THE ACCURACY OF ESTIMATES OF AREAL MEAN RAINFALL

Peter HUTCHINSON
Civil Engineering Dept., University of Aston, Birmingham, U.K.
Formerly
Geography Dept., University of Otago, Dunedin, N.Z.*

SUMMARY
Since data from adjacent raingauges are correlated, and therefore not independent, the simple method of calculating the standard error of the mean rainfall on an area cannot be used. A method for calculating this standard error, using time series analysis, derived originally by Jowett is presented, and examples from two areas, one of which is an experimental catchment, are given. It is shown that the errors of estimating areal mean rainfall can be considerable. These errors can be reduced in large areas by installing additional gauges, but in small experimental catchments it is the nature of the raingauge itself, rather than the design or density of the network, which provides the greater source of error.

RESUME
LA PRÉCISION DES ESTIMATIONS DE LA PRÉCIPITATION SPATIALE MOYENNE
Comme les données de pluviomètres adjacents sont en relation réciproque, et de ce fait ne sont pas indépendantes, la méthode simple de calcul de l'erreur standard de la précipitation spatiale moyenne ne peut pas être utilisée. Une méthode de calcul de cette erreur standard, utilisant l'analyse des séries temporelles, imaginée à l'origine par Jowett, est présentée et des exemples de deux aires sont donnés, l'une d'elles étant un bassin expérimental. Il est montré que les erreurs d'estimation de la précipitation spatiale moyenne peut être considérable. Ces erreurs peuvent être réduites pour de grandes surfaces en ajoutant des pluviomètres additionnels, mais pour les petits bassins expérimentaux, c'est la nature même du pluviomètre, plutôt que le plan et la densité du réseau, qui introduit la plus grande source d'erreurs.

1. INTRODUCTION
For many hydrological purposes it is becoming increasingly important to establish accurately the mean rainfall over an area for particular periods of time. The several methods of doing this are well known, but because rainfall measurement using raingauges is a sampling method the estimates of the mean rainfall are subject to sampling error. The present level of hydrological knowledge is such that these errors cannot be disregarded and must be assessed. Crawford and Linsley (1966), for example, have been able to simulate, by digital methods, the performance of hydrological systems to a high degree of accuracy, yet ultimately the usefulness of such models depends on a knowledge of the accuracy of the input data. It is essential, therefore, not only to be able to determine the mean rainfall over an area, but also to be able to estimate the accuracy of this determination. It is also useful, when planning experimental and representative basins and other hydrological networks, to be able to determine the raingauge network density required to achieve a given accuracy of estimates of the mean value.

The difficulty in estimating the errors lies in the fact that the observations are not independent, since correlations between adjacent gauges will usually be high, and therefore normal methods of estimating the error of a mean cannot be used.

Previously, several methods have been used to overcome this difficulty. McGuiness (1963) has compared the mean values calculated from sparse network data formed by

* Paper communicated on the author's behalf by G. H. Jowett, Professor of Statistics, University of Otago, New Zealand.
random selection from a dense network with the mean value estimated from the latter, and determined, for Coshocton, Ohio, the relationship

\[ E = 0.03 P^{0.54} G^{0.24} \]

where:
- \( E \) is the 'error';
- \( G \) is the gauging ratio in square miles per gauge;
- \( P \) is the mean rainfall in inches for the network.

This method assumes that the determination of mean rainfall from the dense network is liable to negligible error, and suffers from the practical difficulty of needing a dense network of gauges, which may not be available in a particular geographical area. Sutcliffe (1966) has derived the formula

\[ S = \sqrt{\frac{1-r}{n}} S_x, \]

for the standard error of the estimate, \( S_x \) being the standard deviation; \( n \), the number of gauges; and \( r \) the correlation coefficient between data for the gauge network for two independent time periods. This formula has been applied to annual and monthly data, but not, as far as is known, to daily rainfall. The difficulty with this method, particularly for short time periods, is that no unique result can be arrived at, since \( r \) will vary according to which two time periods are chosen.

A third approach is that used by Hershfield (1967) and Hutchinson (1969a). Without trying to assess the errors of the estimate of the mean, these authors have postulated that the estimate will be sufficiently accurate, if the network is so designed that adjacent gauges are highly correlated, Hershfield taking a value of \( r = 0.9 \) to be a desired standard.

A method similar to that described in this paper has been developed by Czelnai and others (1963). Their structural function is very similar to the serial variation function used in this paper.

The present study takes as its theoretical basis the method first developed by Jowett (1955), being an application of that work, and is an extension of the research reported by Hutchinson (1969a, b; 1970).

2. Mathematical Derivation

This section of the paper is a condensation of the relevant parts of Jowett's (1955) paper, and readers are referred to the latter for the full derivation.

The data, for a particular time period, day, month or year, are considered to be drawn from a two dimensional "time series", in which 'time' is replaced by space as represented by an orthogonal co-ordinate system. Where desired, a set of similar time periods can be taken, in which case the data become an ergodic array of such series. The former case would apply to the determination of the error of a single time period, while the latter is more useful in developing general relationships for an area.

Bayley and Hammersley (1946) have given the formula for the sampling variance of the mean \( \bar{x} \) of a section of an evenly spaced time series \( x_i \) with mean \( \mu \), variance \( \sigma^2 \) and serial correlation \( \rho_s \) as

\[ \sigma^2 = \frac{\sigma^2}{n} \sum_{s = -(n-1)}^{+(n-1)} \left( 1 - \frac{|s|}{n} \right) \rho_s \]  

where \( n \) is the number of consecutive terms from the series, and \( s \) is the lag.
If the serial variation parameter $\delta_s$ is defined as
\[
\delta_s = E \frac{1}{2} (x_i - x_{i+s})^2 = \sigma^2 (1 - \rho_s) \quad (2)
\]
the formula can be rewritten
\[
\sigma^2_s = \sigma^2 - \frac{1}{n} \sum_{s=(n-1)}^{s=(n-1)} \left( 1 - \frac{|s|}{n} \right) \delta_s \quad (3)
\]
now $\nu$, the within sample variance of the given section, is given by
\[
\nu = \frac{1}{2n^2} \sum_{i,j=1}^{n} (x_i - x_j)^2 \quad (4)
\]
then
\[
E(\nu) = \sigma^2 - \frac{\sigma^2}{n} \sum_{s=(n-1)}^{s=(n-1)} \left( 1 - \frac{|s|}{n} \right) \rho_s \quad (5)
\]
\[
= \frac{1}{n} \sum_{s=(n-1)}^{s=(n-1)} \left( 1 - \frac{|s|}{n} \right) \delta_s \quad (6)
\]
Since the right-hand sides of equation (3) and (6) are identical, it will be possible to estimate $\sigma^2_s$ if an estimate of $\sigma^2$ can be made.

From equation (2) it may be seen that $\delta_s \to \sigma^2$ as $s$ increases, so that $\delta_s$ may be used provided that $\delta_s$ attains a limit within the data used, i.e. for $s < n$.

Hence in practice, it is necessary to calculate the totals of squared differences
\[
D_s = \sum_{i=1}^{(n-s)} (x_i - x_{i+s})^2 \quad (7)
\]
for $s = 1, 2, \ldots (S_0 - 1)$.

This gives an estimate $v_{s_0}$ as follows
\[
v_{s_0} = \left[ \sum_{i,j(|j-i| \geq s_0)} \frac{1}{2} (x_i - x_{i+j})^2 \right] \left[ \sum_{i,j(|j-i| \geq s_0)} 1 \right]
\]
\[
= \left[ n \sum_{i=1}^{n} (x_i - \bar{x})^2 \right] - \frac{s_0-1}{s=1} \sum_{s=1}^{s_0-1} D_s \right/ (n - s_0) (n - s_0 + 1) \quad (8)
\]
which may be used in the equation
\[
\sigma^2_s \sim v_{s_0} - \nu \equiv \text{Est.} \sigma^2_s \quad (9)
\]
It is necessary to decide on a value for $s_0$, and hence $v_{s_0}$, and this may be done using the serial variation statistic, $d_s$, where
\[
d_s = \frac{1}{2} D_s / (n - s) \quad (10)
\]
$d_s$ is also known as the Mean Semi-squared Differences (M.S.S.D.) This calculation is usually laid out in the form of table 1, while a limiting value of $s_0$ can be chosen according to the point at which $d_s$ reaches its limiting value, this choice being more easily made by plotting the serial variation curve, $\delta_s$ against $s$. 

2.5 205
### Table 1. Analysis of Semi-squared Differences for the Sampling Variance of a Single Mean

<table>
<thead>
<tr>
<th>Nature of Differences</th>
<th>Sum of Semi-squared Differences</th>
<th>Number</th>
<th>Mean Semi-squared Difference</th>
<th>Value</th>
<th>Expected Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\sum_{i,j=1}^{n} \frac{1}{2}(x_i - x_j)^2$ ((= n \sum_{i=1}^{n} (x_i - \bar{x})^2))</td>
<td>$n^2$</td>
<td>$v$</td>
<td>$\frac{1}{n} \sum (1 - \frac{</td>
<td>s</td>
</tr>
<tr>
<td>Lag 0</td>
<td>0</td>
<td>$n$</td>
<td>0</td>
<td>$n_s = -(n-1)$</td>
<td></td>
</tr>
<tr>
<td>Residual (</td>
<td>Lag</td>
<td>≥ 1)*</td>
<td>$\sum_{</td>
<td>i-j</td>
<td>\geq 1} \frac{1}{2}(x_i - x_j)^2$</td>
</tr>
<tr>
<td>Lags±1</td>
<td>$D_1 = \sum_{i=1}^{n-1} (x_i - x_{i+1})^2$</td>
<td>$2(n-1)$</td>
<td>$d_1$</td>
<td>$\delta_1$</td>
<td></td>
</tr>
<tr>
<td>Residual (</td>
<td>Lag</td>
<td>≥ 2)*</td>
<td>$\sum_{</td>
<td>i-j</td>
<td>\geq 2} \frac{1}{2}(x_i - x_j)^2$</td>
</tr>
<tr>
<td>Lags±2</td>
<td>$D_2 = \sum_{i=1}^{n-2} (x_i - x_{i+2})^2$</td>
<td>$2(n-2)$</td>
<td>$d_2$</td>
<td>$\delta_2$</td>
<td></td>
</tr>
<tr>
<td>Residual (</td>
<td>Lag</td>
<td>≥ $s_0$)*</td>
<td>$\sum_{</td>
<td>i-j</td>
<td>\geq s_0} \frac{1}{2}(x_i - x_j)^2$</td>
</tr>
<tr>
<td>Lags±$s_0$</td>
<td>$D_{s_0} = \sum_{i=1}^{n-s_0} (x_i - x_{i+s_0})^2$</td>
<td>$2(n-s_0)$</td>
<td>$d_{s_0}$</td>
<td>$\sigma^2 + 0(e)$</td>
<td></td>
</tr>
</tbody>
</table>

* By subtraction.
### Table 2. Analysis of Semi-squared Differences

<table>
<thead>
<tr>
<th>Nature of Differences (Separations)</th>
<th>Between Stations for Corresponding Storms</th>
<th>Correction for Station Means</th>
<th>Deviations from Station Means</th>
<th>Revised MSSD</th>
<th>Revised Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sum</td>
<td>Number</td>
<td>Sum</td>
<td>Number</td>
<td>MSSD</td>
</tr>
<tr>
<td>All</td>
<td>149.6</td>
<td>10000</td>
<td>31.58</td>
<td>900</td>
<td>0.03509</td>
</tr>
<tr>
<td>Residual Lag 1 (5 miles)</td>
<td>149.6</td>
<td>4350</td>
<td>31.58</td>
<td>870</td>
<td>0.03630</td>
</tr>
<tr>
<td></td>
<td>1.443</td>
<td>40</td>
<td>0.1665</td>
<td>8</td>
<td>0.03061</td>
</tr>
<tr>
<td>Residual Lag 2</td>
<td>149.2</td>
<td>1310</td>
<td>31.54</td>
<td>852</td>
<td>0.03644</td>
</tr>
<tr>
<td></td>
<td>1.845</td>
<td>120</td>
<td>0.1978</td>
<td>24</td>
<td>0.00824</td>
</tr>
<tr>
<td>Residual Lag 3</td>
<td>117.3</td>
<td>1390</td>
<td>31.21</td>
<td>838</td>
<td>0.03725</td>
</tr>
<tr>
<td></td>
<td>8.047</td>
<td>250</td>
<td>2.112</td>
<td>59</td>
<td>0.02642</td>
</tr>
<tr>
<td>Residual Lag 4</td>
<td>139.3</td>
<td>3900</td>
<td>30.10</td>
<td>780</td>
<td>0.03731</td>
</tr>
<tr>
<td></td>
<td>8.211</td>
<td>310</td>
<td>1.655</td>
<td>62</td>
<td>0.02669</td>
</tr>
<tr>
<td>Residual Lag 5</td>
<td>133.1</td>
<td>3590</td>
<td>27.45</td>
<td>718</td>
<td>0.03823</td>
</tr>
<tr>
<td></td>
<td>10.80</td>
<td>380</td>
<td>2.413</td>
<td>76</td>
<td>0.03176</td>
</tr>
<tr>
<td>Residual Lag 6</td>
<td>120.3</td>
<td>3210</td>
<td>25.03</td>
<td>642</td>
<td>0.03899</td>
</tr>
<tr>
<td></td>
<td>7.577</td>
<td>300</td>
<td>1.607</td>
<td>60</td>
<td>0.02678</td>
</tr>
<tr>
<td>Residual Lag 7</td>
<td>112.7</td>
<td>2910</td>
<td>23.43</td>
<td>582</td>
<td>0.04025</td>
</tr>
<tr>
<td></td>
<td>12.87</td>
<td>340</td>
<td>2.168</td>
<td>68</td>
<td>0.03189</td>
</tr>
<tr>
<td>Residual Lag 8</td>
<td>99.88</td>
<td>2570</td>
<td>21.26</td>
<td>514</td>
<td>0.04136</td>
</tr>
<tr>
<td></td>
<td>10.38</td>
<td>310</td>
<td>2.699</td>
<td>62</td>
<td>0.04354</td>
</tr>
<tr>
<td>Residual Lag 9</td>
<td>89.49</td>
<td>2260</td>
<td>18.56</td>
<td>452</td>
<td>0.04106</td>
</tr>
<tr>
<td></td>
<td>14.80</td>
<td>460</td>
<td>3.449</td>
<td>99</td>
<td>0.03749</td>
</tr>
<tr>
<td>Residual Lag 10</td>
<td>70.65</td>
<td>1800</td>
<td>15.11</td>
<td>360</td>
<td>0.04198</td>
</tr>
<tr>
<td></td>
<td>14.74</td>
<td>360</td>
<td>2.612</td>
<td>72</td>
<td>0.03677</td>
</tr>
<tr>
<td>Residual Lag 11</td>
<td>55.91</td>
<td>1440</td>
<td>12.50</td>
<td>298</td>
<td>0.0434</td>
</tr>
<tr>
<td></td>
<td>19.91</td>
<td>340</td>
<td>2.574</td>
<td>68</td>
<td>0.03511</td>
</tr>
</tbody>
</table>
Where the M.S.S.D’s are erratic, it is necessary to revise the estimate of \( v \). This is done by sketching in a smoothed serial variation curve as shown in figure 1, replacing \( d_1, d_2 \ldots d_{sv} - 1 \) and \( v_{so} \) by the fitted values \( \delta_1, \delta_2 \ldots \delta_{sv} - 1, \delta_{so} \) and, and working backwards to give a revised estimate of \( v \). This is shown as the two right-hand columns of table 2, which is the same example as shown in figure 1. The reasons for doing this are explained in the original Jowett (1955) paper.

![Serial Variation Curve. Example of Grouped Daily Data from the Otago Area.](image)

In order to extend the method to the application to an array of series, it is only necessary to account for the variance of the replication means \( \bar{x}_1, \bar{x}_2 \ldots \bar{x}_k \), where \( k \) is the number of replications. This is done by replacing semi-squared differences by semi-squared deviations of differences, which in practice involves carrying out a supplementary analysis of semi-squared differences on the \( n \) pattern totals, dividing each sum of S.S.D’s by \( k \) and to subtract the result from the corresponding sum of S.S.D’s. This is shown as ‘correction for station means’ in table 2.

Finally the method is generalised to apply to spatial, unevenly-spaced data. Provided the data are isotropic, which is true in the examples used, the linear distance between each pair of gauges is calculated, and the data grouped according to the distances, the group intervals corresponding to the lag. In the example shown, the lag is taken as 5 miles (8 km).

In the examples given below, the square root of the calculated variance is used, and called the standard error of the mean.
3. Application of the Method

Three applications of the method are shown. These have been chosen to show the range of data which can be treated, and to give some results which are of interest in the New Zealand context.

3.1. Otago Test Area

Rainfall data from part of the Otago Province, in the south of the South Island, New Zealand, have been used in previous analyses of various kinds by the author. The area covers 2300 square miles (6960 sq. km) being the catchment of the Taieri River and some adjacent areas (fig. 2), its topography varying between flat and sharply accentuated, with elevations from sea level to 5500 ft (1677 m). Data from 30 raingauges, distributed as shown in the figure, were used, annual and monthly values being used for the period 1955 to 1964, and daily values for the year 1966. Apart from their general interest, this and previous analyses are pertinent to a water resources survey of the Taieri River which is currently being carried out by the New Zealand Ministry of Works, with a view to the provision of irrigation and town water supply.

![Location of Otago Rainfall Stations](image)

**Figure 2. Location of Raingauges, Otago Test Area**

3.1.1. Daily Data

The mean rainfall for each of 100 days on which rainfall was general over the area was calculated as the arithmetic mean for all the gauges, no weighting coefficients being used. The means were then grouped into intervals of 0.05 inches (1.27 mm) up to 0.25 inches (6.35 mm) and greater than 0.25 inches (6.35 mm), giving, for each group a set of two-dimensional spatial time series. Using the method outlined above, and exemplified in figure 1 and table 2, the standard error of the mean daily rainfall for each group was 2.9
calculated. The standard error was then plotted against daily mean rainfall, giving the figure 3. This shows that the standard error depends to some extent on the mean value, a result which corresponds to conclusions reached in other parts of the world (McGuiness, 1965; Herbst and Shaw, 1969).

However, the values of the standard error, which give a figure of approximately 60% throughout the range, when recalculated as the coefficient of variation are considerably higher than values quoted elsewhere, even for areas of a similar size. The reasons for this are due to the wide variations in topography, which reduces inter-station correlation, and due to the variations of storm type experienced. All three of Gibbs (1964) types occur in this region, contributing factors being the synoptic situations, convective action, and land-sea effects.

The purpose of this analysis was to provide a working rule which could be used for daily data in this area. In addition to the grouping technique described above, an attempt was made to derive a standard-error/mean-daily-rainfall relationship, using the 100 days individually. The results, however were too widely scattered to be useful.

3.1.2. Monthly Data

For many purposes, particularly for the assessment of water resources, it is more convenient to work with monthly rainfall total than with daily values. The period covered was ten years, giving 120 monthly data sets.

The standard error of the estimate of the mean was calculated for each month separately, but without revising the estimate of \( \nu \), due to the massive computation involved. There is, therefore, some error, but this is not large, since the serial variation curves were less erratic than that shown in figure 1. The estimates of the mean varied between 0.799 ins (20.3 mm) and 4.828 in. (122.5 mm) while standard errors ranged from 0.10 in (2.54 mm) to 1.98 in (50.2 mm).
The accuracy of estimates of areal mean rainfall

Overall, there was slight association between the mean and its standard error. When broken down into seasons, and with standard error transformed logarithmically, winter (June, July and August) showed the highest correlation \( r = 0.608 \) with autumn also being significant \( (0.513) \), figure 4. Spring showed a weak correlation \( (0.351) \), but there

![Figure 4](image-url)

**Figure 4.** Variation of Standard Error with Mean Monthly Rainfall, Autumn and Winter, Otago Test Area
was no apparent association for summer \( (r = 0.107) \). This may be explained by the nature of the rain-producing mechanisms. In summer, there are a greater proportion of smaller, highly variable convectional rainstorms, than in winter, which is dominated by synoptic type rainfall, i.e. while it is generally acknowledged that there is a direct relationship between the mean and its standard error, this is more apposite to synoptic rainfall. With convectional rainfall, consequent upon its nature, and taken on a monthly basis over a specified area, the relationship is much less marked. This is illustrated for systematic variation in rainfall in a previous paper (Hutchinson, 1968), which shows more definite relationships during winter months.

While it is possible to calculate the standard error for each month, it is interesting, for the purposes of deriving a general rule for the area, to note the seasonal variation.

![Seasonal Variation of Standard Error](image)

**Figure 5. Seasonal Variation of Standard Error, for individual months, Otago Test Area**

Figure 5 shows quantities relating to the ten values for each month of the year for the whole period, and, although the seasonal variation is not a smooth one, it is possible to see that the estimates of the monthly mean rainfall are generally less accurate for the autumn/winter period than for the rest of the year. This may well be due to some of the precipitation in winter being in the form of snow, with the consequent drifting and highly inaccurate catch. It is known, for example, that the effect of wind on the catch is greater with snow than with rain. Recalculated as coefficients of variation, the summer figure of 12% and winter figure of 35% compare unfavourably with the study carried out by Czelnai and others (1963), which, for a similar area and similar number of gauges, gave corresponding values of 5% and 3%.
The accuracy of estimates of areal mean rainfall

Evidently in this area there are insufficient raingauges to give accurate estimates of daily or monthly rainfall, a conclusion which coincides with the opinions of the New Zealand Ministry of Works, who have installed a number of additional gauges in the area.

A further point to be considered is illustrated in figure 6. For practical purposes it is unlikely that it would be possible to set up a network sufficiently dense to give a satisfactory accuracy for every month. With any given network, the mean areal rainfall for some months can be judged quite accurately, while for others the limits are greater than the mean itself. To the question: "How many gauges are needed to give a result to a specified accuracy?", must be added another question: "For what proportion of the time are we satisfied to accept this accuracy?".

If the distribution in figure 6 is typical, then we would have to accept that it becomes increasingly difficult to assess accurately the mean for over 80% of the total months.

3.2. Moutere Experimental Basin

Moutere experimental basin is situated in the Nelson Province in the north of the South Island of New Zealand, its purpose being to compare the effects of different agricultural practices on the hydrological regimes. The data are taken from ten raingauges in No. 5
The accuracy of estimates of areal mean rainfall, shown in figure 7. It would be expected that such a high network density would give a very accurate determination of mean rainfall, and the purpose of this analysis was to calculate the possible errors involved. Data from 254 individual storms were grouped according to storm mean rainfall, giving 7 groups with from 5 to 25 in each group. The lag chosen was 2 chains (40.234 metres), but the network was such that no distance fell within lag 1, while at lag 2 the value \( \delta s \) had been attained in most cases. Thus it was not necessary to calculate a revised estimate of \( s \). The estimated values of \( \sigma_x \) plotted against mean storm rainfall gave a similar result as for the Otago data in that the standard error increases with storm rainfall (fig. 8). For the highest showers (0.00-0.05 in. mean value (0.0347 in., 0.00-1.28 mm mean value 0.88 mm) the standard error was calculated as 0.00144 in., (0.036 mm) which, taking 2 s.e. as the usual criterion gave limits of \( \pm 0.00288 \) in. (0.072 mm) for the true value of the mean, and which gives limits in percentage terms of 8.2\%. The corresponding values for the heaviest group of storms (1.50 in. mean 2.022 in., 32 mm mean 51.2 mm) are 0.045 in. (1.14 mm) for the s.e., \( \pm 0.090 \) in. (2.28 mm) for the limits equivalent to 4.5\%. For some purposes it may be considered that these values are too high, and that the network density should be increased. It is known, however, that there are systematic errors associated with the catch of a single raingauge, shown by many workers, as well as a purely random source of error (Hutchinson, 1969b) which is of the same order of magnitude as the errors calculated above. The addition of further raingauges, therefore, may not improve the accuracy of the determina-

---

**Figure 8. Variation of Standard Error with Mean Daily Rainfall, Moutere Experimental Station**

---
tion of mean rainfall over such a small area. This, apparently, may only be done by improving the accuracy of the raingauge itself.

4. CONCLUSIONS

The techniques described in this paper are not new; utilisation of the technique however had previously been held up because of the massive amount of calculation involved. With the aid of a computer, however, the method is now convenient to use. With a suitable programme, as shown in the Appendix, all that is required are the spatial co-ordinates of the raingauges, and the rainfall data.

The results, taken from two contrasting areas, one a large area with many sources of rainfall variation, the other a small area where variations in rainfall would be expected to be very small, show that possible errors in the determination of areal mean rainfall for a storm can be quite high. In one case it is suggested that the accuracy can be increased by increasing the raingauge density, but it is considered doubtful if this would be of benefit in the other area.

5. ACKNOWLEDGEMENTS

The author is particularly indebted to Professor G. H. Jowett, both for advising on the statistical analysis, and for reading this paper; and to Mr B. Grigg, who wrote the computer programme. Acknowledgements are also due to the Commissioner of Works, Wellington, for permission to publish the data from Moutere, and to Messrs. A. G. Gillingham and F. Crimp, Ministry of Works, Nelson, for the collection of these data.

Acknowledgement is also made to the New Zealand Meteorological Service for the data collected from the Otago test area.

REFERENCES


The accuracy of estimates of areal mean rainfall

APPENDIX

// J.G. TIMSER (1943) B. GRIGG, UNIVERSITY OF OTAGO COMPUTING CENTRE
// OPTION LINK
// EXEC FORTRAN
C THIS PROGRAM CONDUCTS AN ANALYSIS OF SEMI-SQUARED DIFFERENCES FOR UP
C TO M SPATIAL, UNEVENLY-SPACED, NONSTATIONARY SERIES OF N POINTS.
C
C CLASS-- IF THE SERIES IS SPATIAL OR UNEVENLY-SPACED, CLASS SPECIFIES
C THE WIDTH OF THE GROUPED SETS OF DIFFERENCES REMOVED FROM
C THE RESIDUAL SSD'S.
C LIMITATIONS -- N = THE NO. OF POINTS IS PRACTICALLY UNLIMITED.
C N = THE NO. IN EACH SERIES FOR EACH POINT MAY BE 100.
C
C CLASS M & N ARE READ IN FIRST FOLLOWED BY MAP COORDINATES FOR UNEVEN
C SERIES THEN THE N SERIES IN CORRESPONDING ORDER.
C
DIMENSION DIST(100), DS(30), X(100), Y(100), KOUNT(30)

DEFINE FILE 9(2800, 100, U, JK)
READ(1, 100) CLASS, N, M
NI=N-1
5 DO 10 I=1, N
10 WRITE(9, 1) X(I), Y(I)

C CALCULATE THE DISTANCES BETWEEN ALL N POINTS & STORE ON DISC.

DO 30 J=1, NI
READ(9, 10) X(I), Y(I)
NJ=N-J
DO 20 I=1, NJ
READ(9, 10) X(J), Y(J)
20 DIST(I)=SQRT((X(I)-X(J))*(X(I)-X(J))+(Y(I)-Y(J))*(Y(I)-Y(J))
WRITE(9, 500) J, IDIST
30 CONTINUE

CALL SKIP

LMAX=0

C READ IN THE SERIES FOR EACH OF THE N POINTS AND STORE EACH SERIES AS A RECORD
C OF UP TO 100 UNITS. DEPENDING ON THE WAY THE DATA IS ARRANGED ON THE CARDS
C THIS SECTION MAY HAVE TO BE CHANGED. IT IS POSSIBLE TO SORT THE DATA INTO
C GROUPS FOR ANALYSIS AS THE CARDS ARE READ IN.

DO 40 I=1, N
READ(1, 120) (Y(J), J=1, M)
40 WRITE(9, 1) Y(I)

C INITIALIZE O -- THE VECTOR WHICH WILL CONTAIN THE SSD'S AT LAG J BETWEEN
C POINTS WITHIN YEARS.
C DS - THE VECTOR OF SSD'S AT LAG J BETWEEN POINTS
C KOUNT - COUNTS THE NO. IN EACH CLASS

32 DO 31 J=1, 30
KOUNT(J)=0.0
DS(J)=0.0
31 DO 30 I=1, NI
DO 30 J=1, NJ
30 SUMY=0.0
DO 50 J=1, M
50 SUMY=SUMY+Y(J)

C READ THE BASE SERIES OFF THE DISC.
READ(9, 1) Y(I)
READ(9, 500) I
DIST SUMY=0.0
50 DO 70 J=1, NI
70 DIST(J)=CLASS+1.0
KOUNT(L)=KOUNT(L)+1

C READ THE OTHER SERIES OFF THE DISC IN TURN WITH THE CORRESPONDING DISTANCE

2.17 217
APPENDIX (continued)

IF (LMAX - L) > 154, 54, 55
54: LMAX = L
55: SUMX = 0.0
C  CALCULATE THE SSD'S
DO 60 K = 1, M
D(L) = (Y(K) - X(K)) * (Y(K) - X(K)) + D(L)
60 SUMX = SUMX + X(K)
70 DS(L) = (SUMY - SUMX) / (SUMY - SUMX) / M + D(L)
80 CONTINUE
SUMX = 0.0
SUMY = 0.0
DO 90 J = 1, 30
SUMY = SUMY + D(J)
90 SUMX = SUMX + DS(J)
C  CALCULATE APPROPRIATE DIVISORS AND PRINT THE RESULTS.
INDEX2 = N*M
INDEX1 = INDEX2 / M
SUM = SUMX + SUMY
SSD1 = SUMX / INDEX1
SSD2 = SUMY / INDEX2
IF (INDEX1) > 160, 170, 180
160 SSD1 = SUMX / INDEX2
170 IF (INDEX4) > 180, 185, 190
180 DBAR1 = DS(1) / INDEX4
185 IF (N) > 186, 187, 188
186 DBAR2 = DSUM / N
188 IF (N) > 189, 190, 191
190 WRITE (3, 210) SUMY, INDEX1, SUMX, INDEX2, SSD1, SUM, NO, SSD2
WRITE (3, 220) INDEX1 / M, INDEX2, NO, DBAR1, DBAR2
SSD1 = SSD1 / INDEX2
SSD2 = SSD2 / INDEX4
END