Transactions, American Geophysical Union Volume 35, Number 4 August 1954

A METHOD FOR DETERMINING THE MINIMUM DURATION OF WATERSHED EXPERIMENTS

Jacob L. Kovner and Thomas C. Evans

Abstract--A simple graphic solution is described for approximating the length of time required to detect significant differences between treatments on experimental watersheds.

Introduction--An increasing awareness of the importance of water to local and national welfare has prompted numerous small watershed studies of the relationship between streamflow and forest cover. The experimental methods in use are quite varied, but in general resemble the familiar pattern of comparing some element of streamflow on treated watersheds with that of untreated or control watersheds.

One very simple and popular design, used extensively at the Coweeta Hydrologic Laboratory in North Carolina, employs two watersheds with similar streamflow characteristics. Ordinarily this also means similar elevations, aspects, climates, and soil complexes. Prior to treatment, measurements of streamflow (such as yield, storm peaks, sedimentation, and quality) are collected on both watersheds for a number of years known as the period of calibration. These data are used to establish the regression of a streamflow characteristic of one watershed upon the other. By this method, variation not associated with treatment (for example, differences such as temperature and rainfall for different years) is placed under statistical control in the subsequent test of difference in streamflow characteristics between the calibration and treatment periods. Following calibration, one of the watersheds is treated according to the specifications of some land-use or forest management practice, and measurements of streamflow continued until the investigator is satisfied that an adequate test of treatment effects has been made.

Analysis of the two sets of data, one set for each period, is by the method of covariance [SNEDECOR, 1946], which supplies a significance test of the adjusted difference due to treatment. WILM [1944, 1949] has demonstrated the method, using an example of annual discharge or yield from a control and treated watershed. In addition, by making the logical and simple assumption that equal samples (for example, streamflow records of equal duration) are drawn from each of the two periods, he has arrived at a method of successive approximations for estimating how long the calibration period should be.

A more general solution can be derived for the length of both the calibration and treatment periods. Under the conditions of the test it need not be assumed that the periods are of equal length. On some watershed studies, calibration data have been accumulated for relatively long periods. Here, of course, treatment may be applied at once, and it is of interest to the investigator to anticipate how long the post-treatment observation should be continued. It is also frequently necessary to estimate appropriate time intervals for spacing successive treatments, as in the case of removing a forest cover in a series of partial cuts. Furthermore, when the effect of the treatment is short-lived, longer periods of observation tend to mask the effect. Again, when funds, personnel, and the physical resource in land are limited, estimates of the length of post-treatment period are powerful aids to long-time planning and assignment of project priorities.

It is the purpose of this paper to propose a method of estimating the minimum length of the treatment period, and of a whole watershed experiment, and to arrive at a simplication of WILM's [1949] technique by substituting a graphic solution for the more laborious method of successive approximations.

Method of analysis--With a slight modification to admit periods of unequal length, the equation for the square of a mean difference after WILM [1949] may be written

$$d^2 = s_{y.x}^2 \, F \, [1/n_1 + 1/n_2 + (\overline{X}_1 - \overline{X}_2)^2/S_x^2] \quad (1)$$

where S_x^2 is the within treatment pooled sum of squares of the independent variable, n_1 and n_2 are the number of observations in the calibration and treatment periods, respectively, $s_{y \cdot x}^2$ is the experimental variance developed from residual deviations from regression, and F is the variance ratio.

Some simplification of (1) is possible by substituting the approximation

$$(\bar{X}_1 - \bar{X}_2)^2 \simeq F\ sx^2\ (1/n_1 + 1/n_2)$$

and remembering that

$$sx^2 = Sx^2/n_1 + n_2 - 2$$

By appropriate substitution and rearrangement (1) resolves to

$$\frac{s_{y \cdot x}^2}{d^2} = \frac{n_1 n_2}{n_1 + n_2} \left[\frac{1}{F\,(1 + \frac{F}{n_1 + n_2 - 2})} \right] \qquad (2)$$

wherein F has degrees of freedom equal to unity and $n_1 + n_2 - 3$.

Because it leads to better graphical presentation, the form in (2) is preferred to the more common $d/s_{y \cdot x}$.

With n_1 and an estimate of $s_{y \cdot x}^2$ both taken from the calibration data, and an experience value for the smallest worthwhile difference, d, (2) might be solved for n_2 by the method of successive approximations. A simpler solution is possible by graphics. Figure 1 shows the relation (2) as a family of curves for values of n_2 with F at a probability level of 0.05. Sample calculations for deriving plotting points are shown in Table 1, which contains the solution for $n_2 = 5$. Similar computations at P = 0.01 would complete the usual working range of significance.

Table 1--Sample computations for $s_{y \cdot x}^2/d^2$ for $n_2 = 5$, P = 0.05

n_1	$n_1 + n_2 - 2$	$n_1 + n_2 - 3$ (D. F.)	F	$F\,(1 + \frac{F}{n_1 + n_2 - 2})$	$\frac{n_1 n_2}{n_1 + n_2}$	$\frac{s_{y \cdot x}^2}{d^2}$
3	6	5	6.61	13.8920	1.8750	0.1350
4	7	6	5.99	11.1156	2.2222	0.1999
5	8	7	5.59	9.4963	2.5000	0.2633
6	9	8	5.32	8.4647	2.7273	0.3222
7	10	9	5.12	7.7414	2.9167	0.3768
8	11	10	4.96	7.1965	3.0769	0.4276
9	12	11	4.84	6.7920	3.2143	0.4733
10	13	12	4.75	6.4857	3.3333	0.5140
11	14	13	4.67	6.2279	3.4375	0.5520
12	15	14	4.60	6.0108	3.5294	0.5872
13	16	15	4.54	5.8285	3.6111	0.6196
14	17	16	4.49	5.6758	3.6842	0.6492
15	18	17	4.45	5.5500	3.7500	0.6756
16	19	18	4.41	5.4336	3.8095	0.7011
17	20	19	4.38	5.3392	3.8636	0.7237
18	21	20	4.35	5.2509	3.9130	0.7451
19	22	21	4.32	5.1684	3.9583	0.7660
20	23	22	4.30	5.1041	4.0000	0.7837

With these curves, the estimated minimum length of the post-treatment period is readily solved if n_1, $s_{y \cdot x}^2$, and d are known, for it is necessary only to compute $s_{y \cdot x}^2/d^2$, and enter the curves at the coordinate value determined by n_1 and $s_{y \cdot x}^2/d^2$. The curve nearest the coordinate point is the required estimate of n_2.

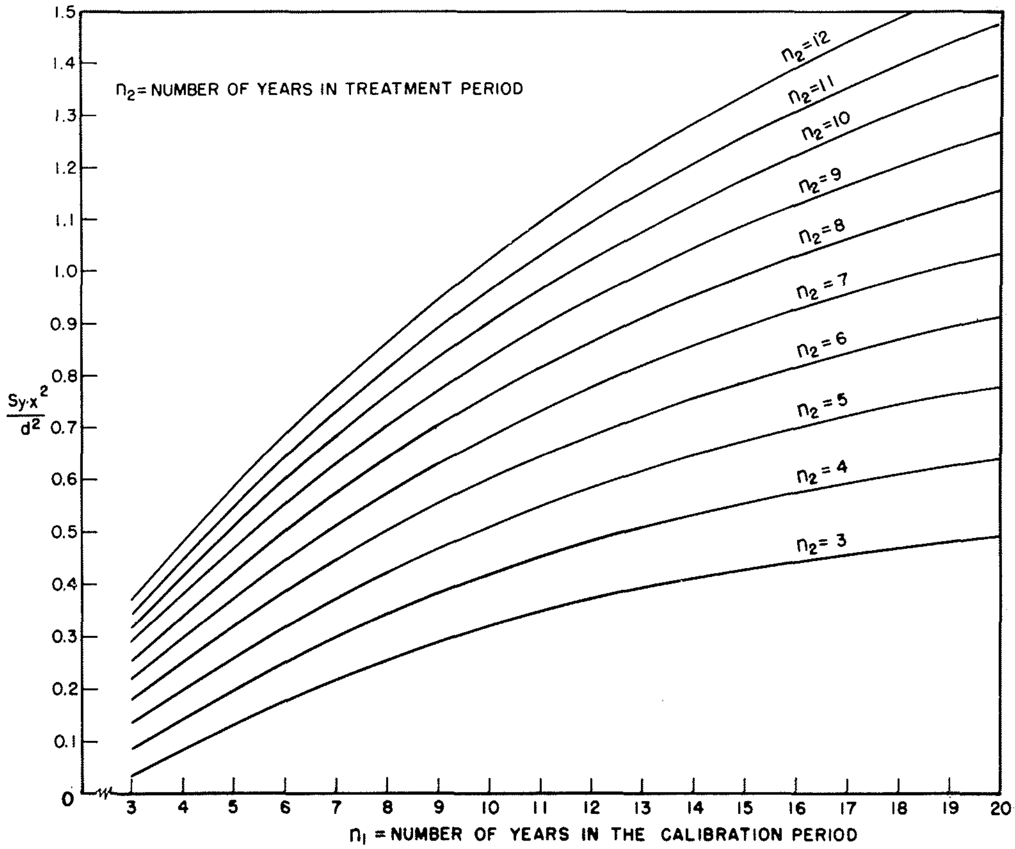

Fig. 1--Graphic solution of the equation

$$\frac{s_{y \cdot x}^2}{d^2} = \frac{n_1 n_2}{n_1 + n_2} \left[\frac{1}{F\,(1 + \frac{F}{n_1 + n_2 - 2})} \right]$$

for the purpose of estimating length of calibration and treatment periods

Of course, of the three quantities, only n_1 is known precisely before the treatment is installed. The variance $s_{y \cdot x}^2$ is an estimate of the pooled residual variances around regression to be obtained after the post-treatment period, and d is generally predetermined by practical considerations based on experience [WILM, 1949].

Examples of the method--The data in Table 2, showing annual water yields, were obtained from records at the Coweeta Hydrologic Laboratory. Data are given for both a calibration and treatment period so as to permit an estimate of n_2 and an empirical check of the estimate.

The error variance for the calibration period $s_{y \cdot x}^2$ is 2.389; n_1 is 12. An experience value for the smallest worthwhile difference, d, may be taken as five per cent of the mean yield for the calibration period of the watershed subsequently chosen for treatment. Then $0.05 \times 28.916 = 1.446$ area-inches, which provides

$$s_{y \cdot x}^2/d^2 = 2.389/2.090 = 1.143$$

Entering Figure 1 with $n_1 = 12$ and $s_{y \cdot x}^2/d^2 = 1.143$, the curve nearest this point is $n_2 = 12$. It is estimated, therefore, that the treatment would have to run 12 years in order that a difference of 1.446 area-inches may be judged significant at a probability level of 0.05.

Table 2--Annual discharge from Watersheds A and B (discharge in area-inches)

Calibration period 1934-1945	Watershed A	Watershed B	Treatment period 1946-1952	Watershed A	Watershed B
1934	33.89	35.53	1946	12.17	16.01
1935	13.16	15.38	1947	15.03	20.83
1936	12.20	12.96	1948	25.50	28.07
1937	22.57	23.58	1949	46.76	45.88
1938	32.87	33.82	1950	38.33	39.00
1939	27.89	31.64	1951	30.48	31.58
1940	16.73	17.75	1952	51.87	52.43
1941	35.36	34.99			
1942	20.82	23.74			
1943	30.71	31.35			
1944	53.38	49.13			
1945	38.79	37.12			
Total	338.37	346.99		220.14	233.80
Average	28.198	28.916		31.449	33.400

Similar computations at other values of d (Table 3) show that if the experiment must be concluded within, say five years, then a treatment effect less than 2.00 area-inches cannot be expected to appear significant. If, on the other hand, the smallest worthwhile difference could be increased to ten per cent, a three-year test is likely to detect a difference of approximately 3.0 area-inches.

Table 3--Solution for length of treatment period

Precision as per cent of mean discharge on Watershed B	Least significant difference, in area-inches, d	d^2	$s_{y \cdot x}^2/d^2$	n_1 years	n_2 years
5	1.446	2.091	1.143	12	12
6	1.735	3.010	0.794	12	7
8	2.313	5.350	0.447	12	4
10	2.892	8.364	0.286	12	3

By way of verification, an analysis of covariance was run on the treatment data on Table 2, testing for significance of the adjusted mean yields. The analysis was applied first at the end of the three-year treatment period 1946-1948, and then successively for the 1946-1949, 1946-1950, 1946-1951, and 1946-1952 treatment periods, as would be done in actual practice. The analysis shows that it was necessary to run the treatment for seven years, 1946-1952, in order to show a real difference of 1.66 area-inches due to treatment. This is 5.7 pct of the calibration mean yield on Watershed B. Referring to Table 3, we find that the estimate of $n_2 = 7$ years for a least significant difference equal to six per cent of this mean is in agreement with actual results.

The curves can also be used to determine how long the calibration period should be when we have an estimate of $s_{y.x}^2$, however obtained. This is an alternate method to that proposed by WILM [1949] using successive approximations. In this case it is only necessary to take equal samples, that is $n_1 = n_2$. Thus $s_{y.x}^2/d^2$ is calculated as previously. Then we enter the vertical scale at this value and move horizontally along this ordinate until that curve is intersected such that $n_1 = n_2$.

As a numerical example, the third line of Table 3 shows d = 2.313 area-inches and $s_{y.x}^2/d^2 = 0.447$; we find that along the 0.447 ordinate the condition that $n_1 = n_2$ is satisfied for the value 6.5 years. Note that we can interpolate between the parametric curves. Therefore, seven years is probably a good estimate of how long the calibration period should be.

It is interesting to see how the results obtained by this method compare with those given by WILM [1949, p. 275]. In his example we find that $s_{y.x}^2/d^2 = 0.266$. Reference to our Figure 1 shows directly that $n_1 = n_2 = 5$, that is, the calibration period is estimated at five years. This agrees precisely with the solution by Wilm. It will be found that the other examples by Wilm can be solved by the use of Figure 1.

Finally, the researcher might want to specify the number of samples in the two periods and have some information regarding the precision or the least significant difference which could be detected. This can be done by entering the curves with values of n_1 and n_2 and reading the corresponding value of $s_{y.x}^2/d^2$; if an estimate of the numerator, $s_{y.x}^2$, is available, d^2 and d are readily calculated.

For example, let us take $n_1 = 12$ and $n_2 = 3$. Then from Figure 1 the corresponding ordinate value $s_{y.x}^2/d^2$ is 0.37. Using $s_{y.x}^2 = 2.39$ from Table 2, $d^2 = 2.39/0.37 = 6.46$, or $\underline{d} = 2.54$ area-inches.

Summary--The investigator is frequently presented with the practical problem of determining the number of observations required to show a significant difference between two means if the difference is greater than a fixed amount and some estimate of the variance is known. In hydrologic research on streamflow with experimental watersheds, a similar problem arises. A simple graphic method has been presented which provides the following information: (1) Length of the treatment period following a known extended calibration period; (2) length of the calibration period in advance; and (3) the smallest difference due to treatment which can be detected as significant at a chosen probability level, when the length of two periods has been assumed. This method may be applied in other problems where, in effect, we are comparing two or more regression lines.

References

SNEDECOR, G. W., Statistical methods, 4th edition, Iowa State College Press, 485 pp., 1946.

WILM, H. G., Statistical control of hydrologic data from experimental watersheds, Trans. Amer. Geophys. Union, pt. 2, pp. 618-622, 1944.

WILM, H. G., How long should experimental watersheds be calibrated? Trans. Amer. Geophys. Union, v. 30, pp. 272-278, 1949.

Southeastern Forest Experiment Station,
Federal Building,
Asheville, North Carolina

(Communicated manuscript received August 20, 1953; open for formal discussion until January 1, 1955.)