SOIL MOISTURE SAMPLING PLAN FOR WATERSHEDS

February 1964 ARS 41-87
Agricultural Research Service

UNITED STATES DEPARTMENT OF AGRICULTURE

SOIL MOISTURE SAMPLING PLAN FOR WATERSHEDS 1/

J. L. McGuinness and J. B. Urban 2/

INTRODUCTION

This report is intended to introduce the hydrologist who has little statistical background to the general concepts of sampling theory as they apply to soil moisture measurements in hydrology. Frequent random selection of sampling sites in a watershed is usually not practical. As a rule, the procedure is to choose fixed sites, install instruments, and read them periodically. Furthermore, soil moisture evaluations need to be related to soil profile characteristics. In fact, the hydrologist must know the surface and subsurface characteristics of the watershed and have a clear idea of the objectives of the soil moisture program before the statistical layout can be planned.

The purpose for which the sampling is made will influence the intensity of sampling. If the main purpose of the sampling is to estimate the "true" average moisture content of the soil to some specified depth, samples may be required at numerous points in the watershed. If only a rough index of watershed moisture is needed, fewer sampling sites may be required.

A guide for designing a soil moisture sampling plan for any specific watershed area, based on an example from the Coshocton Station, is described in this report.

RELIABILITY OF AVERAGE SOIL MOISTURE FOR A WATERSHED

For routine hydrologic work, adequate determination of soil moisture at a specific point in a watershed presents little difficulty. Shaw and Arble's bibliography (2) 3/ lists 629 references on soil moisture determination and measurement.

The estimate of "true" watershed moisture is an average of several widely varying point estimates, each one of which is considered reasonably precise. For a given watershed, an estimate of the "true" average moisture content to a specified soil depth is usually required. Measurements are made at several

---

1/ Contribution from the Cornbelt Branch, Soil and Water Conservation Research Division, Agricultural Research Service, U. S. Department of Agriculture, Coshocton, Ohio, in cooperation with the Ohio Agricultural Experiment Station, Wooster, Ohio.

2/ Statistician and Geologist, respectively.

3/ Underscored numbers in parentheses refer to Literature Cited at end of this report.

locations in the watershed and, with proper precautions, the moisture determined at each of these points will be close to the "true" moisture of the points. In normal watershed soils, however, it is likely that the value of the point determinations will vary widely from one location to another.

Other than to increase sampling intensity, the best way to improve the reliability of the watershed average must lie in the method of selecting the number and location of the sampling points. Obviously, the precise point estimates are not much use unless they can be combined into a reasonably good estimate of average watershed moisture. Increasing the precision of the point readings only will do little good to meet this particular need.

What is a reliable estimate of average soil moisture? The estimate is only an estimate and, therefore, to some degree uncertain. This uncertainty, however, should be kept as small as practicable. If the estimates of watershed average moisture obtained from repetitive sampling are consistently higher or lower than the "true" average, the sample estimates are biased. Thus, the ideal is to obtain (1) an unbiased estimate, and (2), an estimate with minimum uncertainty. If these two conditions are met, the average may be termed reliable and merit trust and confidence. The principles of statistics and sampling theory can be very useful in the design of a moisture sampling program for deriving a reliable average.

Delineation of Homogeneous Subareas

Variability in soil moisture is expected from point to point in the watershed. Soil moisture is not the same from place to place because of variations in slope, aspect, soil type, vegetative cover, clay content of the soil, and distribution of rainfall.

The entire watershed area may be thought of as a collection of subareas, each somewhat homogeneous in itself. Grass on one soil type near the top of a slope may have a different soil moisture regimen than grass on another soil type near the bottom of the slope. Of course, there will be some variation in moisture content from one point to another point within a homogeneous subarea. Complete knowledge is not available on factors that influence soil moisture. However, the variation in a subarea should be less than the variation between points in different subareas.

The statistical objective in delineating subareas is to insure that, when the total variability is partitioned, most of the variation is between subareas rather than between sampling points in subareas. Theoretically, differences between perfectly defined subareas do not add uncertainty to the estimate of average watershed soil moisture. The variability within the perfectly defined subareas wholly determines the uncertainty associated with the estimate of mean moisture. The success or failure of any soil moisture sampling plan is critically dependent upon a good delineation of homogeneous subareas that make up the watershed to be sampled.

The delineation of subareas having a homogeneous moisture regimen can be difficult because each watershed is a unique problem. The advice of a soil specialist in this phase of the procedure is desirable. Although the subareas must be delineated on a hydrologic basis, the soil specialist's knowledge of soil density, texture, and other factors that affect soil moisture is invaluable. In addition, he can recognize other local factors that affect soil moisture.

As an example, the moisture regimen of the soil subareas in a meadow watershed on the Coshocton Station may be different (fig. 1). The basis for delineating subareas may be differences in soil type, volume weight, or other factors peculiar to the watershed. The influence of aspect or other factors upon the moisture regimen is expected to be uniform within any particular subarea, but it may differ between subareas.

It is not necessary to be certain that two subareas have different moisture relationships. Mere suspicion is enough to qualify them as separate entities. Homogeneity between subareas carries little penalty, but heterogeneity within a subarea cannot be tolerated.

The procedures outlined in this report are equally applicable to the sampling of subsoil layers, although only the topsoil will be used to illustrate the method. A good sampling plan for determining watershed moisture in the topsoil may be inadequate for sampling the subsoil. Thus, for the watershed of figure 1, a second sampling plan may be required for the subsoil layer, and a third plan may be needed for the next underlying layer.

Point measurements

Having delineated homogeneous subareas of the soil layer to be sampled, the next step might be to select the type of sampling, i.e., gravimetric, resistance blocks, neutron surface meter, etc. Cost, feasibility, and the experimenter's previous experience will affect the selection of the sampling method. For the sake of illustration, assume that sampling will be done by use of electrical resistance blocks.

For the watershed shown in figure 1, how many resistance blocks should be allotted to each sampling point to obtain reasonably precise estimates of the "true" soil moisture at these points? Previous experience can help answer this question.

Calculation of the number of resistance blocks required per sampling point is straightforward (3, p. 60), once the standard error of a set of block readings, s_{WP}, is obtained. The amount by which the average of the blocks varies from the "true" moisture value can be estimated from the standard error. The standard error might be computed from previous samplings in other areas; it might be obtained from a preliminary sampling made just for this purpose; or it might be estimated on the basis of other workers experiences as reported in the literature. In any case, the value used for s_{WP} is an estimate - not a precise figure. The calculated number of blocks required is also not precise, although the estimate will be adequate unless an unreasonable value of s_{WP} was chosen. Since the cost of additional blocks at a sampling point is small, oversampling may be practical.

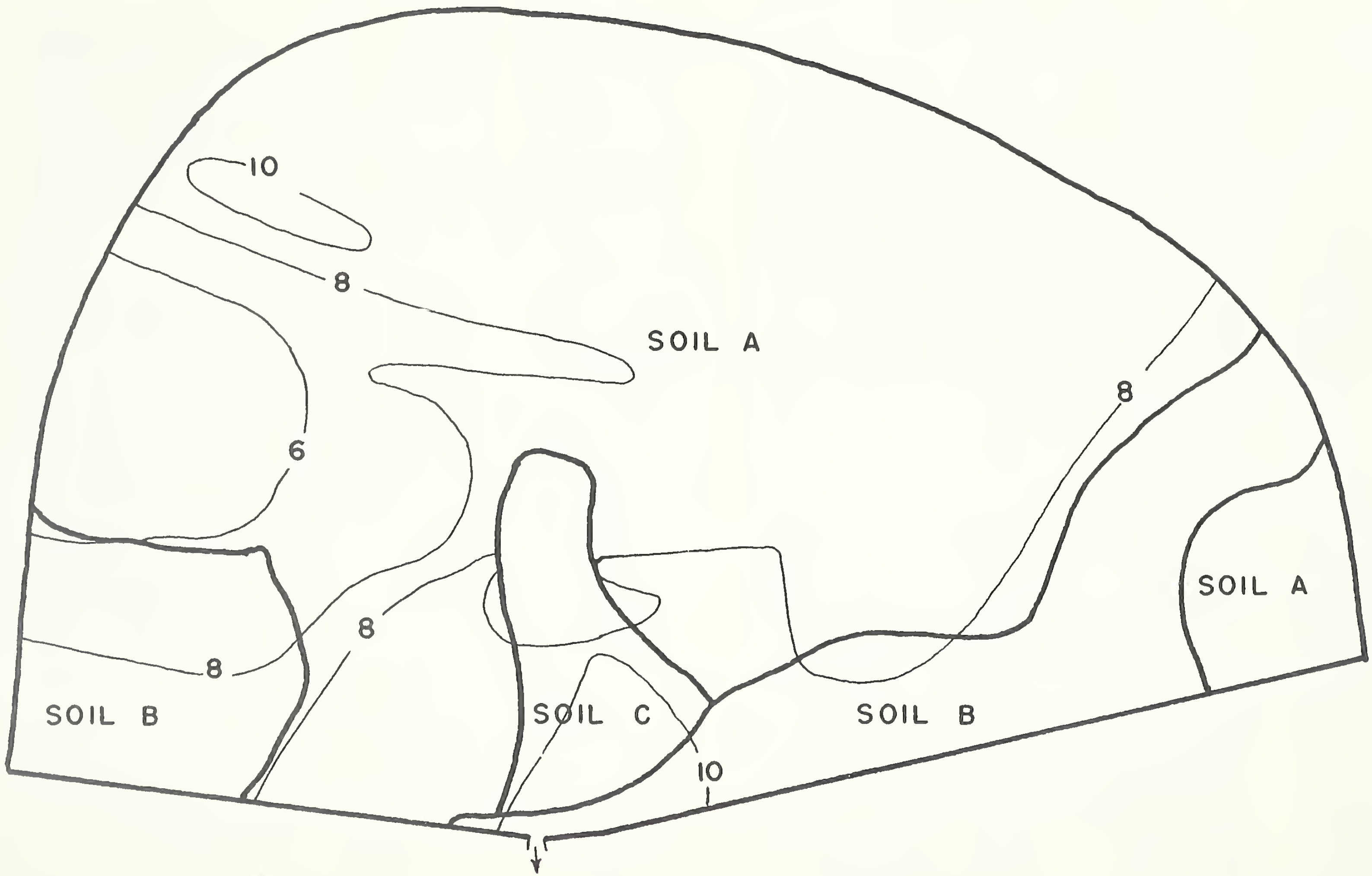

Figure 1. Fine lines show thickness in inches of topsoil in a meadow watershed; heavy lines are the boundaries of a single soil.

Number of sampling points

The method of obtaining an adequate point measurement having been determined, the next step is to determine the number of sampling points in the watershed area. The variation in thickness of the topsoil layer and the extent of homogeneous subareas were shown in figure 1. The volume of topsoil in each subarea can be computed from delineations similar to those shown in figure 1 since both area and depth are specified. The proportionate volume of topsoil in each subarea in the watershed is shown in the following tabulation.

Subarea:	Topsoil Volume, Percent
Soil A	74.4
Soil B	19.6
Soil C	6.0
Watershed Total	100.0

The tabulated percentages are the relative weights of the topsoil volume in each homogeneous subarea. It seems logical to apportion the sampling points among the subareas in proportion to these relative weights. If prior knowledge shows that one subarea is likely to be less homogeneous than the others, then that subarea could be allotted additional sampling points. The B-soil subarea (fig. 1) is likely to be in this category. In the absence of such information, proportional sampling is indicated.

How many sampling points should be placed in the A-soil subarea (fig. 1), which constitutes about 74 per cent of the watershed topsoil? Sampling theory can be of some help in answering this question.

The immediate objective is to arrive at a reliable average soil moisture for the topsoil of the watershed, i.e., an average close to the "true" average and an average with a minimum amount of uncertainty attached to it. The initial objective will be discussed later in this report. Consideration will be given first to the uncertainty of the average, or in statistical terms, the standard error of the mean.

In describing the block layout at the sampling point earlier in this report, reference was made to the source of an estimate of the standard error of a set of block readings, s_{WP}. From the same source, an estimate can be obtained of the standard error of several sets of block readings, s_{BP}, which is a measure of the variability between sampling points within a homogeneous subarea. Assume that the allowable error, L, in the estimate of the amount of soil moisture in the topsoil layer is 0.1 inch, and set the fiducial limits at 95 per cent, $t = 2$, (to be about 95 per cent sure that the sample estimate is within $L = 0.1$ inch of the "true" value).

The required sample size may be computed from the formula (3, p. 501):

$$n = t^2 \, s_{BP}^2 \, / \, L^2$$

where n = the number of sampling points.

The degree of confidence (95 per cent sure) determines the numerical value of t, which is tabled in most statistical texts. Using a rough estimate of 0.03 for s_{BP}^2, it follows that about 12 sampling points should be allocated to soil subarea A of the watershed shown in figure 1.

The formula given previously produces a good approximation rather than a precise number. The formula is theoretically exact, but the value used for s_{BP} must inevitably be an estimate. It would not be at all amiss to add one or two sampling points to the computed number to allow for the possibility of a poor estimate of the value of s_{BP}.

Having allocated 12 sampling points to A-soil subarea, we use the topsoil percentages to determine that four sampling points are needed for the other two subareas on the basis of proportional allocation. Of these four points, three should be located in the B-soil subarea and one in the C-soil subarea. However, as noted previously, the B-soil subarea is likely to be less homogeneous than the others, so one extra sampling point can be allocated to it.

On this particular watershed, the possible variation in soil moisture due to position-on-slope needs investigation. From an inspection of figure 1, it looks feasible to install a line of sampling points from the top to the bottom of the watershed in the A-soil subarea. These sampling points should be closely spaced to provide a detailed definition of the slope effect if it does indeed exist. Therefore, two extra sampling points are allotted to the A-soil subarea to help define the influence of position-on-slope.

In summary then, 14 sampling points are allocated to the A-soil subarea, 4 to the B-soil subarea, and 1 to the C-soil subarea - a total of 19 sampling points (fig. 2).

Location of sampling points

If each subarea were perfectly homogeneous, the location of sampling points would not matter because each possible location in a subarea would give the same moisture reading. If this were true, only one sampling point per soil subarea would be needed. However, the delineation of subareas is hardly that accurate.

Intuitively, a set of sampling points well spaced throughout the subarea seems more desirable than a set of points unevenly spaced. Mathematically, it can be shown that if there is a trend in the moisture regimen across the subarea due to topography or rainfall pattern or other factors, the well-spaced sampling points are most efficient in defining it (5). They are also more likely to detect any small section which may differ from the remainder of the subarea. Cochran's (1) work indicates that a systematic selection of well-spaced samples should result in a smaller variance than if a stratified random selection of samples were used. In addition, prior knowledge of the expected moisture regimen is somewhat sketchy. These are persuasive arguments for the use of well-spaced sampling points. The disadvantages of a systematic placing of sampling points will be discussed later.

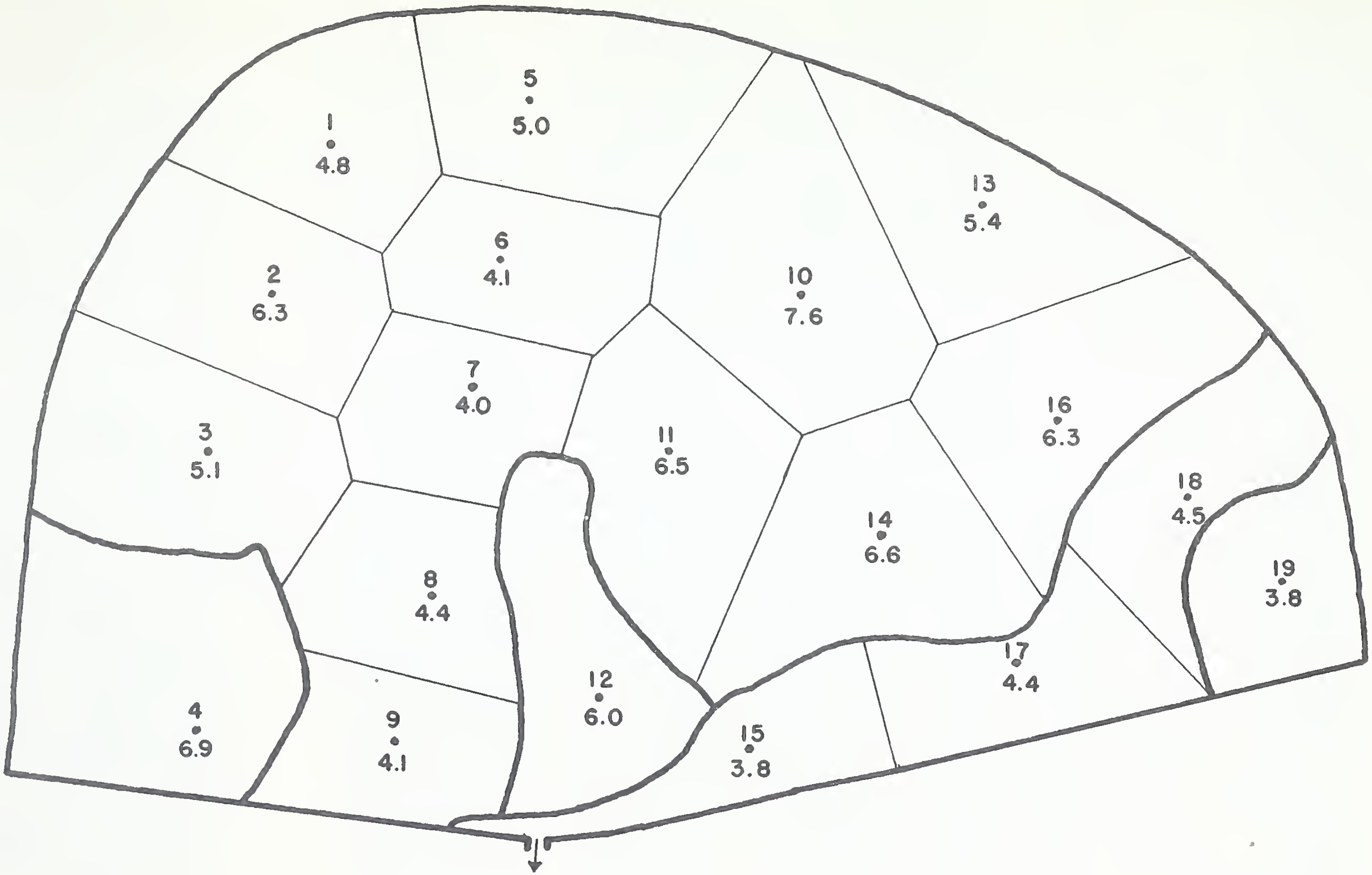

Figure 2. Location of moisture block sampling sites for the determination of topsoil moisture in soil subareas A, B, and C. The proportionate extent of the topsoil volume in each sampling area is also shown. (See also figure 1.)

Plots 5 to 9 in A-soil subarea (fig. 2) were placed to detect the influence of position-on-slope. This is the only line along which this effect could be measured within a single subarea. For the whole watershed, the percentage of topsoil volume represented by each sampling point ranges from 3.8 to 7.6.

The method of assigning area to a sampling point in order to obtain weighting factors for the volume of soil represented by the point can be described by illustration. Plot 4 is located in an isolated part of B-soil subarea (fig. 2). The limits of the soil area represented by this plot are defined by the heavy line separating the B- and A-soil subareas. The limits of plots 12 and 19 are similarly defined.

Now consider plot 6 in the large A-soil subarea. So far as is known, the moisture regimen near plot 6 should be similar to that in the vicinity of the surrounding sampling points. Perpendicular bisectors are drawn between the sampling point of plot 6 and the surrounding sampling points so that any point within the polygon around plot 6 is closer to the center of plot 6 than to the center of any other plot. This method of assigning areal representativeness to a point reading was used by Thiessen (4) for determining precipitation averages. The volume of topsoil represented by each point is the product of polygon area and its average soil depth.

A good scattering of sampling points has been achieved within the limitations of the soil subarea and position-on-slope restrictions (fig. 2). Some further adjustments in location could probably be made. Plots 1 and 3 could be moved closer to plot 2 so that each of these points would represent 5.4 per cent of the topsoil. Such minor adjustments do not appear warranted.

COMPUTATION OF AVERAGE WATERSHED SOIL MOISTURE

The average watershed soil moisture of the topsoil layer is computed as a weighted average. Each point average is multiplied by the per cent of volume of the topsoil layer that it represents, and these products are summed to obtain the watershed average. Calculation of the uncertainty of this average, the standard error of the weighted mean, is straightforward (3, p. 506).

SIZE OF WATERSHED

It may be surprising that the sampling plan shown in figure 2 has been determined without considering size of watershed as a separate factor. Size of watershed area enters the picture when the standard error between plots in a subarea, s_{BP}, is estimated. If the subarea is made up of sections that lie within an area of about an acre or so, the data used to estimate s_{BP} should be from a set that occupies approximately an acre. If the subarea is spread out over a square mile, the data used to estimate s_{BP} should come from plots located relatively far apart so that the increased variation associated with the larger area is included in the estimate.

Sampling will naturally be more difficult as size of area increases. A thousand or more subareas may be found in a square mile. The limited resources usually available may make it impossible to install enough sampling points to estimate average soil moisture with sufficient reliability. In this case, the experimenter might install a permanent, systematic layout consisting of widely spaced sampling points and supplement this basic layout with temporary sites as resources permit. After enough observations have been taken to allow correlation of the permanent and temporary site readings, the temporary sites can be moved to other subareas of the watershed and the process repeated.

BIAS

The systematic placement of well-spaced sampling sites shown in figure 2 was designed to obtain a reliable estimate of average watershed soil moisture. This reliable estimate should closely approximate the "true" average and have a minimum amount of uncertainty. Discussion of the approximation of the "true" average was deferred to this point in the report while the uncertainty concept was being developed.

Suppose the 14 sampling sites in the A-soil subarea had been chosen at random. Then the estimate of mean soil moisture for the subarea would be unbiased. This means that if numerous samplings were made and there was a rerandomizing of the site selection each time, the average of the differences between the estimated and "true" moisture would tend toward zero. This result would be reassuring if it were practical to rerandomize after each sampling. Actually, the sites will be chosen, the resistance blocks will be installed, and many samplings will be made by use of this one installation. Thus, a random selection of sampling sites in this case will not assure an unbiased estimate; neither will a systematic selection of the sampling sites.

However, it is difficult to visualize how the systematic sampling could be even deliberately biased. Any sections of the watershed that might have higher or lower moisture regimens than surrounding sections have already been delineated into subareas. Within the limits of the restrictions imposed by this systematic plan, the experimenter might ask whether the 14 sampling sites could be placed in the A-soil subarea in such a way as to deliberately obtain a moisture estimate for the subarea that was higher than the "true" moisture content. If this can be done, it indicates that the subarea is still not homogeneous, and that further delineation is required.

DISCUSSION

The problem of sampling soil moisture in a watershed is analogous to the mathematical description of a three-dimentional geometric system where x and y specify location and z represents the moisture content. It has been shown (5) that the most efficient exploration of the z-surface and the estimation of its magnitude are accomplished by sampling in a systematic pattern.

The only real objection to systematic sampling is that there is no completely reliable method for estimating the amount of uncertainty - the standard error - of average watershed soil moisture. If the usual method of computing the standard error is used, confidence in the result must be somewhat less than for random sampling. However, in view of the care taken in delineating subareas and the inability to introduce deliberate bias, the degree of faith in the standard error should be only a little less than if the sampling were random in fact.

SUMMARY

This report deals with the determination of soil moisture, which is one of the difficult problems in hydrologic measurement. Subareas of uniform moisture regimen were delineated on the basis of a wealth of background information on soils, geology, and slope-drainage conditions. The combination of statistical methods with this knowledge is the essence of this report. Statistical methods alone or standard field and laboratory investigations by themselves are not sufficient to provide a sound approach for the determination of average watershed soil moisture.

The systematic selection of well-scattered sampling points insures an estimate of mean soil moisture close to the "true" value. Careful delineation of homogeneous subareas insures an estimate of soil moisture with a near-minimum of uncertainty. The major objective of sampling - to obtain a reliable estimate of average soil moisture - has been achieved at the cost of a slight degree of confidence in the standard error computation.

All too often, data are analyzed without full recognition of the method initially used in sampling and collecting the data. This report is intended to point up the importance of systematically gathering data for hydrologic analysis and for the necessity of carefully evaluating data that presumably represent "true" field conditions.

LITERATURE CITED

(1) Cochran, W. G.
1946. Relative accuracy of systematic and stratified random samples for certain class of populations. Ann. Math. Statis. 17: 164-177.

(2) Shaw, M. D., and Arble, W. C.
1959. Bibliography on methods for determining soil moisture. Engin. Res. Bul. B-78, 152 pp. The Pennsylvania State University, University Park, Pa.

(3) Snedecor, G. W.
1956. Statistical methods. Ed. 5, 534 pp. The Iowa State College Press, Ames, Iowa.

(4) Thiessen, A. H.
1911. Precipitation averages for large areas. Monthly Weather Rev. 39: 1082-1084.

(5) Tyler, G. W.
1953. Numerical integration of functions of several variables. Canad. Jour. Math. 5: 393-412.