Calibration of Five Small Forested Watersheds

K. G. REINHART

Abstract—On the Fernow Experimental Forest in West Virginia, five forested watersheds, 38 to 96 acres in area, have been gaged by the U. S. Forest Service since 1951. The watersheds are close together, for the most part contiguous, and are similar in topography, soil, forest cover, and stream-flow characteristics. The methods and problems of calibration should interest other investigators in this field because seldom has a group of watersheds been calibrated for forest watershed research to determine the effects of a range of cutting treatments. After calibration, four watersheds will be cut over, using four different cutting practices. One watershed will be used as a control. Effects of treatment on stream flow will then be measured. Methods of Wilm and of Kovner and Evans were used to determine length of calibration period based on an analysis of annual, monthly, peak, and low flows. Problems of treatment assignment are discussed.

Introduction—The effect of different types of forest management on stream flow is being investigated by the Mountain State Research Center of the Northeastern Forest Experiment Station, U. S. Forest Service, on the Fernow Experimental Forest in Tucker County, near Parsons, W. Va.

Five forested watersheds, ranging from 38 to 96 acres in area, have been continuously gaged since May 1951 (Fig. 1; 80 chains = one mile). V-notch weirs (120°) and continuous water-stage recorders are used. Water quality, precipitation, and soil-moisture content are also being measured.

The watersheds, supporting a 50-year old stand of hardwood timber, have been undisturbed for several decades. Slopes are steep; soils are stony and for the most part shallow.

For calibration of these watersheds, a relationship was established between stream-flow measurements for a period of years in each watershed to be treated and concurrent measurements in another watershed, the control. An analysis was made testing the predictive value of the relationship, to determine (1) the standard error of estimate of the prediction, and (2) the number of observations after treatment required to establish the significance of a difference considered to be of practical importance.

Analytical methods for determining the adequacy of calibration of annual discharge from small watersheds have been presented by *Wilm* [1949] and by *Kovner* and *Evans* [1954]. In this paper application of these methods to peak flows and low flows as well as annual discharge is discussed.

After calibration, four of the Fernow watersheds will be cut over, each by a different system. These systems range from a commercial clearcut with skid roads at maximum operable grades to a conservative selection system (intensive forestry) with skid-road grades not exceeding ten per cent and with provisions for road drainage. One watershed will be left in its undisturbed state as a control. After treatment, stream-flow measurements will be continued to determine the effect on water quantity, distribution, and quality.

During the calibration period, stream-flow characteristics of one watershed are related to another, leading to expressions for predicting values of one from values of the other. We are using straight-line regressions after determining that they fit the data well. For example, based on five years of record, the equation $Y = 0.35 + 0.91X$ enables us with a certain degree of accuracy to predict the annual flow in area-inches of Watershed 1 (Y) from that of Watershed 3 (X).

Length of calibration period—One problem faced, as in all such experiments, was to determine the proper length of the calibration period based on acceptable error. Too short a calibration period would lead to inconclusive results. Too long a calibration period results in unnecessary expense. Our problem is complicated by the fact that the period for measurement of treatment effects will be of limited duration: after logging, the watersheds will recover and effects will diminish. We are thinking in terms of a period of two or three years during which effects of treatment will be at or near a maximum.

Kovner and Evans prepared a graphical solution of Wilm's equation for determing the minimum length of the calibration period. Modified to include a larger number of observations, their graph is reproduced in Figure 2. The horizontal axis shows the number of observations in the calibration period. The vertical axis is the squared ratio:

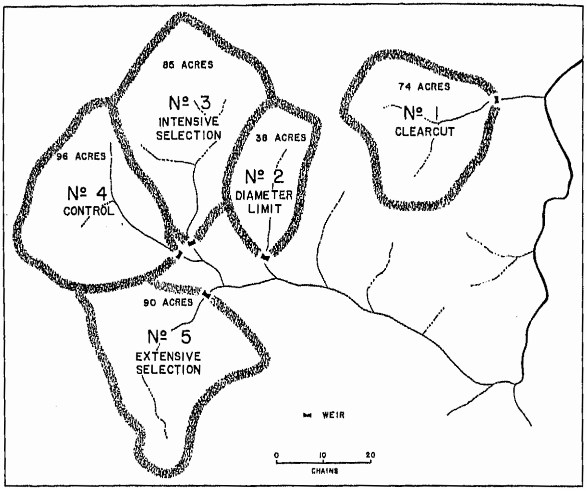

FIG. 1 – Locations of the five gaged watersheds on the Fernow Experimental Forest, showing proposed treatments

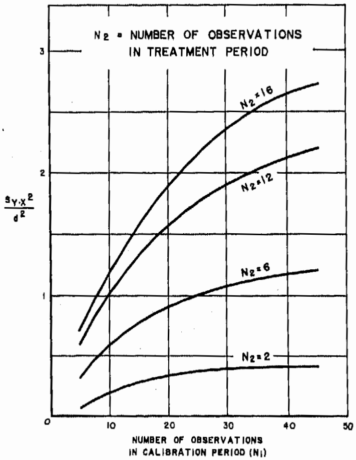

FIG. 2 – Graph for determining duration of calibration and treatment periods, revised from *Kovner* and *Evans* [1954]

standard error of estimate divided by the smallest worthwhile difference. The graph shows a family of curves. Given the number of observations in the calibration period and a value for $s_{y \cdot x}^2/d^2$, the number of observations needed in the treatment period can be read or interpolated.

To use this method, certain assumptions must be made. First, a value must be assigned for d, the smallest worthwhile difference. This might be the smallest difference expected as a result of treatment, based on experience, or a difference due to treatment of such a magnitude that if it is not reached the investigator is willing to forego determination of its significance. Another assumption made, and open to some question, is that variances after treatment will be of the same magnitude as those before treatment.

Analysis—Five years of data were analyzed as a basis for deciding the length of the calibration period and assignment of treatments to individual watersheds. Annual and monthly water flows, peak flows, and low flows were analyzed. Water quality was not analyzed because from previous studies at the Fernow Experimental Forest and elsewhere there seems little doubt that changes in quality, such as turbidity, that result from treatment will be significant.

Mean annual flow ranged from 20 to 27 area-inches for the five watersheds. Regression equations were computed to predict annual flow of each of the watersheds from each of the others; best predictions were obtained with Watershed 3 as the control.

TABLE 1 – *Calibration of four watersheds on Watershed 3; correlation coefficients and errors of estimate*

Item	Watershed							
	1		2		4		5	
	Cor. coef.	Error of est.	Cor. coef.	Error of est.	Cor. coef.	Error of est.	Cor. coef.	Error of est.
Annual flow (error of estimate in area inches)	0.99993	0.055	0.995	0.592	0.996	0.493	0.9988	0.326
Peak flow (error of estimate in csm[a])	0.984	3.72	0.990	3.67	0.990	2.42	0.975	4.86
Low flow (error of estimate in number of days flow less than 0.05 csm[a])	0.998	3.92	0.995	5.35	0.997	4.32	0.995	5.03

[a] Cubic feet per second per square mile.

For the four regressions predicting flow from Watershed 3, the errors of estimate ranged from 0.05 to 0.59 area-inch of water (Table 1). Correlation coefficients ranged from 0.995 to 0.99993. Using the procedure described, with five years in the calibration period and Watershed 3 as the control, one or two years would be required in the treatment period of each of the other watersheds to establish significance at the five per-cent level of a ten per-cent change in flow of the treated watershed (Table 2). We will be willing to forego determination of its statistical significance if change in flow is less than ten per cent.

With five years of record, there are of course only five observations of annual flow for each watershed, a small number for regression analysis. Therefore, consideration was given to analyzing monthly flows, giving 60 observations in five years. Analysis similar to that already given indicated that, with Watershed 3 as the control, only two to six months would be required to establish significance of a ten per-cent change at the five per-cent level. However, there are objections, often referred to as 'serial correlation', to use of monthly data in this way. For this reason reliance was not placed upon analysis of data by months.

Basic data used in the peak-flow analysis were the peaks at each watershed on 45 occasions when flow of Watershed 3 exceeded ten cu ft/sec mi². Peak flows ranged up to 180 cu ft/sec mi² and averaged between 21 and 29 cu ft/sec mi² for the five watersheds. Regressions were computed to estimate peaks of each watershed from peaks of each of the others. With Watershed 3 as the control, analysis indicates that seven to 16 peaks would be required in the treatment period to establish the significance (at five per-cent level) of a ten per-cent change due to treatment. In the calibration period,

TABLE 2 – *Calibration of four watersheds on Watershed 3; years required in treatment period with calibration period of five years*

Item	Smallest worthwhile difference	Watershed			
		1	2	4	5
	pct	yr	yr	yr	yr
Annual flow	10	1	2	2	1
Peak flow[a]	10	2 (14)	2 (11)	1 (7)	2 (16)
Low flow	20	2	3	2	4

[a] Years required based on average occurrence of peaks in calibration period. Required number of observations (peaks) given in parentheses.

TABLE 3 – *Calibration of four watersheds on Watershed 4; years required in treatment period with calibration period of five years*

Item	Smallest worthwhile difference	Watershed			
		1	2	3	5
	pct	yr	yr	yr	yr
Annual flow	10	2	3	2	1
Peak flow	10	3 (25)	2 (19)	1 (8)	3 (24)
Low flow	20	1	3	3	3

there was an annual average of nine peaks of the size included in the analysis. Thus the required treatment period might be about two years.

More difficulties were encountered in the analysis for low flows than in the others. First, it was difficult to decide upon a good measure of low flows. After trial runs using several different approaches, analysis was based upon number of days in the year when flow was 0.05 cu ft/sec mi² or lower. For the five years on five watersheds, number of such days varied from 3 to 166. Analysis indicates that, with Watershed 3 as the control, two to four years in the treatment period will be required to establish the significance at the five per-cent level of a 20-pct change in number of days. We are still trying to find a better measure for low flows.

Decisions made—Based on the analysis, we decided to wait until one additional year of record, making six in all, has been obtained before starting treatment. The measurement of low flows gives the most concern. However, results of experiments elsewhere indicate the probability that the effect of treatment on low flows may be greater proportionally than for annual flows.

For several reasons, treatments assigned to individual watersheds were not at random. This made our statistician advisors uneasy, but they finally agreed. Watershed 3 was originally chosen as the control because generally it gave the highest correlations; Watershed 4, well correlated with Watershed 3, was designated for the intensive forestry practice, the least drastic treatment.

However, a proper road for the intensive practice could not practicably be constructed in Watershed 4 without crossing into Watershed 3. Therefore, Watershed 4 was made the control, and Watershed 3 was assigned the intensive forestry practice since a road could be built into Watershed 3 without infringing upon adjacent watersheds. (Table 3 shows years required in treatment period with Watershed 4 as the control). Watershed 1 was chosen for clear-cutting, a destructive practice, for the compelling though unscientific reason that of all five watersheds it is the only one situated below the reservoir of the nearby town of Parsons. Watershed 1 was well correlated with Watershed 4. Thus, the two extremes of treatment were assigned to the watersheds best correlated with the control. The intermediate treatments, a diameter-limit cut and extensive forestry, were assigned to Watersheds 2 and 5.

Conclusions—Five years of record showed remarkably good correlation in stream-flow characteristics of the five gaged watersheds. The methods of Wilm, and Kovner and Evans proved suitable for the analysis of data in this experiment and led to an apparently sound evaluation of necessary length of calibration period. A final evaluation of the methods used cannot be made, of course, until after the treatment period.

References

Kovner, J. L., and T. C. Evans, A method for determining the minimum duration of watershed experiments, *Trans. Amer. Geophys. Union*, **35**, 608–612, 1954.

Wilm, H. G., How long should experimental watersheds be calibrated? *Trans. Amer. Geophys. Union*, **30**, 272–278, 1949.

Northeastern Forest Experiment Station, U. S. Forest Service, 102 Motors Avenue, Upper Darby, Pennsylvania.

(Communicated manuscript received January 20, 1958; open for formal discussion until March 1, 1959.)