311.15
551.501.45

A SIMPLIFIED TECHNIQUE OF STATISTICAL MAXIMISATION OF RAINFALL WITH NORMAL DISTRIBUTION OF LOGS

by S. Karelsky

Central Office, Bureau of Meteorology, Melbourne

(Manuscript received August 1961)

Abstract: Measurements of rainfall by a rain gauge represent the actual amount of water with a relative error of the order 15 to 20 per cent of the measured value. Calculations of maximum rainfall with better relative accuracy is therefore pointless. It means that all calculations could be made with values of rainfalls, logarithms etc. expressed by numbers with no more than two significant figures and the amount of technical work required for "maximisation" can be greatly reduced. This paper offers a method of "speeding" up these calculations.

1. INTRODUCTORY REMARKS

The greater the relative accuracy of any calculated physical quantities required, the more difficult and time-consuming the calculations needed, and the more complicated and lengthy are the tables (statistical, physical, logarithmic etc.) which should be used. The amount of writing (or typing) of values of the quantities involved increases rapidly with increasing number of significant figures representing these values.

On the other hand the relative accuracy of measurements of any physical (or meteorological) quantities has a limit implied by

(i) the imperfection of the measuring instruments,

(ii) the qualities of the measured quantity and of the medium which should be characterised by this quantity (for example, the temperature of the free air is fluctuating continuously, but the temperature of a solid body is more or less "steady"),

(iii) the difficulties of observations connected with measurement of large scale phenomena by a "point" instrument.

Relative errors in results of measurements and observations which cannot be overcome in practice for the three reasons above, lead us to a choice of a unit of measurement which should be used for a sensible representation of these results by numbers with a limited number of significant figures.

Thus, for example, the relative error in measurements of distances has a limit 0.01 per cent of the measured value when precise instruments and methods are used. This error means that $\frac{1}{10000}$ of the measured value may be incorrect and any measured distance can be represented by numbers with 5 significant figures with possible absolute error of the order ± 0.5 of the unit chosen. Ordinary measurement of distances have the relative error about ± 0.1 per cent (± 1 in. absolute error in every 1000 in.$\simeq$83 feet, of measured distance). It would be pointless therefore to use 1 in.

as a unit for measuring distance, say between Melbourne and Sydney along a railway line, and represent this distance as say 621 miles 12 yards 2 feet and 5 inches. The error ±½ mile is insignificant in this case and the distance should be quoted as 621 miles ±½ mile.

2. UNITS AND ACCURACY OF MEASUREMENTS OF RAINFALL

Measurements of rainfall by a rain gauge have the following errors:

(i) Pure instrumental and "reading" errors. The measuring glass has the capacity of ½ in. of water accumulated in the rain gauge. The observer can read off amounts in the glass with an absolute error of ±½ pt. (0.005 inch). The relative error is therefore ±1 per cent for ½ in. rainfall and greater for smaller values.

(ii) Errors arising from loss of water in the rain gauge due to evaporation, to incomplete emptying of the rain gauge etc. These errors can vary depending on many circumstances which cannot be taken into account quantitatively.

It is believed that the total error caused by (i) and (ii) may amount to 5 per cent of the measured value.

The standard rain gauge has an opening 8 in. in diameter (approximate area 50¼ sq. inch or 0.4 sq. feet). We assume that the height of water in this rain gauge represents the average height of the water layer over an area of the order of 1 sq. mile or 27,878,400 square feet. It is obvious that the accuracy of this representation cannot be high. This was realised already at the end of the last century and experimental simultaneous measurements of rainfalls by a number of rain gauges evenly distributed on a relatively small flat area around Berlin have been made by Hellman (1890) whose work was quoted by Hann and Suering (1926, p. 335). Hellmann concluded: "The difference between monthly totals can reach in some months up to 5 per cent for rain gauges placed evenly on a flat area at distances less than ½ km (approximately 1500 feet). The difference increases on individual days with squally weather and thunderstorms to such an extent that it can reach 100 per cent or even more. It is not possible to draw isohyets for these days with an accuracy of 10 mm (40 points)".

It means that:

(a) The measurement of the water in the rain gauge by observers with an error ±½ point and with 1 point as a unit, is justified for accurate recording of the amount of water actually found in the rain gauge.

(b) 1 point as a unit is much too fine for representation of the real amount of rainfall even over an area as small as ½ sq. mile. For rainfalls exceeding 1 in. and up to 3 in., a unit of 20 points (0.2 in.) with absolute error ±10 points and relative error ±10 per cent to ±3.3 per cent, would be more than satisfactory. For rainfalls over 3 in. a unit of ½ inch can be accepted without diminishing the relative accuracy of assessment of the amount of water fallen over an area of the order ½ to 1 sq. mile around the rain gauge. For example, a rainfall of 543 points having an error of ±25 points (½ in. unit) could represent any value of rainfall between 518 and 568 points (roughly 5.2 in. - 5.7 in.).

We can state now that:

(1) Series of yearly maximum 24, 48 or 72 hour rainfalls exceeding 1 in. and recorded with absolute error of

measurement ±½ point (0.005 inches) could be rounded of to the nearest 10 points (±5 points) or could be distributed into class intervals of 10 points each, counting frequency for each class interval (relative error on average will be less than ±5 per cent).

(2) Without diminishing the accuracy of results for practical purposes we can take class intervals of 20 points each for all those series, the lowest value of which is about 2 in. or divide all values of these falls by 20, round results to the nearest new unit of 20 points and count frequencies of these new values, with the average relative error less than 5 per cent.

(3) For a series of maximum rainfalls where the lowest value exceeds 3 in. and the highest value is over 10 in. we can take class intervals of 50 points each or introduce a new unit of ½ in. and round results up to the nearest whole numbers. The relative error of the final result of our calculation will be of the same order of ±5 per cent and even less, for the greatest value would exceed 10 in. with the absolute error ±25 points, i.e. with relative error less than ±2.5 per cent of the greatest value.

Having a series of 50 or more maximum 24, 48 or 72 hour values collected from records and distributing them into equal class intervals centred around values with only 2 significant figures, or introducing new units 10, 20 or 50 points each and expressing the collected values in the new units chosen, we will obtain a short sample of 2 figured values with their frequencies. After distributing them in ascending order we can find any function of these values with the same relative error as the values themselves. Assuming, for example, the normal distribution of logarithms, we can use only 2 figured logarithms (2 figured mantissae). But the table of 2 figured logarithms is very short and all work with this table will take only a few minutes (Appendix 1).

3. SIMPLIFIED TECHNIQUE OF CALCULATION OF RAINFALL WITH PROBABILITY $\frac{1}{100}$ AND $\frac{1}{1000}$ IN THE CASE OF NORMAL DISTRIBUTION OF LOGS OF MAXIMUM YEARLY 24, 48 AND 72 HOURS RAINFALL

One of the rules of approximate calculations is that all calculated values and tables used should have the relative errors of the same order as the relative errors of the results of observations. Application of this rule greatly reduces the number of calculations and tabulations required.

Consider as an example the values of yearly maximum 48 hours rainfall at Kiandra (N.S.W) for 77 years, the logarithms of which are assumed to be distributed normally. A full table of these values, in ascending order, and their logarithms with 3 figured mantissae (on many occasions even 5 figured logarithms are used) is given in the Appendix 2.

Formal statistical treatment, which takes considerable time even with calculating machines, leads to the following results:

Rainfall probability	Expected value	99% confidence limits
$\frac{1}{100}$	$E(P_{100})$ = 706 pts.	Lower = 611 pts. Upper = 817 pts.
Rainfall probability $\frac{1}{1000}$	$E(P_{1000})$ = 859 pts.	Lower = 716 pts. Upper = 1030 pts.

P_{100} = 706 or P_{1000} = 859 have an absolute error of $\pm\frac{1}{2}$ point, which is practically meaningless and unnecessary. The statistical reliability is 99 per cent that the maximum rainfall will be equal or exceed any value between 611 points and 817 points with probability $\frac{1}{100}$. It means that the value 706 is only one probable value in the "confidence" interval 611 to 817, or over 2 in., and accuracy up to $\frac{1}{2}$ point is pointless. For P_{1000} the confidence interval increases to 3 in. We can conclude that values P_{100} and P_{1000} and their confidence limits could have been calculated with accuracy $\pm\frac{1}{4}$ in. (= 25 points), or even $\frac{1}{2}$ in., and still would satisfy practical requirements.

The 77 values may be either

(i) distributed into equal class intervals 50 points each with mid-point values 250 pts., 300 pts... etc. or

(ii) each of these 77 values could be divided by 50 with rounding up results to the nearest whole number.

In the case (i) we express mid-point values 250, 300 .... in new units of 50 points each. In both cases we arrive at series of only 11 whole numbers 5, 6....15 with their frequencies (Table 1).

TABLE 1.

Class Interval	Central Value in points	Central Value in units $\frac{1}{2}$ in. x_i	Frequency f_i
225-274	250	5	4
275-324	300	6	14
325-374	350	7	20
375-424	400	8	16
425-474	450	9	7
475-524	500	10	3
525-574	550	11	5
575-624	600	12	3
625-674	650	13	1
674-724	700	14	1
725-774	750	15	2

We take now a new variate y_i = Log x_i using 2 figured log s (Appendix 1) and calculate the mean and the standard deviation of y_i by routine method*.

TABLE 2.

x_i	f_i	$y_i = \log x_i$	$t = y_i - \bar{y}^1$	$f_i t$	$f_i t^2$
5	4	0.70	-0.20	-0.80	0.16
6	14	0.78	-0.12	-1.68	0.25
7	20	0.84	-0.06	-0.12	0.072
8	16 $\bar{y}^1$ =	$\boxed{0.90}$	0.00	0.00	0.00
9	7	0.95	+0.05	+0.35	0.018
10	3	1.00	+0.10	+0.30	0.030
11	5	1.04	+0.14	+0.70	0.098
12	3	1.08	+0.18	+0.54	0.097
13	1	1.11	+0.21	+0.21	0.044
14	1	1.15	+0.25	+0.25	0.063
15	2	1.18	+0.28	+0.56	0.16
	N =77			$\Sigma f_i t$ = +0.31	0.992 $= \Sigma f_i t^2$

---

* It should be pointed out that this simplified and also full formal treatment are not rigorous mathematically. In both cases we are dealing with not equally distant logarithms of rainfall values. It is believed that errors arising from this are within limits of accuracy of data.

Assumed mean $\bar{y}^1 = 0.90$

True mean $\bar{y} = \bar{y}^1 + \frac{\Sigma f_i t}{N} = 0.90 + \frac{0.31}{77}$

$= 0.90 + 0.0040 \simeq 0.90$

Standard deviations, $S = \sqrt{\frac{\Sigma f_i t^2}{N} - \frac{(\Sigma f_i t)^2}{(\ N\ )}} = \sqrt{\frac{0.992}{77} \quad (0.0040)^2}$

$= \sqrt{0.0123} = 0.11$

Now $y_{100} = \log P_{100} = \bar{y} + 2.32S$

$= 0.90 + 0.255 = 1.155$ (**)

$P_{100} = 14.1$ in $\frac{1}{2}$ in. units

and $y_{1000} = \log P_{1000} = \bar{y} + 3.09\ S$

$= 0.90 + 0.34 = 1.24$

$P_{1000} = 17$ in $\frac{1}{2}$ in. units

or $P_{100} = 7.1$ in. and $P_{1000} = 8.5$ in.

Now, (Standard error of $\log P_{100})^2 = 3.71 \frac{S^2}{N}$

and (Standard error of $\log P_{100}) = \left(\frac{0.0123 .\ 3.71}{77}\right)^{\frac{1}{2}}$

$\simeq \underline{+}0.025$.

99% Confidence limits calculated with interpolation of 2 figured mantissae are,

$\text{Log } P_{100} = 1.155 \pm 0.025 \times 2.58 = 1.155 \pm 0.062 = \begin{matrix} 1.093 \\ 1.217 \end{matrix}$

$P_{100}\text{min.} = 12.3$ units, or approx. 6.2 inches

$P_{100}\text{max.} = 16.5$ units, or approx. 8.3 inches ..

with an accuracy of ± 25 points.

Also (Standard error of $\log P_{1000})^2 = \frac{S^2}{N} .\ 5.77 = 0.000917$

and (Standard error of $\log P_{1000}) = \pm 0.030$

$\text{Log } P_{1000} = 1.24 \pm 0.030 \times 2.58 = 1.24 \pm 0.08 = \begin{matrix} 1.16 \\ 1.32 \end{matrix}$

$P_{1000}\text{min.} = 14$ units $= 7$ in.

$P_{1000}\text{max.} = 21$ units $= 10.5$ in.

We compare now results obtained by formal method of treatment with results obtained by simplified method:

	P_{100}	P_{100} min.	P_{100} max.	P_{1000}	P_{1000} min.	P_{1000} max.
Formal	7.06 in.	6.11 in.	8.17 in.	8.59 in.	7.16 in.	10.30 in.
Simplified	7.1 in.	6.2 in.	8.3 in.	8.5 in.	7.0 in.	10.5 in.

It is seen that the accuracy of the simplified method is quite satisfactory.

** To ascertain the absolute accuracy of ±0.5 of the second significant figure it is advisable to make interim calculations on occasions up to third significant figure with simple interpolation in two figured logarithms, especially when this third figure is exactly 5. Final results should be again rounded to the nearer second significant figure.

4. CONCLUSIONS

(1) All maximum rainfall values could be represented by 2 or 1 figured numbers by introducing an appropriate unit (10 points, 20 points, 50 points etc.) depending on the lowest and the highest values of maximum in the sample.

(2) In the case of log normal distribution of rainfall, logarithms with 2 figured mantissae can be used with simple interpolation in the calculations of P_{100}, P_{1000} values and their confidence limits, without reducing the real accuracy of assessments.

REFERENCE

Hann-Suering	1926	Lehrbuch der Meteorologie, Leipzig 1926.

APPENDIX 1.

Table of Logarithms, 2 figured mantissae

	0	1	2	3	4	5	6	7	8	9
1	.00	04	08	11	15	18	20	23	25	27
2	.30	32	34	36	38	40	41	43	45	46
3	.48	49	50	52	53	54	56	57	58	59
4	.60	61	62	63	64	65	66	67	68	69
5	.70	71	72	72	73	74	74	76	76	77
6	.78	78	79	80	81	81	82	83	83	84
7	.84	85	86	86	87	88	88	89	89	90
8	.90	91	91	92	92	93	93	94	94	95
9	.95	96	96	97	97	98	98	99	99	.00

APPENDIX 2.

48 hour rainfalls maxima - Kiandra, New South Wales

No.	x_i	$y_i = \log_{10} x_i$	y_i^2
1	230	2.362	5.579
2	255	2.407	5.794
3	257	2.410	5.808
4	263	2.420	5.856
5	280	2.447	5.988
6	283	2.452	5.012
7	285	2.455	6.027
8	285	2.455	6.027
9	289	2.461	6.057
10	294	2.468	6.091
11	300	2.477	6.136
12	301	2.479	6.145
13	303	2.481	6.155
14	307	2.487	6.185
15	311	2.493	6.215
16	316	2.500	6.250
17	318	2.502	6.260
18	320	2.505	6.275
19	327	2.515	6.325
20	328	2.516	6.330
21	331	2.520	6.350
22	335	2.525	6.376
23	335	2.525	6.376
24	340	2.531	6.406
25	342	2.534	6.421
26	343	2.535	6.426
27	343	2.535	6.426
28	346	2.539	6.447
29	348	2.542	6.462
30	350	2.544	6.472
31	351	2.545	6.477
32	356	2.551	6.508
33	365	2.562	6.564
34	365	2.562	6.564
35	366	2.563	6.569
36	366	2.563	6.569
37	372	2.571	6.610
38	374	2.573	6.620
39	377	2.576	6.636
40	379	2.579	6.651
41	384	2.584	6.677
42	385	2.585	6.682
43	386	2.587	6.693
44	394	2.596	6.739
45	397	2.599	6.755
46	400	2.602	6.770
47	400	2.602	6.770
48	403	2.605	6.786
49	403	2.605	6.786
50	412	2.615	6.838
51	413	2.616	6.843
52	414	2.617	6.849
53	417	2.620	6.864
54	423	2.626	6.896

APPENDIX 2 (contd)

No.	x_i	$y_i = \log_{10}x_i$	y_i^2
55	431	2.634	6.938
56	437	2.640	6.970
57	450	2.653	7.038
58	452	2.655	7.049
59	455	2.658	7.065
60	461	2.664	7.097
61	466	2.668	7.118
62	468	2.670	7.129
63	487	2.688	7.225
64	490	2.690	7.236
65	510	2.708	7.333
66	527	2.722	7.409
67	529	2.723	7.415
68	530	2.724	7.420
69	537	2.730	7.453
70	540	2.732	7.464
71	578	2.762	7.629
72	600	2.778	7.717
73	603	2.780	7.728
74	656	2.817	7.935
75	722	2.859	8.174
76	745	2.872	8.248
77	773	2.888	8.341