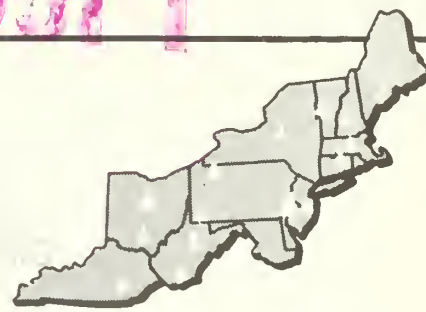
USDA FOREST SERVICE RESEARCH NOTE NE-175
1973

Northeastern Forest Experiment Station



FOREST SERVICE, U.S. DEPT. OF AGRICULTURE, 6816 MARKET STREET, UPPER DARBY, PA. 19082

THE PROBLEM OF EXTREME EVENTS IN PAIRED-WATERSHED STUDIES

Abstract.—In paired-watershed studies, the occurrence of an extreme event during the after-treatment period presents a problem: the effects of treatment must be determined by using greatly extrapolated regression statistics. Several steps are presented to help insure careful handling of extreme events during analysis and reporting of research results.

Much of our present knowledge of forest hydrology is the result of experiments on paired gaged watersheds. The statistical method most commonly used in these studies is relatively straightforward and was first proposed by Wilm (*1949*). Data from a calibration period are used to develop regression equations between streamflow from a control watershed (additional variables may be included if necessary) and streamflow from a nearby similar watershed on which a forest treatment will be performed. After treatment, deviations of the treated watershed values from the calibration regression are considered to represent treatment effects if the deviations fall outside specified confidence intervals placed about the regression line. Illustrations of this technique can be found in papers by Hibbert (*1969*), Reinhart *et al.* (*1963*), and Goodell (*1958*).

A problem in paired-watershed studies is the occurrence of an extreme event—a streamflow total or other variable that falls well away from the expected span of measured values. Causes of extreme events are overabundant precipitation, unusually rapid snow-water contribution (frequently combined with rainfall), or prolonged drought. Because they are rare, extreme events are often of particular interest in studying the effects of forests and forest treatments on streamflow. However, the current statistical methods used in paired-watershed studies are not well adapted to handling the extreme event, particularly if it occurs in the after-treatment period. Calibration statistics must be extrapolated to accommodate the extreme event, thus providing a potential source of error in determining treatment effects on streamflow.

The best protection against the extreme event is an all-encompassing calibration. Such a calibration is usually obtained by a lengthy period of measurement (*Wilm 1949*). Longer calibrations, in addition to providing more precise regressions and confidence intervals, increase the chances of including a full range of weather events for a particular time period (both wet Augusts and dry Augusts, for example). However, costs and the desire to obtain quick research results frequently dictate that calibration be ended after the shortest practicable period.

In the East, treatment is usually performed after a calibration of 5 to 8 years. The exact length of calibration is generally arrived at with the aid of a graphical solution developed by Kovner and Evans (*1954*), which takes into account the variation and range of the measured data. In addition, long-term precipitation records are usually studied to further insure that the calibration period has encompassed a nearly full range of weather events. Despite these precautions, experience has shown that some streamflow values from the control watershed during the after-treatment period can be expected to fall outside the range encountered during calibration.

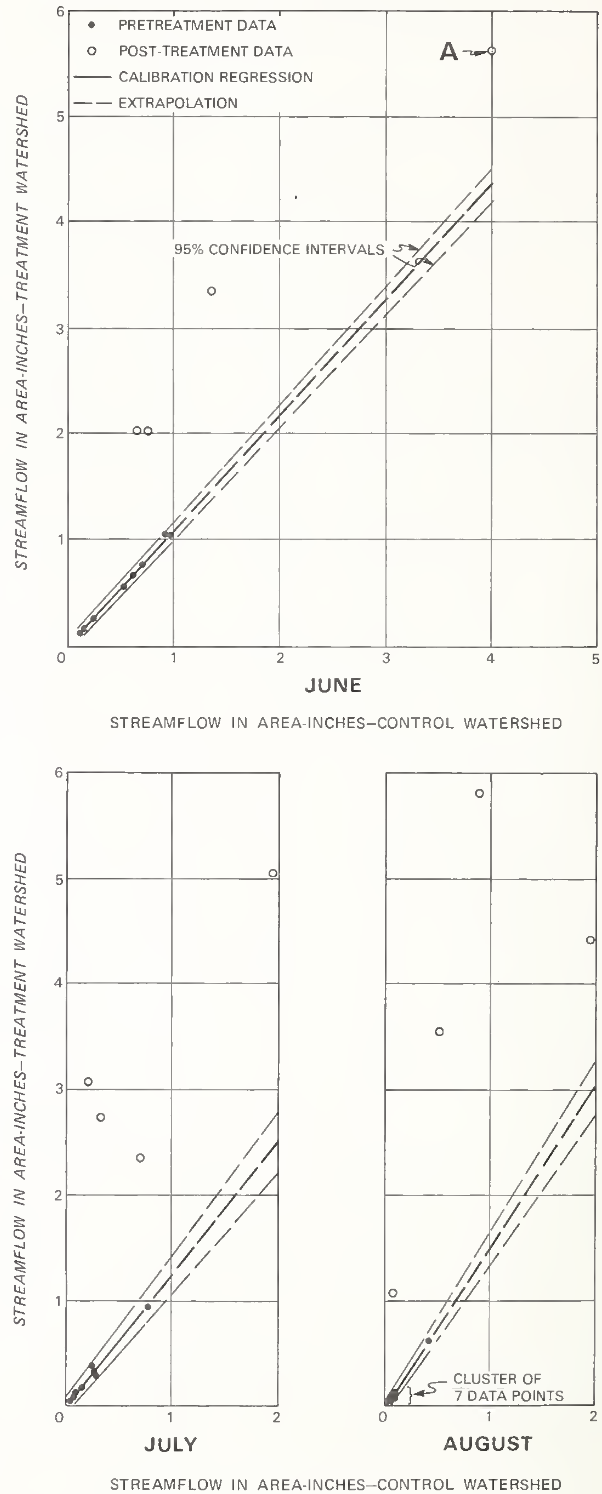
A case in point is a study now in progress at the Hubbard Brook Experimental Forest in central New Hampshire. After an 8-year calibration, a 39-acre watershed was cleared of its forest cover and sprayed with herbicides for three successive summers in an attempt to obtain a measure of maximum increases in water yield (*Hornbeck et al. 1970*). During the first 44 months after treatment, 12 monthly streamflow values for the control watershed fell outside the ranges encountered during calibration. At least three of the monthly flows were large enough to be classed as extreme events.

The extreme event problem is not unique to Hubbard Brook. It has also been experienced at other hydrologic research stations in the East, including the Fernow Experimental Forest in West Virginia and the Coweeta Hydrologic Laboratory in North Carolina. The extreme event problem may also be familiar to researchers in other disciplines, who use regression to determine effects of a treatment.

Problem Illustration

When the after-treatment values for the control watershed fall outside the range of flows encountered during calibration, determination of treatment effects must be based on extrapolated regression statistics (fig. 1). In the case of an extreme event, the extrapo-

Figure 1. — Streamflow and regression data for Hubbard Brook paired-watershed experiment. Streamflow of the control watershed was the only independent variable used in calibrating this watershed.



lation may have to be several times the range covered during calibration. Because the researcher cannot be certain that the regression model is valid over a wider range than the calibration sample, the extrapolation procedure is risky. Watershed differences in soil depth, area, precipitation distribution, or vegetation could conceivably cause pronounced change in slope of the regression line, beginning at some point beyond the range sampled during calibration.

Streamflow totals from the Hubbard Brook paired watersheds for June, July, and August illustrate the problem that arises when after-treatment values fall outside the range of flows measured during the calibration. For each of these 3 months, at least one after-treatment streamflow value from the control was more than double the maximum value measured during calibration (fig. 1). The highest June after-treatment value (point A in fig. 1) clearly falls into the extreme event class.

Intuitively, watershed researchers using extrapolated regression statistics probably would not be concerned about determining treatment effects on the July and August points. The highest control watershed values for these months exceed the maximum calibration values by only about 1 inch, which is not likely to greatly change the slope of the regression line. But less confidence could be placed in a test of the highest June after-treatment value (point A in fig. 1), which exceeds the maximum control watershed calibration value by more than 3 inches.

The real peril of the extreme event problem lies in the reporting of gaged watershed results. The effect of a particular forest treatment is usually given quantitatively as the difference between the regression line and the actual measured streamflow. For example, point A in figure 1 would be reported as a 1.2-inch increase in streamflow for June 1968. If, in reality, the regression line were not linear, but changed slope and became less steep at around 2.0 inches, the effect of treatment would actually be an increase greater than 1.2 inches. On the other hand, if the regression line swung more sharply upward, there may not have been any significant change in streamflow for the month. The area in which it becomes unsafe to use an extrapolated regression depends on many factors and must be determined by a researcher who is thoroughly familiar with the watersheds and the statistics being studied.

Analyzing the Extreme Event

When an extreme event occurs, the first step should be to assure the validity of the measurement. Simple comparisons among a group of gaged watersheds will usually point out gross errors. Comparable precipitation for the watersheds can also be checked to determine that amounts are radically higher or lower than those obtained in the calibration period.

If the extreme event proves valid, the next step should be to check the calibration regression for any signs of non-linearity. In most cases, the calibration data will fit well, particularly because paired watersheds are selected so that they will be as similar as possible. But if the calibration fit is poor or the standard error is large, it may be necessary to exclude analysis of the extreme event or at least be fully aware that any conclusions based on extrapolation will be weak.

When the calibration fit is satisfactory, a final step is to look for reasons why the watersheds might react differently for an extreme event than they did for the range of calibration measurements. Where such reasons are not readily apparent, some idea of reaction for extreme events may be obtained if a spare untreated watershed is available. Calibration statistics can be determined and post-treatment streamflow values for the spare untreated watershed can be estimated in the same manner as for the treated watershed. In the case of an extreme event, if the estimated flow is close to actual streamflow for the spare untreated watershed, more confidence can be placed in determinations of streamflow changes on the treated watershed.

This procedure was tried for the June streamflow values presented earlier (fig. 1). For the extreme event, a predicted streamflow of 3.9 inches was obtained for a nearby untreated watershed. Actual flow from the watershed was 4.1 inches. The difference in the two values is not statistically significant, illustrating that the prediction equation holds up well for the extreme event. Thus more confidence can be placed in the extrapolated estimate of flow for the treated watershed.

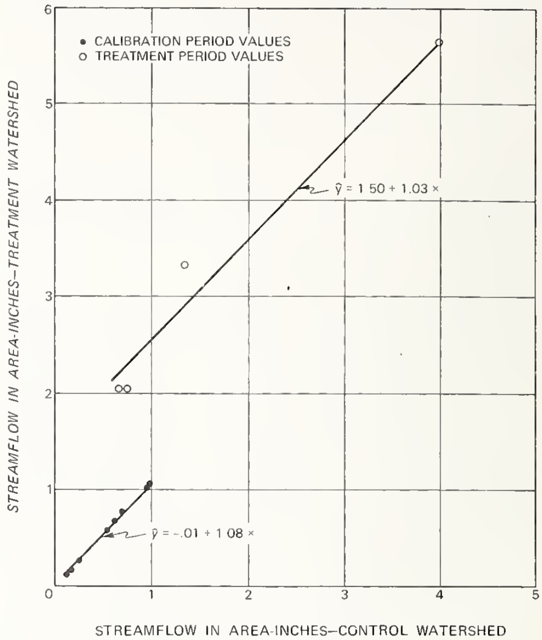
As an alternate approach to analyzing the extreme event, paired regressions can be used (*Freese 1967, pp. 68-70*). This method tests before- and after-treatment regressions for significant differences, but it is adaptable only to experiments in which the before and after vegetative cover is relatively constant. Figure 2 shows the paired regressions for the Hubbard Brook watersheds for the month of June. F-tests show no significant change in slope, but show a highly significant change in regression level or intercept. Thus the effect of treatment would be reported as an increase of 1.5 inches in June streamflow (differences in intercept for the two regression lines).

A conceivable advantage of paired regression analysis is that this technique tests the after-treatment observations as a group instead of individually. Also, the plotted regressions give a clearer illustration of how the extreme event is related to the remainder of the data. However, the extreme event problem remains because there is still no indication of how the before-treatment regression might change over a greater range of data.

The above discussion has been concerned with the extreme event occurring in the after-treatment period. Occurrence in the calibration period can also be troublesome. A monthly streamflow of several area-inches during the calibration period of the months shown in figure 1 would obviously carry heavy weight and could make important changes in the slope and intercept of the regression. In such cases, a minimum precaution should be the use of a second-order polynomial model for the calibration regression.

In summary, extreme events are an inherent problem in paired watershed studies. The only real solution is to greatly lengthen the calibrated period. Because this solution is generally impractical, I suggest some approaches to help insure that extreme events are treated carefully during analysis and reporting of research results.

Figure 2. — Paired regressions for determining effect of treatment on June streamflow totals.



Literature Cited

Freese, Frank.
1967. ELEMENTARY STATISTICAL METHODS FOR FORESTERS. U. S. Dep. Agr. Handb. 317. 87 p.

Goodell, Bertram C.
1958. A PRELIMINARY REPORT ON THE FIRST YEAR'S EFFECTS OF TIMBER HARVESTING ON WATER YIELD FROM A COLORADO WATERSHED. USDA Forest Serv. Rocky Mt. Forest and Range Exp. Sta. Pap. 36. 12 p.

Hibbert, A. R.
1969. WATER YIELD CHANGES AFTER CONVERTING A FORESTED CATCHMENT TO GRASS. Water Resources Res. 5: 634-640.

Hornbeck, J. W., R. S. Pierce, and C. A. Federer.
1970. STREAMFLOW CHANGES AFTER FOREST CLEARING IN NEW ENGLAND. Water Resources Res. 6: 1124-1132.

Kovner, J. L. and T. C. Evans.
1954. A METHOD FOR DETERMINING THE MINIMUM DURATION OF WATERSHED EXPERIMENTS. Amer. Geophys. Union Trans. 35: 608-612.

Reinhart, K. G., A. R. Eschner, and G. R. Trimble, Jr.
1963. EFFECT ON STREAMFLOW OF FOUR FOREST PRACTICES IN THE MOUNTAINS OF WEST VIRGINIA. USDA Forest Serv. Res. Pap. NE-1. 79 p.

Wilm, H. G.
1949. HOW LONG SHOULD EXPERIMENTAL WATERSHEDS BE CALIBRATED? Amer. Geophys. Union Trans. 30: 272-278.

—JAMES W. HORNBECK
Research Forester
USDA Forest Service
Northeastern Forest Experiment Station
Durham, N. H.