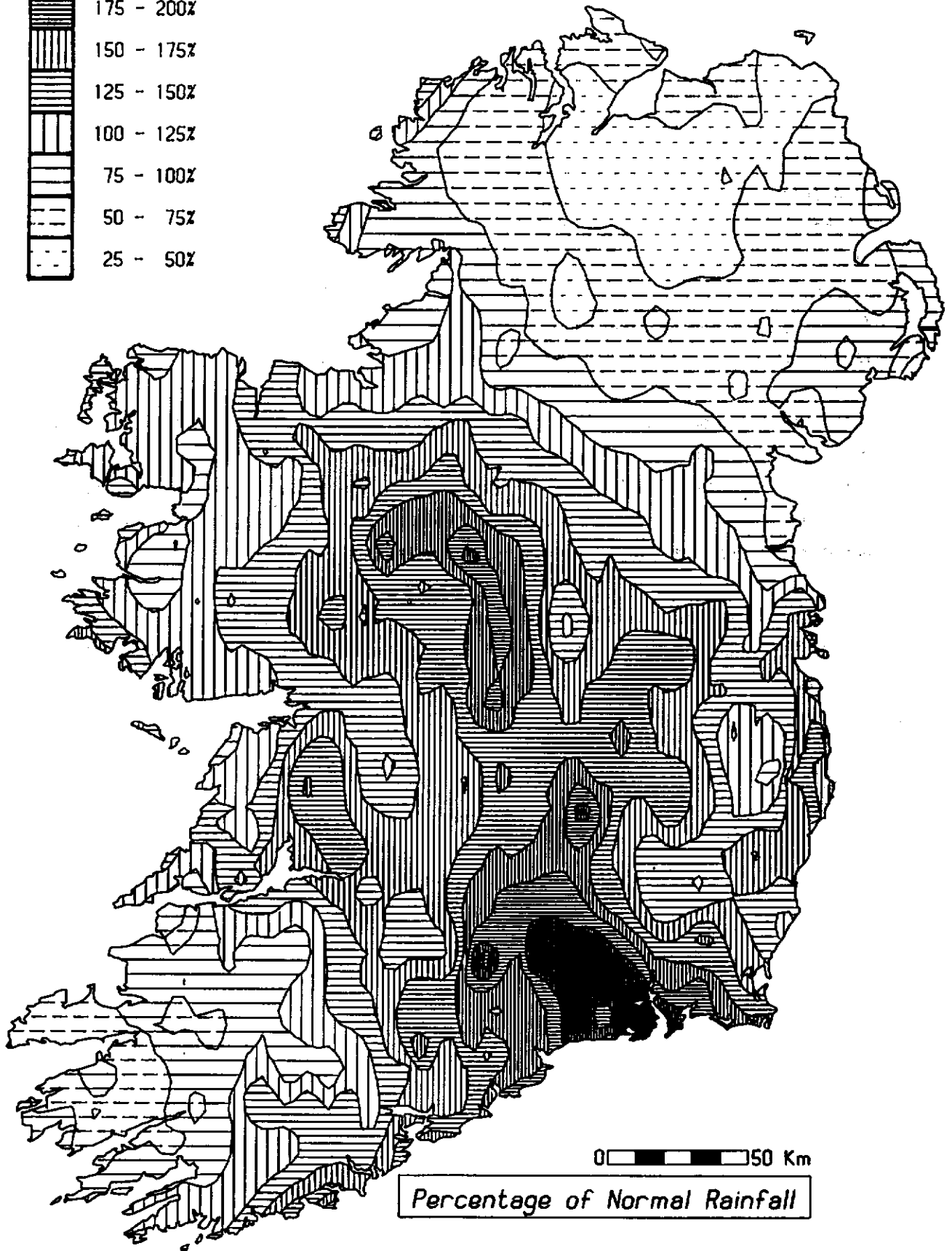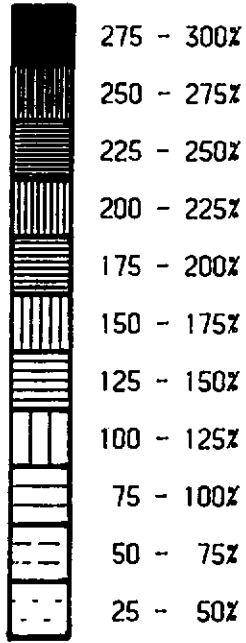## Meteorological Service - Glasnevin - Dublin - Ireland

Abstract

An objective method of drawing climatological maps of sunshine, temperature, rainfall percentage and rainfall amount, which is used operationally at the Irish Meteorological Service in Dublin, is described. The interpolation is based on McLain's method of distance weighted least squares quadratic approximation (McLain 1974). The observations are interpolated directly to a regular (rectangular) mesh, without the use of a first-guess, and the resultant field is contoured and shaded using a standard contouring package.

This basic interpolation gave some spurious results near the coast and in data void areas when it was applied in its original form. These problems were overcome by the introduction of dummy values in such regions and a method was developed for calculating such values automatically based on nearby observations and various climatological normals.

The method has been in operational use since February 1985 and the results are of quality comparable to the manual method which this computer method has replaced.

June 1986

Percentage of Normal Rainfall

275 – 300%	
250 – 275%	
225 – 250%	
200 – 225%	
175 – 200%	
150 – 175%	
125 – 150%	
100 – 125%	
75 – 100%	
50 – 75%	
25 – 50%	

0 ⌷⌷⌷⌷⌷ 50 Km

1. Introduction

This paper describes the technique of objectively analysing monthly climatological maps by computer which is used operationally at the Irish Meteorological Service. The method is used to draw maps of mean air temperature, mean daily sunshine, rainfall as a percentage of normal and rainfall amount. Data from scattered observation stations is interpolated to a regular rectangular grid, using McLain's technique of distance weighted least squares quadratic approximation (McLain 1974). The grid is then contoured and shaded using a standard package (Hamilton 1985). In implementing McLain's algorithm, we have introduced a number of improvements — in particular that of sorting the observations by their grid number — to make the method computationally more efficient.

This objective (computer) method replaces a manual method which was used from the birth of the Meteorological Service (in 1936) until February 1985. In computerising the process we wanted a technique that would produce charts comparable in quality to those produced manually, that would be reasonably efficient, and that would need no manual intervention. Another requirement was for a method that would not require a preliminary field or 'first guess' since we had no digitising facilities for entering such a field.

Three basic approaches to the problem of objective analysis are (a) surface fitting, (b) successive corrections (Cressman 1959) and (c) optimal interpolation (Gandin 1965). The last two methods are better than surface fitting especially where there are large data sparse areas. However, with good data coverage simple surface fitting methods can be quite satisfactory (Gustafsson 1981) and this paper discusses such a method (McLain 1974).

The Cressman technique is efficient and easy to implement. However, it has the disadvantages that it needs a first guess and a lot of fine tuning to give good results.

The optimum interpolation method uses weighting parameters based on various spatial correlations. These parameters can be calculated objectively provided a large statistical data base is available. However, in the absence of such a data base, we would have to choose these parameters by experiment and so optimum interpolation would be no better than the Cressman method in this regard. Also, the requirement for a realistic first guess is a disadvantage.

The surface fitting methods are based on fitting some type of function — generally a polynomial — to the data. McLain's technique is of this type and we decided to start by investigating this method. It is easy to implement, since,

In its original form, it has just a single arbitrary parameter. Also, it can produce good results without the need to use a first-guess.

We tried out McLain's method for analysing sunshine and temperature and found that it gave such good results that we did not investigate more complicated approaches such as optimal interpolation. The same method was successfully applied in analysing percentage of normal rainfall. In the case of rainfall amounts, however, we found it necessary to do some simple modelling of orographic effects to produce acceptable analyses. When such modelling was introduced the analyses were of a high quality. This method is now used operationally in the Irish Meteorological Service.

2. Mathematical Background

The basic problem of interpolation is as follows. We have n data values (or observations) $[z_1, z_2, z_3 \ldots z_n]$ specified at the positions $[(x_1, y_1), (x_2, y_2), (x_3, y_3) \ldots (x_n, y_n)]$ respectively. We wish to estimate the value of z at the position (a,b) using the information contained in these n data values.

McLain's method consists of fitting a second degree polynomial to the observations and then evaluating it at the point (a,b). Thus, consider the general polynomial of second degree :

$$P(x,y) = c_{00} + c_{10}x + c_{01}y + c_{20}x^2 + c_{11}xy + c_{02}y^2 \tag{1}$$

Our aim is to choose the coefficients c_{rs} so that the polynomial will be as accurate a fit as possible, in the usual least squares sense, to the data. The difference between McLain's method and the usual least-squares approximation, is that in his method data points (x_i, y_i) close to (a,b) carry more weight than distant points. More precisely we choose the coefficients c_{rs} to minimise the quadratic form

$$Q = \sum (P(x_i, y_i) - z_i)^2 \cdot w((x_i - a)^2 + (y_i - b)^2) \tag{2}$$

where w is a weight function, such as $w(d^2) = 1/d^2$, which is large when (a,b) is close to (x_i, y_i) and small when it is remote. The minimisation is achieved in the usual way by solving the (in our example, six) linear equations

$$\frac{\partial Q}{\partial c_{rs}} = 0 \tag{3}$$

In the (six) unknown coefficients c_{rs}. Having found these coefficients we can evaluate the height of the surface at (a,b) as $P(a,b)$.

Combining equations (2) and (3) and performing the differentiations we obtain the following set of equations for the coefficients :

$$
\sum_i
\begin{bmatrix}
w_i & X_i w_i & Y_i w_i & X_i Y_i w_i & X_i^2 w_i & Y_i^2 w_i \\
X_i w_i & X_i^2 w_i & X_i Y_i w_i & X_i^2 Y_i w_i & X_i^3 w_i & X_i Y_i^2 w_i \\
Y_i w_i & X_i Y_i w_i & Y_i^2 w_i & X_i Y_i^2 w_i & X_i^2 Y_i w_i & Y_i^3 w_i \\
X_i Y_i w_i & X_i^2 Y_i w_i & X_i Y_i^2 w_i & X_i^2 Y_i^2 w_i & X_i^3 Y_i w_i & X_i Y_i^3 w_i \\
X_i^2 w_i & X_i^3 w_i & X_i^2 Y_i w_i & X_i^3 Y_i w_i & X_i^4 w_i & X_i^2 Y_i^2 w_i \\
Y_i^2 w_i & X_i Y_i^2 w_i & Y_i^3 w_i & X_i Y_i^3 w_i & X_i^2 Y_i^2 w_i & Y_i^4 w_i
\end{bmatrix}
\begin{bmatrix}
c_{00} \\ c_{10} \\ c_{01} \\ c_{11} \\ c_{20} \\ c_{02}
\end{bmatrix}
=
\sum_i
\begin{bmatrix}
z_i w_i \\
X_i z_i w_i \\
Y_i z_i w_i \\
X_i Y_i z_i w_i \\
X_i^2 z_i w_i \\
Y_i^2 z_i w_i
\end{bmatrix}
$$

$$....(4)$$

where we have made the substitutions $X_i=(x_i-a)$ and $Y_i=(y_i-b)$ and where the summations (over the observation points) should be performed before solving the system of equations. Note that, when interpolating to a grid, this set of equations must be solved for each individual grid point.

McLain, in his paper, discusses a number of possible weighting functions. The use of a simple $1/d^2$ (or to avoid arithmetic overflow $1/(d^2+\epsilon)$ for small ϵ), does not prove particularly satisfactory. He believes this is probably because the data points remote from (a,b) have too much weight. The more rapidly decreasing function $\exp(-\alpha d^2)$, for some suitable constant α, gives much more accurate results. If the data points are subject to experimental error the use of such a function with perhaps a smaller value of α, leads to a smoothing of the surface. However, if the observations are exact, then when (a,b) is very close to some (x_i,y_i) we would expect that to dominate absolutely. Hence McLain recommends the function :

$$
W(d^2) = \frac{e^{-\alpha d^2}}{\epsilon + \alpha d^2} \tag{5}
$$

where α is a constant (of the order of the inverse of the square of the average distances between neighbouring data points) and where ϵ is a small

constant to prevent arithmetic overflow at d=0. The significance of α can be seen by introducing the scale length d_{scal} defined by :

$$d_{scal} = \frac{1}{\sqrt{\alpha}} \qquad (6)$$

so that the weighting function can be written in the alternative form :

$$W(d^2) = \frac{e^{-(d/d_{scal})^2}}{\epsilon + (d/d_{scal})^2} \qquad (7)$$

This function behaves like $1/d^2$ near an observation point (and grows rapidly with decreasing distance from the observation) but behaves like $\exp(-\alpha d^2)$ further away from the point (where it falls off rapidly). Figure 1 shows a graph of the function and compares its behaviour with $1/d^2$ and $\exp(-\alpha d^2)$. Note, that for $\alpha d > 2$ the function is negligible.

The distance, d, between the grid point (a,b) and the observation point (x_i, y_i) can be calculated from the latitude and longitude coordinates of the points (Maling 1973). Thus (in units where the Earth's radius is unity),

$$\cos d^2 = \sin \phi_i \sin \phi_g + \cos \phi_i \cos \phi_g \cos \Delta\lambda \qquad (8)$$

where λ is longitude, ϕ latitude, the subscript g refers to the grid point (a,b) and $\Delta\lambda = (\lambda_i - \lambda_g)$. For small $\Delta\lambda$ and small $\Delta\phi$ (where $\Delta\phi = \phi_i - \phi_g$) the distance can be written :

$$d^2 = (\Delta\lambda \cos \phi_0)^2 + (\Delta\phi)^2 \qquad (9)$$

where $\phi_0 = \frac{1}{2}(\phi_i + \phi_g)$. In practice, for a country the size of Ireland, $\cos \phi_0$ does not vary much from North to South, and so we can treat $\cos \phi_0$ as a constant.

3. Programming Considerations

The analysis of (say) a chart of monthly accumulated rainfall is carried out in the following two stages :

(a):The observations are interpolated to a regular latitude/longitude grid covering the whole of Ireland

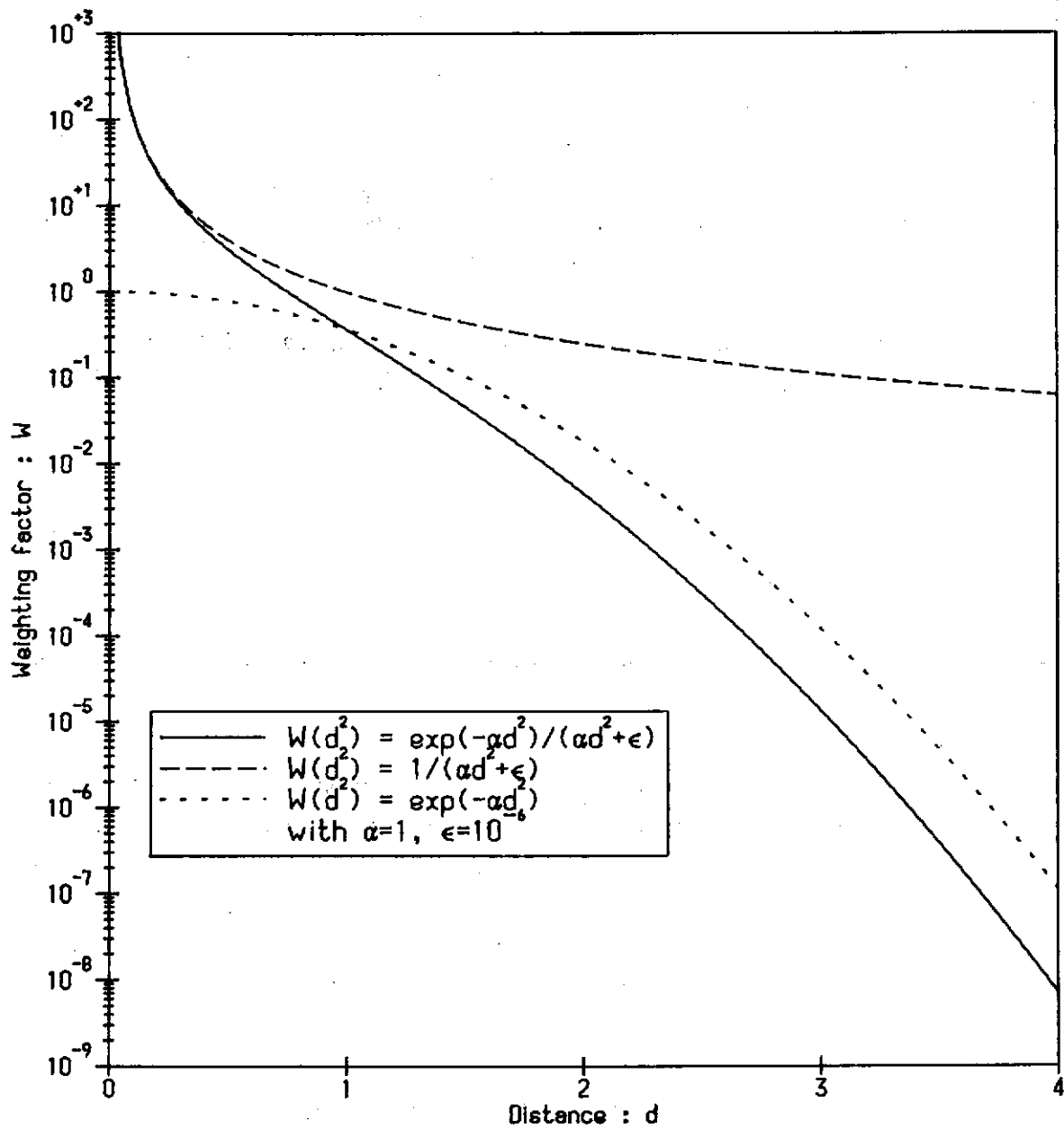(b):The grid data is contoured and the contour lines clipped against the Irish coastline.

Figure 1: Comparison of weighting functions. Note how the weighting function recommended by McLain (solid line) combines the desirable features of the $1/d^2$ function at small distances and the $\exp(-\alpha d^2)$ function at large distances.

The interpolation is the most time consuming part of the process. We have made a number of modifications to McLain's basic algorithm to increase the efficiency of the method. McLain suggests dividing the region into sub-regions and ignoring data points in remote sub-regions. We have taken this a step further by sorting the data points according to the grid-cell-number containing the observation. We define a cutoff radius, d_{cutoff}, and, with the aid of the list of sorted observations, ignore all observations outside the cutoff radius. In a further refinement of these ideas, we have the option of searching for observations in a spiral motion moving outwards from the grid point - when a certain critical number of observations, N_{enough}, have been accumulated the search for further observations is halted and the interpolation is performed. Both these techniques have a significant effect on the efficiency of the interpolation. In addition the introduction of the N_{enough} option has a small effect on the interpolated values since it causes more weight to be given to nearby points as points further away are ignored.

The interpolation is performed in latitude/longitude space using the cosine factor discussed in connection with equation (9). Thus we define a coordinate system :

$$x = \lambda \cos \phi_0 \quad \text{(longitude)}$$

$$y = \phi \qquad \text{(latitude)}$$

(10)

where ϕ_0 = $53 \cdot 35^O$ (since Ireland extends from approximately $51 \cdot 3^O N$ to $55 \cdot 4^O N$).

Much of the cost of the interpolation is the solution of the systems of equations which arise for each grid point. We have optimized their solution by writing a special direct solver with the DO-loops un-rolled for the special case of a system of six equations. Also, to improve the numerical stability of the solution, we shift the origin of the coordinate system to the grid point in question, before calculating the matrix coefficients in equation (4).

4. Operational Production of Sunshine and Temperature Maps

Before the introduction of the interpolation routines, sunshine and temperature maps were drawn up entirely by hand. The process involved the manual plotting of observations and the manual insertion of isolines. The final maps for publication were traced and shaded by a draughtsman. These charts were the first to be generated automatically using the interpolation routines.

At the outset we encountered various difficulties with the interpolation. These

	d_{cutoff}	d_{scal}	N_{enough}	Grid	No. Obs
Sunshine	2.5	1.2	--	75x100	80
Temperature	2.25	1.2	--	75x100	140
Rainfall Percentages	0.7	1.2	18	90x120	470
Rainfall Amounts	0.7	1.2	18	90x120	1100

Table 1: Parameters used operationally in the interpolation of sunshine, temperature, rainfall percentages and rainfall amounts. These values were chosen after much experimentation. d_{scal} and d_{cutoff} are in units of degrees of latitude. The number of observations includes both real observations and 'dummy' values. Sunshine and temperature are analysed using the routine SINTRP and rainfall percentage and rainfall amount using the routine NINTRP. Both these routines are in the PCONTR contouring package (Hamilton, 1985).

were due to (a) topographical considerations for the sunshine maps, as the temperature values used are already reduced to mean sea level, (b) irregular station distribution leading to certain extreme values being 'lost' in the interpolation and (c) data void or 'blank' areas being drawn for spuriously, particularly in coastal regions. With no values beyond the coastline, if a 'blank' area occured between the coast and values, (say) fifty kilometres away, the interpolation for the area would assume the gradient up to that point and continue it to cover that area. This resulted in a completely false representation of values over the area involved particularly if the gradient was already steep.

These difficulties were overcome by the following two modifications :

(a):We adjusted the parameters d_{scal} and d_{cutoff} to put more weight on local extreme values and to exclude more distant points. This caused the isolines to become jagged and angular with the (30 x 45) mesh we were using which was not fine enough for a satisfactory result. We found eventually that a (75 x 100) grid suited the number of observations – 80 for sunshine and 120 for temperature (including some manually entered values from Northern Ireland). The (90 x 120) grid which proved successful in the interpolation of greater amounts of data (i.e. rainfall percentages and amounts) proved too tight a mesh for the number of observations involved.

(b):We inserted 'dummy' values at various points. This was as a result of comparisons between the objective method and the manual method (using many monthly data-sets), which showed constant disagreement in a small number of areas. These were invariably in 'blank' mountainous areas, were sunshine values are always considerably lower than surrounding areas, and in 'blank' coastal regions for reasons already outlined. The values where calculated using a mixture of surrounding station values and climatologically-comparable station values and then used in the interpolation. It was found that using between 2 and 8 such dummy values was sufficient to produce a resultant interpolation that compared favourably with the hand-drawn maps.

5. An Example of a Sunshine Map and of a Temperature Map

Figure 2 shows a computer analysis of sunshine data for March 1985. Data from 74 stations and 2 dummy values (off the West coast of Ireland) were interpolated to a (75 x 100) grid covering the area $10 \cdot 6^{\circ}W$ to $5 \cdot 4^{\circ}W$ and $51 \cdot 3^{\circ}N$ to $55 \cdot 4^{\circ}N$. The following parameters were used :
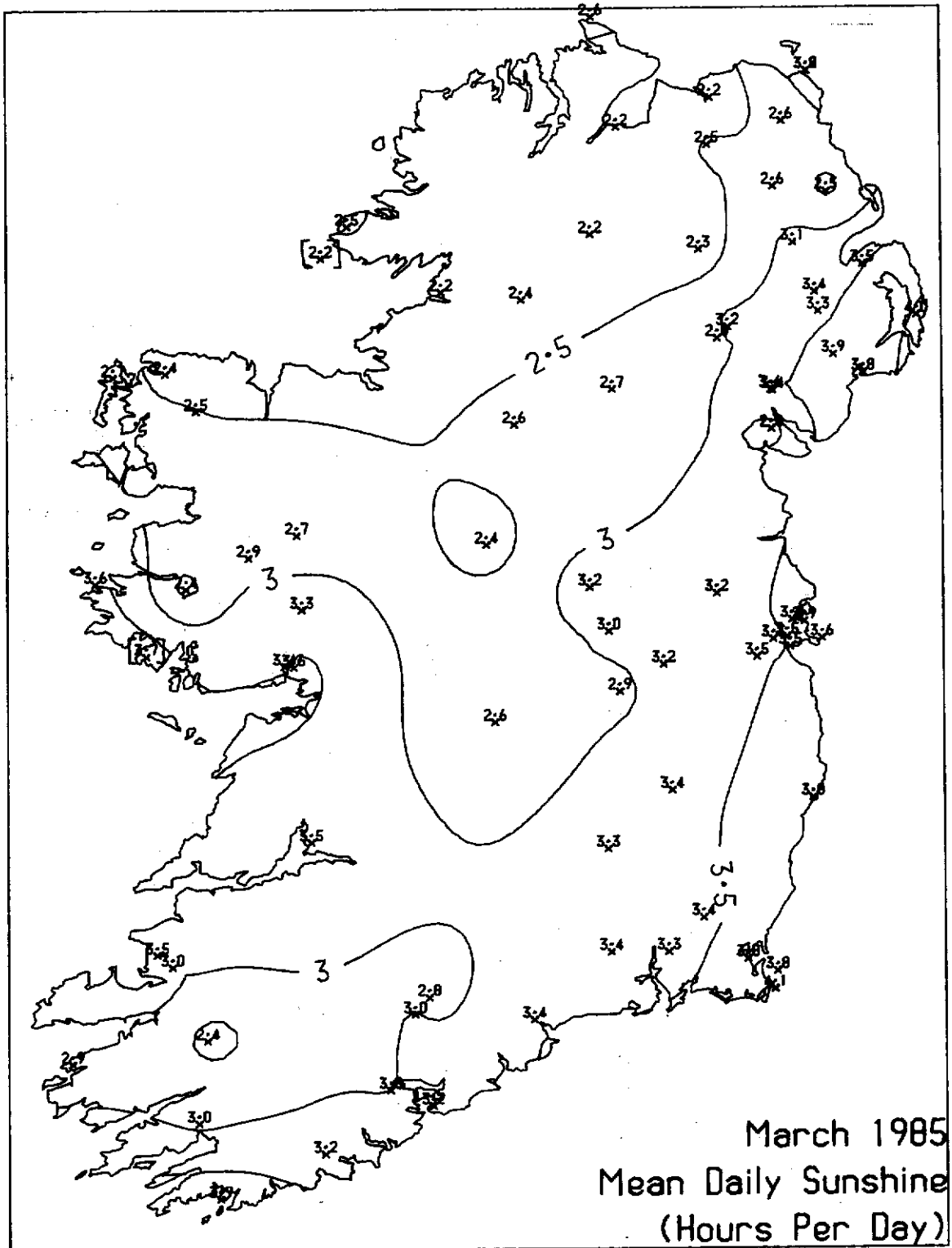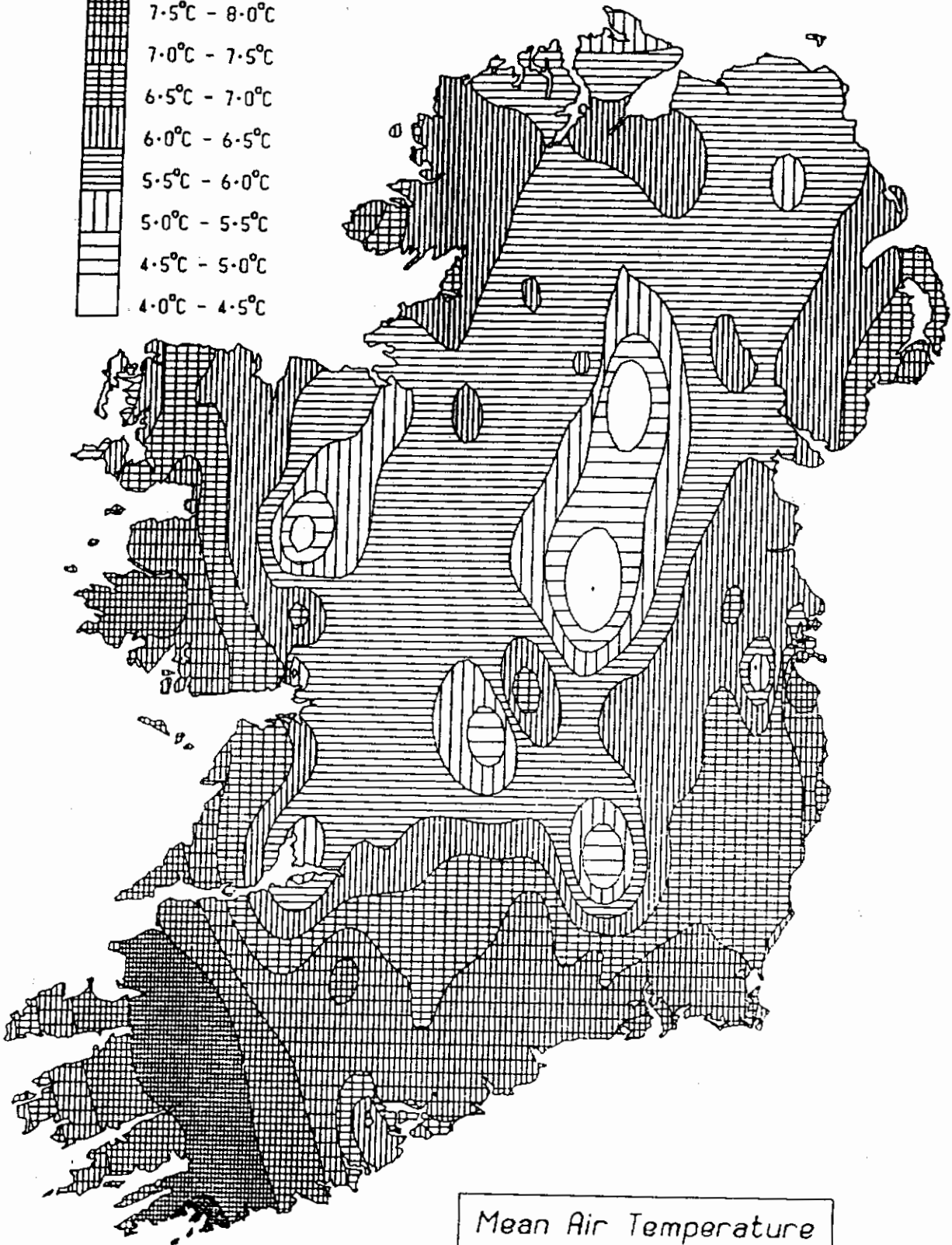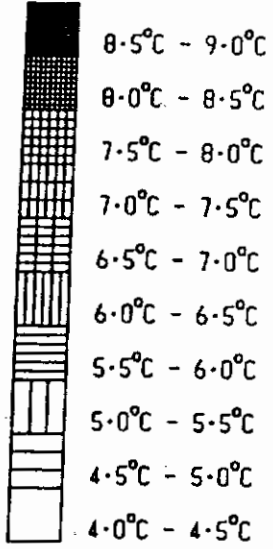
Figure 2: Example of a computer produced analysis of sunshine. Data from 74 stations and 2 dummy values are interpolated to the grid. The two dummy values are off the West coast of Ireland and are enclosed in brackets.

DECEMBER 1985

Legend:
- 8·5°C – 9·0°C
- 8·0°C – 8·5°C
- 7·5°C – 8·0°C
- 7·0°C – 7·5°C
- 6·5°C – 7·0°C
- 6·0°C – 6·5°C
- 5·5°C – 6·0°C
- 5·0°C – 5·5°C
- 4·5°C – 5·0°C
- 4·0°C – 4·5°C

Mean Air Temperature
(reduced to Mean Sea Level)

Figure 3: Example of a computer produced analysis of temperature. This map is contoured and shaded by computer and has been published in this form (Meteorological Service, 1985).

$$\phi_0 = 53 \cdot 35^{\circ} \quad (\cos \phi_0 = 0 \cdot 60)$$

$$d_{scal} = 1 \cdot 2 \qquad\qquad\qquad\qquad (11)$$

$$d_{cutoff} = 2 \cdot 5$$

The results of the interpolation are shown in figure 2 where we have plotted both the observations and a contour plot of the grid point values.

We tested the accuracy of the interpolation by taking the grid point values and interpolating them back to the observation points. We used both bi-linear and bi-cubic interpolation to perform the back interpolation and we got essentially the same results with both methods. The rms difference between the observational values and the analysed values was 3% and the maximum difference was 10%. Thus this method of analysis draws quite well for the observations.

Figure 3 shows a computer analysis of mean air temperature for December 1985. This map has been contoured and shaded and has been published in this form (Meteorological Service, 1985). The interpolation parameters used for generating this plot are summarised in Table 1.

6. Operational Production of Percentage Rainfall Maps

The interpolation procedure was next applied to producing maps of rainfall percentages, using approximately 400 points with fairly even distribution across the country. Percentage values for another 70 stations supplied by the Northern Ireland Meteorological Office were included in the interpolation. While the resulting maps were found to compare closely with those drawn up manually, localised highs and lows tended to be omitted. This problem was overcome by adjusting the interpolation parameters in such a way as to exclude more distant points from the interpolation. Firstly, the value of the cutoff radius (d_{cutoff}) was decreased and secondly, a threshold value was introduced to predetermine the number of observations to be included (N_{enough}). Given the high density of rainfall stations around the country, these changes still enabled valid interpolations to be made. As percentage values are independent of topographical influences, the maps thus produced are sufficiently accurate to obviate the necessity of including dummy values. See Table 1 for a summary of the interpolation parameters used, and the frontpiece for an example of such a map.

7. Operational Production of Rainfall Maps

The production of rainfall amounts maps using the interpolation routines presented greater difficulties owing to the highly variable pattern of rainfall over small areas. There were two problems, both related to lack of data. Firstly, in some areas near the coast an inaccurate gradient was assumed by the interpolation in the absence of data. Secondly, the analysed values were too low in mountainous areas of high rainfall not represented by an observing station. The former problem was not serious and could be solved by introducing just a few dummy values. However, the latter was much more fundamental and was basic to the modelling of orographic effects.

The first approach to the rainfall analysis was to introduce dummy values to fill in the gaps in the network outlined above. As a first step, around 40 such dummy values were found necessary to provide a reasonable picture of rainfall distribution (figure 4). This approach was unsatisfactory, however, as these values were most easily estimated from inspection of the manually produced map (figure 5) which it was hoped eventually to replace. In addition, keying in these values, together with values for Northern Ireland stations, was a laborious task on a monthly basis.

The next approach, therefore, was to estimate the dummy values automatically. An estimate of monthly amounts for each of approximately 950 stations (including dummy stations) was calculated using a combination of climatological norms and the percentage deviations of nearby stations. The procedure was as follows :

(a):The first step was to estimate n(r,t), the average monthly rainfall for each station as a function of position (r) and of the month of the year (t). Each monthly average was estimated by writing it in the form :

$$n(r,t) = Y(r) \times I(t) \tag{12}$$

where $Y(r)$ is the yearly normal for each station and $I(t)$ is a climatological index of the distribution of rainfall over each month of the year. Note that although monthly as well as annual rainfall normals were available, by using the annual figures alone only one value had then to be estimated for dummy stations representing blank areas. These annual normals were obtained from the published rainfall averages available for Ireland for the period 1951-1980 (Fitzgerald 1984), while dummy annual values were estimated from the accompanying maps.

This method of estimating monthly averages was not sufficiently accurate so we divided the country into 32 regions (the counties of Ireland) and used the

following estimate :

$$N(r,t) = Y(r) \times I_c(t) \tag{13}$$

where there are 32 values of $I_c(t)$ for each of the 12 months.

(b):The next stage was to use these averages to calculate estimated values for a particular month and compare them with the actual observations where available. We took April 1986 as an example and calculated an average percentage rainfall index for each county using observations from real stations. Thus:

$$P_c(t) = 100 \cdot average (R(actual)/N(actual)) \tag{14}$$

where R(actual) is the rainfall observation for the given station and N(actual) the corresponding normal.

(c):Finally we used the percentage index for each county to produce an estimate of each observation :

$$E(r,t) = (P_c(t)/100) \cdot N(r,t) \tag{15}$$

Around 90% of the values produced from this calculation were found to lie within ±20% of the available recorded amounts. The resulting map (figure 6), therefore, produced a reasonable general pattern of rainfall distribution but was unable to show minor deviations in the overall pattern, particularly in the bigger counties where an average percentage value represented a large number of stations.

The next stage of the analysis was to combine these estimated values with real observations. Where a report is available for a station the monthly total is used but in addition an estimated total E(r,t) is used for each of a preselected group of points. The result is to use a combination of around 650 actual amounts together with estimated values for another 450 stations (closed stations, Northern Ireland stations and dummy values). The resulting map (figure 7) achieves most of the accuracy of that produced manually, but without the necessity for any manual intervention.

This method was applied to a number of months (without a recalculation or reselection of the dummy values) and it continued to produce good results. The interpolation accepts both real and estimated values and treats them in the same way.

This process has a number of advantages. It is objective and requires no manual
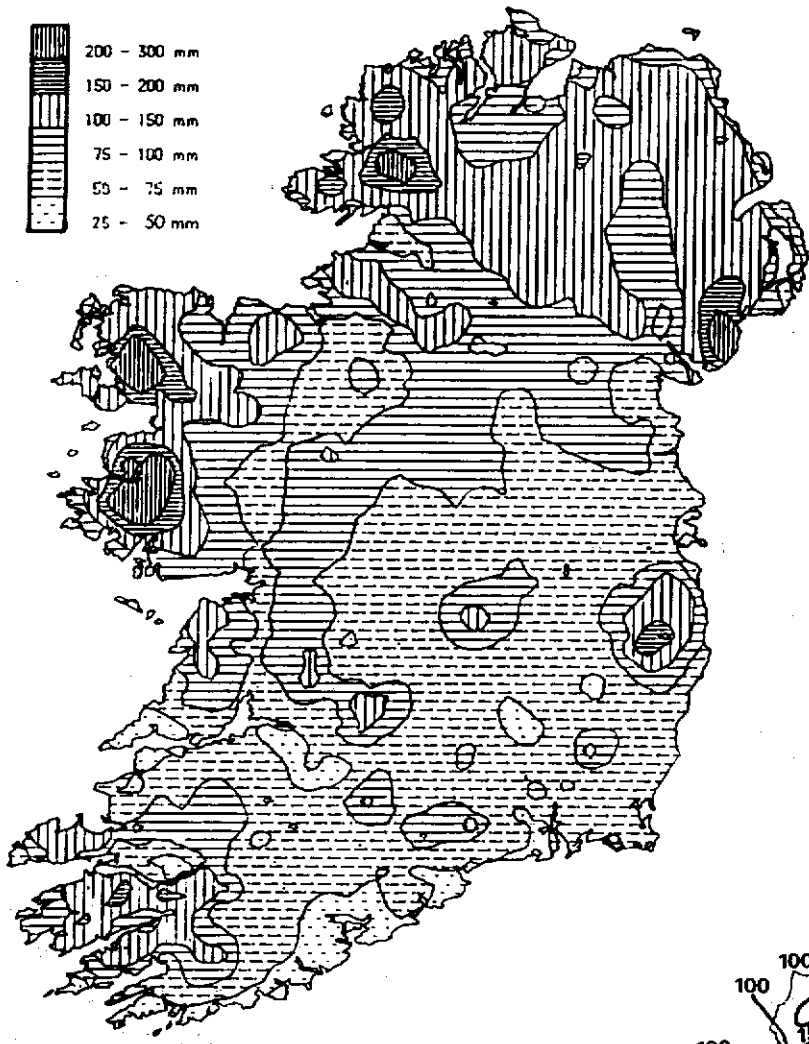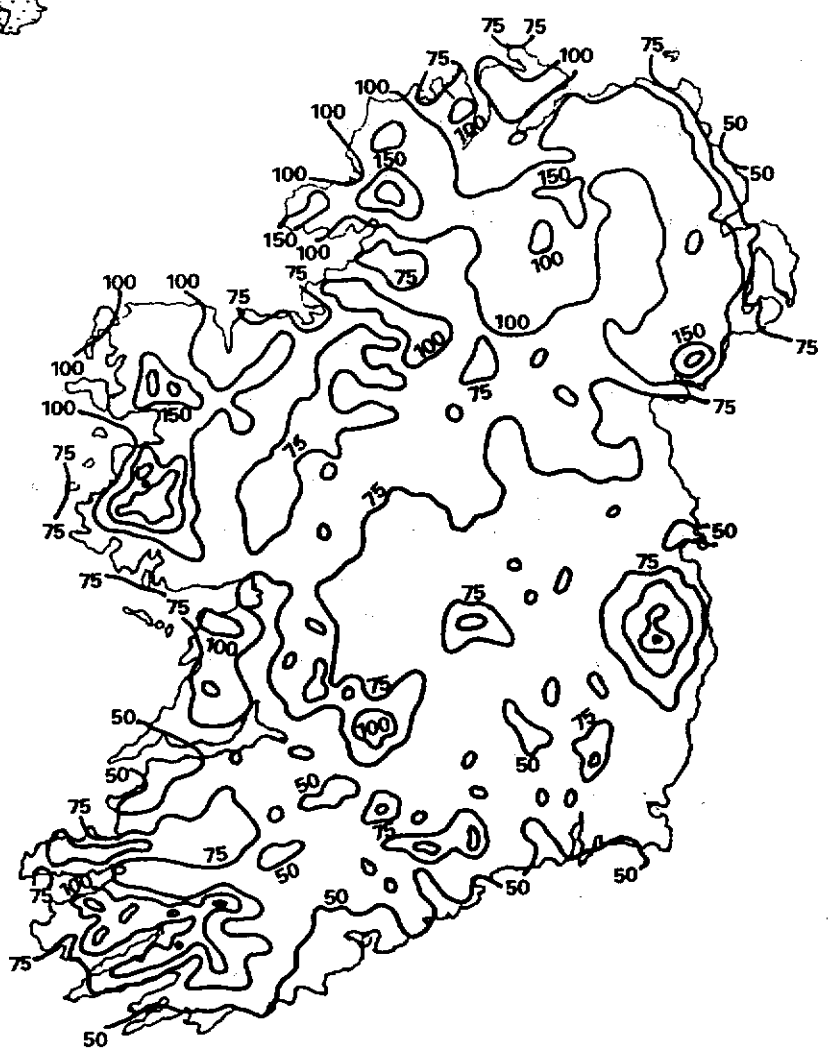
Figure 4: Computer analysis of rainfall amounts using data points and 40 manually estimated dummy values.

Figure 5: Manual analysis corresponding to figure 4.

April 1986

200 - 300 mm
150 - 200 mm
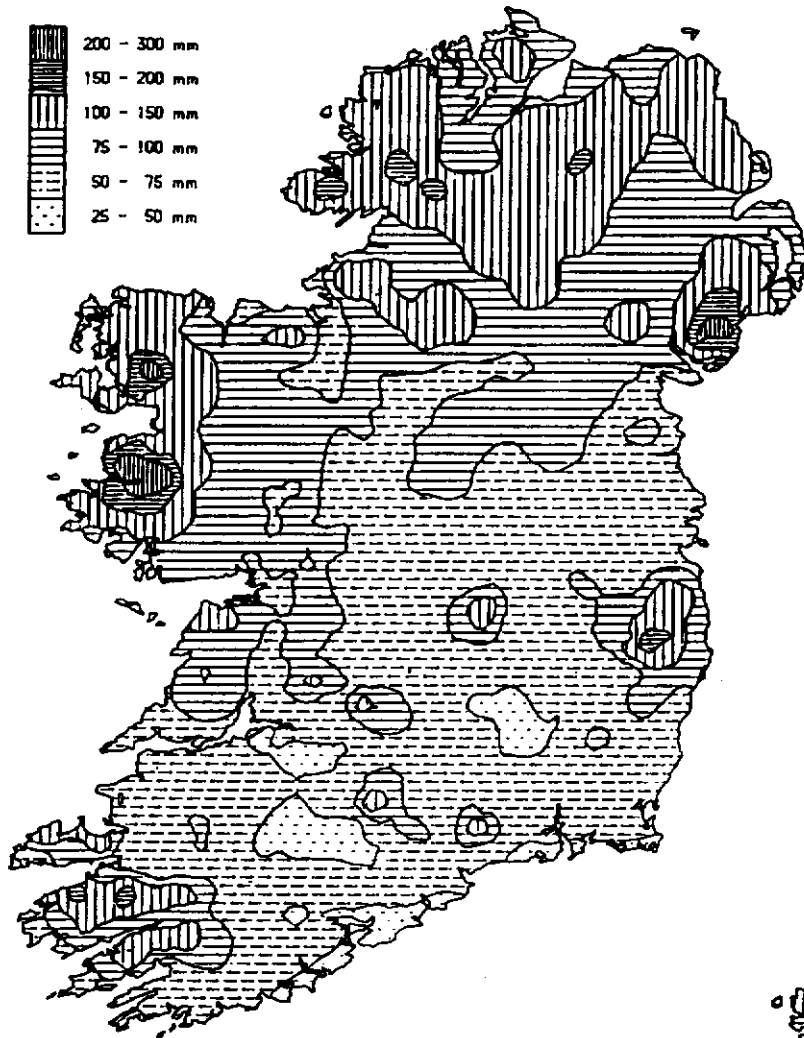100 - 150 mm
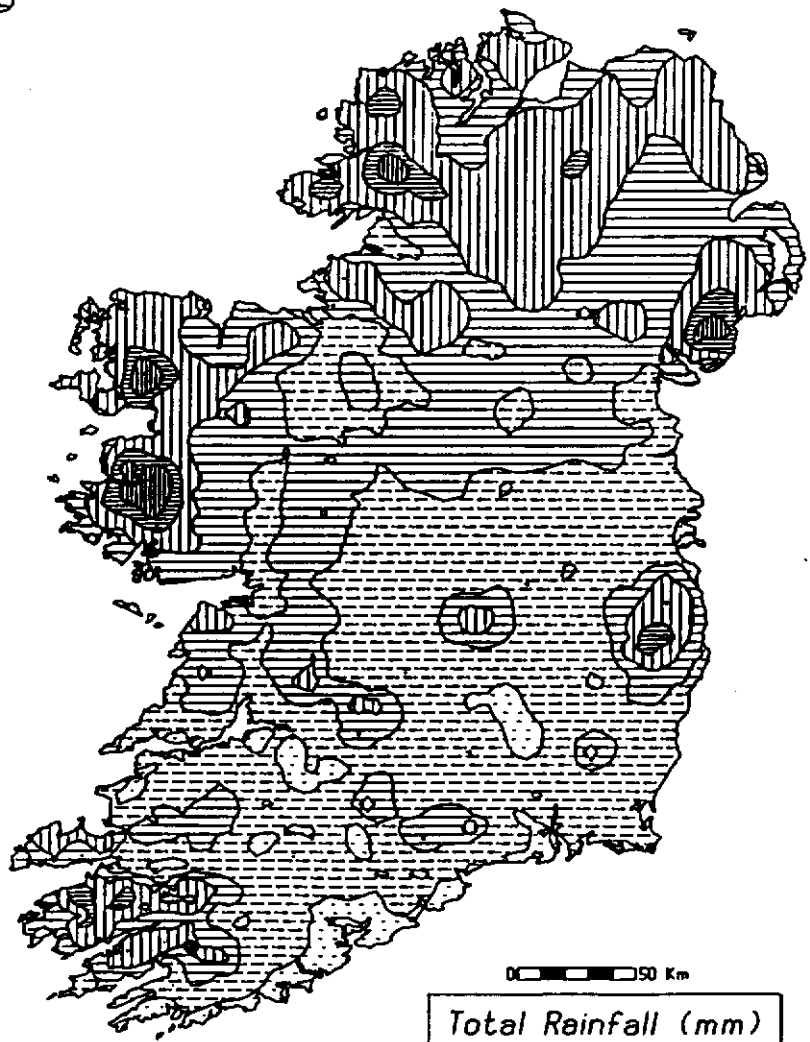75 - 100 mm
50 - 75 mm
25 - 50 mm

Figure 6: Estimated rainfall map using 30 year normals and an average percentage value for each county.

Figure 7: Rainfall map based on real data and estimated values from figure 6. This is the operational product.

0 ▬▬▬▬ 50 Km

Total Rainfall (mm)

intervention. Another advantage is that any slight modification considered necessary can be made simply by the addition of a dummy annual average and an associated grid position. Finally note that this method of modelling the orography is independent of the method of interpolation and could be used with other methods of objective analysis to provide a preliminary field. While there are 450 estimated values included in the interpolation, a first guess would require 10,800 points (one for each point of the 90x120 grid). The estimated values used here are independent of the grid resolution.

The 450 estimated values represent 350 closed stations, 70 Northern Ireland stations and 30 dummy points. The closed stations and Northern Ireland stations are based on real data and it is gratifying that only 30 dummy values have had to be used. The latter are based on a careful reading of the charts of rainfall averages (Fitzgerald 1984) and they mainly model orographic effects in the absence of observations.

The method is reasonably efficient. It takes approximately two minutes CPU time for the interpolation and another two minutes CPU time for the contouring and shading routines on a DEC 2050 computer. See Table 1 for a summary of the interpolation parameters used.

B. Conclusions

This paper has discussed an objective method of drawing climatological maps of sunshine, temperature, rainfall percentage and rainfall amount. The interpolation is based on the method of distance weighted least squares quadratic approximation. The interpolation parameters for the various fields are summarised in table 1.

The interpolation gave some spurious results near the coast and in data void areas when it was applied in its original form. These problems were overcome by the introduction of dummy values in such regions and a method was developed for calculating such values automatically based on nearby observations and various climatological normals.

The method has been in operational use since February 1985 and the results are of quality comparable to the manual method which this computer method has replaced.

Northern Ireland was provided by the Northern Ireland Office of the British Meteorological Office.

References

Cressman, G.P. 1959. 'An Operational Objective Analysis System', *Monthly Weather Review*, Vol. 87, p.367-374.

Fitzgerald, D. 1984. *Monthly and Annual Averages of Rainfall for Ireland 1951-80*, Climatological Note No. 7, Meteorological Service, Dublin 9, Ireland.

Gandin, L.S. 1965. *Objective Analysis of Meteorological Fields*, Jerusalem, Israel program for scientific translations, translated from Russian (Leningrad 1963).

Gustafsson, N. 1981. 'A Review of Methods for Objective Analysis', in Benftsson L., Ghil M., and Kallen E. (eds), *Dynamic Meteorology : Data Assimilation Methods*, Applied Mathematical Sciences Series, Vol. 3, Springer-Verlag, New York.

Hamilton, J.E.M. 1985. 'PCONTR : Contouring Package User's Guide', Internal Memorandum no. 108/85, Meteorological Service, Dublin 9, Ireland.

Maling, D.H. 1973. *Coordinate Systems and Map Projections*, George Philip and Son Limited, London.

Meteorological Service, 1985. *Monthly Weather Report, Part 1: General Weather Report for December 1985*, Meteorological Service, Dublin 9, Ireland.

McLain, D.H. 1974. 'Drawing Contours from Arbitrary Data Points', *Computer Journal*, Vol. 17, No. 4, p. 318-324.