Volume 7 - Number 2 - April 1971

Reprinted from the

Water Resources Bulletin

Journal of the American Water Resources Association

Urbana, Illinois 61801
U.S.A.

TECHNICAL NOTES

COMMENTS ON THE CATCHMENT EXPERIMENT TO DETERMINE VEGETAL EFFECTS ON WATER YIELD[1]

John D. Hewlett[2]

ABSTRACT. So far most applicable knowledge about forests and water yield has come from catchment experiments. Perhaps even more practical information might have been secured during the past twenty years if more and better designed catchment experiments had been undertaken. At the very least, the old question of the main effects of vegetation on total basin water yield should now be settled, and we should be in a position to write management prescriptions containing reliable estimates of differing water yields under crop, pasture, brush and forest lands. As things are, managers and policy makers are being forced to decide what combinations of vegetal cover and land use best favor water yield before the scientific community has fully agreed on some of the salient aspects of the problem. This has led to considerable confusion in the minds of land managers in many regions of the world and may continue to do so for some time. The catchment experiment remains the surest way to furnish each region with practical knowledge of local vegetation-water-yield relations. (KEY WORDS: experimental watersheds (catchments); water yield)

Numerous papers have reported the results of catchment experiments as well as other experience relating to water yield following a change in vegetal cover, but the purpose of this paper is not to repeat the various quantitative estimates of those changes. I will rather address myself to the question of the validity and reliability of the methods used to secure such estimates in the United States and other countries. The only method productive of quantitative estimates on a watershed scale seems to be the catchment experiment, supported to some extent by time trend analyses of streamflow data and correlation analyses of crude records from many drainage basins.

Increases in water yield following the clearing of forest from whole catchments, and conversely, reduction in yields following afforestation of open land, have amounted to as much as 550 mm of streamflow per annum in several paired-catchment experiments [Hibbert, 1967, Rothacher, 1970]. Because the physical and biological reasons for these surprisingly large changes in yield due to vegetal manipulations have not yet been clearly set forth, it seems well to examine the logical basis for the catchment experiment.

To make my position clear at the outset, I will offer the proposition that theoretical or inductive conclusions about the tendency for forest stands to increase, or decrease, or not affect the quantity of water yielded by a natural drainage basin will never satisfy the watershed manager, or the public, until those conclusions have been demonstrated on a scale

[1]Paper No. (TN) 71036 of the *Water Resources Bulletin* (Journal of the American Water Resources Association). Discussions are open until six months from date of publication.

[2]Professor, School of Forest Resources, University of Georgia, Athens, Georgia 30601.

appropriate to the management problem; i.e., on a drainage basin large enough to serve as a primary water supply to a community or industry. There are some theoretical reasons to support each of the several conclusions about forest and water. However, if we take seriously the results from dozens of catchment experiments around the world—particularly those that have been set up specifically to test the hypothesis that vegetal cover affects water yield—then the evidence has become inescapable that well-forested land yields much less water than the same land in fallow, grass, or small crops. To my knowledge, there are no well-documented catchment experiments that indicate otherwise.

In the United States, about 1900, a public controversy arose over the influence of forests on streamflow, erosion, and floods. At issue was the ownership and control of many millions of hectares of forests and wildlands which formed the headwaters of the navigable streams and the source of most of the water supplies. In 1909, the controversy led the new Forest Service and the Weather Bureau of the United States to establish the first paired-catchment experiment on record. The results were reported by Bates and Henry [1928]. In brief, clearing a stand of mixed spruce and aspen from a high mountain catchment in Colorado increased basin water yield for several successive years. From this time on, the advantages of using two or more catchments in such experiments became well-known.

The theory of the paired catchment experiment is basically simple, but has been widely questioned because no thorough treatise of the method has ever been published. Bates and Henry [1928] used simple ratio comparisons between the pair rather than the now-popular regression analyses. Wilm [1944, 1949] computed the minimum number of years necessary to establish the confidence limits around experimental estimates of changes in streamflow. Wicht [1967] carried the design further in an unusually long-term experiment on a group of associated catchments with rotating treatments, the "multiple catchment experiment." Completion of his experiment, begun in 1940, and involving forest planting, grazing and burning, is expected in 1980. The treatments are comparative and the experiment has yielded results appropriate to the objectives of watershed management in a progressive manner from the date of establishment. Kovner and Evans [1954] further refined Wilm's estimates of the minimum number of years of calibration and treatment records necessary to detect changes in streamflow following vegetal cover alterations. Presently watershed researchers are using multivariable regression methods which are merely an extension of Wilm's methods and are briefly described in various papers by Reinhart [1967], Hewlett and Hibbert [1961], and Hewlett *et al.* [1969]. Anderson [1967] has successfully analyzed basin hydrologic records for storm peaks, snowmelt rates and sediment production, using multivariate and factor analyses on dozens of scattered basins in California. However, the quality of the basic records and the intercorrelation of variables in such multi-basin correlation studies have prevented the testing of specific hypotheses about forests and the water yield of basins.

In essence the paired-catchment experiment is based on the simple assumption that the relation between two basins experienced in the past will continue into the future unless some change is made on one of the basins. The necessity to account for meteorological influences in an experiment, e.g., the planting of a fallow catchment to trees, requires at least two basins and preferably two experimental periods of time. There must be a treatment basin and a control basin located adjacent to or near each other. The control basin serves as a climatic standard, to correct the experimental results for climate. A third basin, either a treatment or control basin, is highly desirable as a check on hydrologic continuity within the other two. The basins should be similar in size, shape, geology, exposure and elevation, and at the start they should have been under the same land use or vegetal cover for a number of years. The two experimental periods are usually designated the *calibration period* and the *treatment* *period.* A calibration period is not absolutely necessary unless a high degree of precision and confidence is required. For example, in the multiple catchment experiment, calibration and treatment periods run simultaneously, with one control basin remaining untreated during the entire experimental period. The control basin is maintained as nearly as possible under stable land use and vegetal cover. If yield on this basin during the treatment period is changing due to slow changes in vegetal cover or climate, it is assumed that similar changes would have occurred on the treatment basin had it not been treated. Thus results are quantitatively relative to the management practices being tested.

Some critics of catchment experimentation have hastily concluded that the time required to produce results is excessively long. They overlook the fact that the results are produced progressively from the beginning of a treatment period and that any measured treatment effects are in a scale appropriate to the nature and dimensions of the management problem. Furthermore, while the argument over the time and cost of catchment experiments goes on the years slip by and no substitute appears for the quantitative demonstration of vegetal effects on the water yield of catchments.

A brief review of the data available from the best controlled catchment experiments in the United States indicates that monthly streamflow from the treatment basin in a pair cannot be predicted to an accuracy greater than plus or minus 10 mm at the 95% level of confidence, or about 10 to 20% of the total monthly streamflow in humid parts of Eastern United States. Annual yield on some experimental pairs at the Coweeta Hydrologic Laboratory has been predicted to an accuracy of plus or minus 50 mm of streamflow at the 95% level of confidence, or about 5 to 10% of the average annual streamflow.

Wicht [1967], in reviewing some recent comments on the inadequacy of the statistical tests sometimes applied to paired catchment experiments (with particular reference to the possible lack of normality in the errors about regression), pointed out that the effect of treatment in some cases is so clear from graphical presentation that statistical confirmation may be considered superfluous. This is indeed true, as examination of published experiments will show. The necessity to make such an obvious point stems from the published disappointment of watershed researchers in detecting significant changes in water yield after sometimes minor alterations in catchment conditions. Even if changes in catchment conditions do not appear to alter water yield in amounts greater than about twice the standard error of estimate, we may at least conclude safely that the effect must be smaller than that amount. The important question to answer is this: Was the standard error in the experiment so large that no possible change in yield could have been detected? If the answer is yes, then the experimenter is wise to abandon the catchment method and seek another approach. The decision is critical, of course, because the researcher will be hard put to prove his objectives practical if they lead to changes in water yield so minor that he can never show them on a catchment basis.

There are many concerns in establishing an experimental set of catchments. A major one is the size of the experimental basins. If the basin water divides cannot be verified by drilling, and if the underlying bedrock is not clearly exposed at the point where streamflow is to be measured, the treatment basin should be larger than a first order basin, i.e., a basin containing at least the first order perennial stream. The effect of the treatment should be anticipated also, to be sure that the drying up of the treated basin's streamflow does not confuse the analysis. If this latter condition is not met, there is always question about the influence of treatment on subsurface water yield occurring during the time when there is no flow over the measuring station. As long as flow is above the stream bed, we may assume that the great majority of effect of treatment is quantitatively included in measured streamflow, and that subsurface leaks (underflow) or diversions elsewhere in the basin are constant or are accounted for in the calibration relationship. Possible fallacies in this assumption are subject to test through use of the control basin.

Experience indicates that a pair of basins each about 50 to 100 ha in area is a manageable unit to meet the requirements in many areas. Some terrains should be avoided; karst terrain, flat coastal basins, and basins formed by tilted geologic strata. If these terrains are avoided in selecting an experimental pair, the uncertainties of underground transfers can be reduced by increasing basin size.

The results of water yield studies on experimental basins have been criticized because the areas are too small to allow proper operation of meteorological processes in evapotranspiration and water yield after treatment [Penman, 1963; Rakhmanov, 1966]. In this view, it is feared that results from an experiment on a small basin, will not give estimates of treatment effect that can be applied with any confidence to a 100 ha, 1000 ha, or 10,000 ha basin. Although these criticisms have been almost entirely speculative, there are good reasons to avoid experiments on basins at both extremes of the scale. As Penman [1967] has pointed out, removing large forest trees from a plot leaves a wall of trees around a deep hole in the forest. For several reasons, meteorological as well as edaphic, we could be measuring the effect of a hole in the forest on evaporation, rather than the effect of forest per se. It is, however, unreasonable to apply this criticism to results involving experiments on basins 10 to 20 ha or larger, at least outside snow pack areas. Furthermore, basins of the order of 100 ha are approaching the size of areas often used as watersheds for municipal and industrial water supplies. Any acquired knowledge of their behavior is immediately of practical value whether or not a regional prediction can yet be made.

Basins of the order of 1000 ha or larger will be avoided for experimental purposes simply because it will be impractical to apply an experimental treatment uniformly over such an area. But it will be quite feasible to use such a large basin as a "pilot test" of practices shown to be effective in altering water yield on smaller experimental pairs or groups. In experimenting with the water balance of land areas, we gain some advantages of scale up to a certain size, perhaps in the order of 100 to 200 ha, but above that size we lose measurement control of various components of both the treatment and the hydrologic cycle.

In catchment research, by far the most important instrument is the device for measuring hourly discharge by the stream. Gauging in the natural streambed will seldom if ever serve in water yield experiments. The elements of the gauging device are three: (1) A secure cut-off wall, preferably of concrete, anchored into solid material in such a manner as to eliminate the possibility of time-variable leaks beneath the wall during the experiment. (2) A simple, pretested weir section for controlling the flow of water. (3) An accurate, reliable (battery-operated), solidly-installed, water level recorder. Books on hydrometry contain almost too many designs and variations of streamgauging devices. The experimenter will be wise to study the manuals on weir design, and to specify closely the precision and accuracy levels he expects of the finished device, inspecting the construction personally to see that specifications are met.

The critical requirement in water yield studies by the paired watershed approach is not precision in rating a particular hydraulic structure but in securing a stable underground cut-off. The cut-off need not be perfectly water-tight. Small, steady leaks beneath or in the vicinity of weir walls (or for that matter anywhere in the drainage basin) detract little from the value of the basin as a treatment or control unit, unless the leaks form a very substantial portion of the total water yield. To be sure, such leaks may severely damage the determination of the total water balance of a single basin from measurement of precipitation and streamflow.

However, the most valuable results from research on water yield have come not from complete solution of the catchment water balance but from comparative, partial solutions of paired or grouped catchments. The use of covariance analysis among catchments permits some cancellation of terms and allows quantitative estimates of seasonal changes in water yield following an experimental treatment [Hewlett *et al.*, 1969].

Instrumentation other than the stable streamgauging station is secondary in the catchment experiment. Precipitation should be measured (perhaps "indexed" is a better word because weighted basin precipitation estimates are seldom accurate to 10% of the mean values). The gauges should be located around the experimental area but preferably not within the catchment on which a change in vegetal structure is expected to occur. Experience has shown that cutting trees or reforesting open fields around gauge sites changes the percentage of the gross rainfall caught by any standard gauge.

As long as the Wagon Wheel Gap Experiment was the only catchment result available, we were indeed justified in labeling catchment evidence "circumstantial." But as dozens of similar experimental pairs were added between 1940 and the present, virtual replication has been secured in fact if not strictly in accordance with classical design theory. Many nations have been involved in a world-wide "multiple catchment experiment" for forty years, and the general conclusion is fairly clear, though a combined analysis has never been undertaken.

Many basic questions about hydrologic processes on watersheds will never be answered by catchment experiments. The ultimate model of mass input-output relations of the drainage basin including all influences of land use and cover on these relations, must rely on a better understanding of basic hydrologic processes than we now have. Hydrologists have said this repeatedly over the past twenty years and much progress has indeed been made, e.g., in evapotranspiration-energy relations, in soil water physics, and in source area runoff processes. It remains true, however, that most of the applicable knowledge about forests and water yield has so far come from catchment experiments. Looking backward, it seems to me that a de-emphasis of catchment experimentation in favor of "basic" studies came too soon as a result of some discouragement with long-term research projects that did not appear to be converging rapidly toward final answers. But the situation has quietly changed; the accumulating experience with experimental catchments around the world is leading to re-examination of the idea that the only way to make progress is through basic, non-catchment studies of hydrologic processes. One cannot question the ultimate usefulness of acquired knowledge regardless how secured; only the relative cost can be questioned. At this stage, however, the cost question is difficult to resolve. Perhaps even more practical information might have been secured during the past twenty years if more and better designed catchment experiments had been undertaken. At the very least, I suspect that the old question of the main effects of vegetation on total basin water yield would now be settled. Furthermore, we might well be in a position to write management prescriptions involving quantitative estimates of the differing water yields under crop, pasture, brush and forest lands. Under current pressures toward full utilization of land and water resources, managers and policy makers are being forced to decide what combinations of vegetal cover and land use best favor water yield before the scientific community has fully agreed on some of the salient aspects of the problem. This has led to considerable confusion in the minds of land managers in many regions of the world and may continue to do so for some time. Perhaps the catchment experiment remains the surest way to furnish each region with practical knowledge of local forest-water relations.

REFERENCES

Anderson, H. W. 1967. Watershed modeling approach to evaluation of the hydrologic potential of unit areas. In: Forest Hydrology, W. E. Sopper and H. W. Lull (Eds.), Proc. Intern. Symp. on Forest Hydrol., Pergamon Press, Oxford, pp. 737-748.

Bates, C. G. and A. J. Henry. 1928. Forest and streamflow experiment at Wagon Wheel, Gap, Colo. Monthly Weather Review, Supplement No. 30, Government Printing Office, Washington, 79 pp.

Hewlett, J. D. and A. R. Hibbert. 1961. Increases in water yield after several types of forest cutting. Quart. Bull. Intern. Assoc. Sci. Hydro. 6:5-17.

Hewlett, J. D., H. W. Lull and K. G. Reinhart. 1969. In defense of experimental watersheds. Water Resources Research 5(1):306-316.

Hibbert, A. R. 1967. Forest treatment effects on water yield. In: Forest Hydrology, W. E. Sopper and H. W. Lull (Eds.), Proc. Intern. Symp. on Forest Hydrol., Pergamon Press, Oxford, pp. 527-543.

Kovner, J. L. and Thomas C. Evans. 1954. A method for determining the minimum duration of watershed experiments. Trans. Amer. Geophys. Union 35(4):608-612.

Penman, H. L. 1963. Vegetation and hydrology. Commonwealth Agricultural Bureau, Communication 53, 124 pp.

Penman, H. L. 1967. In Discussion of: Summary of instrumentation and techniques session (R. E. Dils). In: Forest Hydrology, W. E. Sopper and H. W. Lull (Eds.), Proc. Intern. Symp. on Forest Hydrol., Pergamon Press, Oxford, pp. 782-783.

Rakhmanov, V. V. 1967. Role of forests in water conservation. (Israel Program for Scientific Translations, Oldbourne Press, London, 1966), Reviewed by L. Leyton in Nature 214:1960-1061.

Rothacher, Jack. 1970. Increases in water yield following clear-cut logging in the Pacific Northwest. Water Resources Research 6:653-658.

Reinhart, K. G. 1967. Watershed calibration methods. In: Forest Hydrology, W. E. Sopper and H. W. Lull (Eds.), Proc. Intern. Symp. on Forest Hydrol., Pergamon Press, Oxford, pp. 715-723.

Wicht, C. L. 1967. The validity of conclusions from South African multiple watershed experiments. In: Forest Hydrology, W. E. Sopper and H. W. Lull (Eds.); Proc. Intern. Symp. on Forest Hydrol., Pergamon Press, Oxford, pp. 749-760.

Wilm, H. G. 1944. Statistical control of hydrologic data from experimental watersheds. Trans. Amer. Geophys. Union, Part II:618-622.

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