A SAMPLING SYSTEM FOR POPLAR INSECTS1

BY W. R. HENSON?

Abstract

The distribution of poplar foliage over the trees is described for a number of locations in the trans-Divide area of the Rocky Mountains of Canada. The
statistics of some foliage attributes and of the distributions of a number of
organisms infesting foliage are derived. On this basis, a sampling syst preset limits of precision and confidence is designed. The system is tested and its limitations discussed. Techniques for the modification of the system to fit local conditions are suggested. The system outlined will provide in a single operation data of known accuracy on the populations of a number of organisms.

Introduction

During the summer of 1953, a long-term study of insects inhabiting poplar in the trans-Divide region of the Rocky Mountains was begun. The project is designed to test the hypothesis (7) that insects which occupy different types of habitats may often require different climatic conditions for their optimum development, survival, and increase. A very satisfactory tree species for testing this interesting idea is trembling aspen *(Populus tremula'des* Michx.) because it supports a wide variety of insects, and provides a number of different types of insect habitats. In addition, it is widely distributed so that the investigation could be extended to many different regions after techniques have been developed and tested in the trans-Divide area, where differences in climate often occur over short distances.

The first requirement for following population fluctuations is a sampling system which will give adequate measurements of population densities. The present report deals with investigations leading to the development of such a sampling system.

Statistical Investigations on *Populus trernuloides*

General Considerations

The object of the investigations was the development of a sampling system which would provide data on the populations of insects and disease organisms inhabiting the leaves and twigs of trembling aspen. As the use to which this sampling system is to be put will probably enlarge beyond the scope of the project as it is presently envisioned, an attempt was made to place it on as broad a basis as possible. The system had to be adapted for use by a variety of field worlcers of varying degrees of training and skill. Therefore, it was necessary to maintain simplicity of operation even if this meant extra labor in the actual field collection of the samples.

The sampling system had to accommodate local peculiarities of stands and individual trees. Because several kinds of organisms must be sampled at the

 $\mathbf 1$ *Manuscript received August 31, 1954.*

Contribution No. 172, Forest Biology Division, Science Service, Department of Agriculture, Ottawa, Canada. *Forest Biology Laboratory, Calgary, Alberta.*

same time, it had to provide sufficient data on the organism presenting the most difficult sampling problem, even at the expense of oversampling more evenly distributed organisms. The system had to be designed so that the data would appear in such a form that records could be used for the purposes of a variety of special analyses. The density of the organisms sampled, as well as the nature of the comparisons to be made, fixes the manner in which field data are presented for analysis. However, for actual use in the analysis of population fluctuations, field data will probably have to be converted to some form of absolute density expression on a per tree or a per unit area basis.

FIG. 1. The location of sampling points within the Rocky Mountain National Parks of Canada. 1 Mile 0 7 Great Divide

2 Mile 3.5 7 8 Cathedral

3 Mile 14 9 Leanchoil $\begin{array}{ll}\text{Mile} & 0\\ \text{Mile} & 3.5 \end{array}$ 3 Mile
4 Mile
5 Mile Cathedral

Mile

9 Leanchoil
10 Snow Cree 10 Snow Creek
11 Kootenay C 11 Kootenay Crossing
12 McLeod Meadows McLeod Meadows

HENSON: SAlMPLING SYSTEM 423

Location, Materials, and Methods

Trees for study were selected from three areas: Kootenay, Banff, and Yoho National Parks. A number of locations within each national park were sampled, as shown in Fig. 1. Banff Park lies to the east of the Great Divide, whereas the other parks are both to the west of that barrier. Yoho Park sample locations are all within the defile of the Kicking Horse Pass, to the north of the massif formed by the Lake Louise and the Cathedral-Stephen groups of mountains. The sample locations in Kootenay Park are in the valleys of the Vermilion and Kootenay Rivers, separated from the Yoho locations by the above-mentioned massif and its subsidiaries to the south. Though both the locations in Kootenay and those in Yoho are in passes which permit the low-level passage of maritime air (2), their aspects and exposures are very different and there are indications **(3)** that their climates show corresponding differences.

The distribution of the sample trees is given in Table I. The sampling was concentrated at Kootenay Crossing. This was done because time did not permit widespread intensive sampling, and the Kootenay Crossing stand was selected as both suitable and typical. Therefore in the following discussion, the results will be presented first for this stand. The characteristics of the attribute under discussion will be derived for the Kootenay Crossing stand and used as a basis of comparison for the single tree samples taken from other locations and other areas.

Trees of Populus tremuloides were selected randomly within each location. Each tree was felled, its total height recorded, and the orientation and arrangement of its various segments were marked. Thereafter it was cut into smaller parts for easier handling. Branches were treated individually. Records

THE DISTRIBUTION OF SAMPLE TREES BY AREA AND LOCATION

were made to show (a) the height and orientation of the branch; (b) the location of each leaf bunch* as numbered from the trunk to the periphery; (c) the number of leaves in each bunch; and (d) the number of organisms on each leaf and the total for the bunch. These data were recorded in such a way that totals could be obtained rapidly, but the original data on leaf populations could be referred to as required.

The crown of the tree was divided on the basis of height into lower, middle, and upper thirds. In addition, terminals were kept separate in the first analysis. The horizontal classes were the four cardinal compass quadrants. The first problem was to determine the sampling unit. It was felt that if a leaf bunch could be used, the recording of the data could be simplified.

Number of Leaves in the Bunch

The mean number of leaves per bunch was derived for each of the 16 segments of each tree examined. The results were subjected to the analysis of variance on a single tree basis with the height in the crown plus terminals (four classes) and the compass direction (four classes) forming the basis of a four-by-four, single-level analysis. The over-all mean number of leaves per bunch is 4.562 ± 0.266 .

No persistent trend in the number of leaves per bunch could be detected within the tree for level or direction, nor could a trend be detected between trees, locations, or areas in the distribution of number of leaves per bunch. Thus, the leaf bunch may be considered to have substantially the same number of leaves regardless of level within the tree, direction in the crown, tree within the location, location within area or between areas.

The variance ratio for level was of border significance ($F=3.87 : 5\% = 3.86$). This ratio is apparently due to the number of leaves in the terminal bunches. The absolute amount of foliage which is contained in the terminal bunches, however, is very small. The discarding of the leaves in the terminal bunches did not affect the size of the total sampling universe ($\chi^2 = 0.00709$ for degrees of freedom 47/11) and the inclusion of the terminal bunches with those of the upper crown class did not affect the size of the upper crown with respect to the whole $(\chi^2 = 0.00659)$ for degrees of freedom 36/47). Therefore, the terminal tissue was lumped with that of the upper crown, since it was not considered advisable to discard it because of its possible ecological significance. When this was done, the variance ratio for level was reduced considerably $(F = 2.56 : 5\% = 4.07)$.

Distribution of Leaf Bunches

As the leaf bunch is remarkably constant, it is possible to base the examination of foliage distribution and the distribution of organisms on the foliage on the number of bunches and density per bunch instead of on individual leaves. This reduces the labor of recording in the field and speeds calculation.

* *The foliage of poplar is arranged in bunches of leaves growing from the terminal of the current year's growth. The bunches which are lateral to the main branches are compact, those which are terminal are diffuse.*

I-IENS0.V: SAMPLING SYSTEM **425**

The trees of the Kootenay Crossing location were analyzed separately. Differences in the numbers of bunches due to direction in the crown were uniformly insignificant. Differences due to crown level showed intermittent significance. However, when all trees from the location were analyzed together with the means for crown segments used in the analysis, the level in the crown was shown to be significant. A summary of the analysis is given in Table 11.

THE ANALYSIS OF NUMBERS OF LEAF BUNCHES FOR DIRECTION AND HEIGHT IN THE CROWN-MEANS OF KOOTENAY CROSSING TREES

The significance of level persisted when the analysis was extended to include trees from other sources. Table I11 gives a summary of such an analysis which included all trees taken from the area of Kootenay Park.

TABLE I11

THE ANALYSIS OF NUMBERS OF LEAF BUNCHES FOR DIRECTION AND FOR CROWN LEVEL-MEANS OF TREES FROM THREE KOOTENAY PARK LOCATIONS

A similar situation may be demonstrated for the trees from Banff Park. Table IV presents a summary of the analysis for the mean Banff Park values.

TABLE IV

THE ANALYSIS OF NUMBERS OF LEAF BUNCHES FOR DIRECTION AND FOR CROWN

A slightly lower level of significance is shown in the trees from Yoho Park. However, the general distribution is the same as that demonstrated from the other two areas. The summary of the analysis is given in Table V.

TABLE V

THE ANALYSIS OF NUMBERS OF LEAF BUNCHES FOR DIRECTION AND FOR CROWN LEVEL-MEANS FOR TREES FROM THREE YOHO PARK LOCATIONS

It may be demonstrated that the distribution of the leaves is similar in all three parks. A summary of such an analysis is presented in Table VI.

THE ANALYSIS OF NUMBERS OF LEAF BUNCHES FOR DIRECTION AND LEVEL WITHIN CROWN AND FOR AREA-MEANS OF TREES FROM KOOTENAY, BANFF, AND YOHO PARK

The homogeneity of the distributions for the three parks was also tested by a comparison of the residual error variances which were calculated from the mean values from each park. The comparison was made by means of the F ratio. It was found that the residual variances from all three parks were of similar order. This is an indication that there is no reason to suspect a lack of homogeneity between the populations represented by the trees of the three parks.

Summary of the Analysis of Foliage

It has been shown that the number of leaves in each bunch on the trees examined from Kootenay, Banff, and Yoho National Parks is remarkably constant with respect to position in the crown, location, and area. Because of this consistency, the number of bunches rather than the number of leaves was used as a measure of the amount of foliage. The mean number of leaves per bunch was 4.562 ± 0.266 .

HENSON: SAMPLING *SYSTEM* 427

On this basis, it has been shown that the distribution of foliage of poplar in Kootenay, Banff, and Yoho Parks is not significantly variable with respect to its compass direction within the tree, location, or area of origin. However, there is a significant difference in the amount of foliage at various heights in the crowns of the trees.

For sampling purposes, therefore, foliage may be taken at random between trees and with respect to direction within trees at any location. However, the sample must be stratified with respect to level within the crown. This is best done (4) by distributing the sample in the same ratio as is exhibited by the variances calculated for levels in the foregoing analysis. The variances for levels are: lower 112993.58, middle 68012.44, and upper, 39436.47. Roughly, this ratio is $3:2:1$. Therefore, a sample which is taken at random between trees and between directions within the trees, and stratified within the trees in the ratio of lower, 3: middle, 2: upper, 1, with the required number of bunches being taken within each level, will yield a true representation of the trees from the location sampled. The number of bunches required will be established in a later section of this paper after the discussion of insect distributions and other subsidiary points.

Leaf Size

Early in the examination of the foliage from different locations, it was noted that the size of the leaves showed wide variation. At high insect densities, the amount of leaf tissue rather than the number of leaves may well be a more adequate measurement of the substrate. Therefore, the possible need for a correction for leaf size was investigated.

Leaves, taken at random within trees and between trees for each location, were brought to the field station. The outlines of 100 leaves from each location were traced on paper. The area of each leaf was measured with a planimeter and the measurements were checked by direct graphic determination. In addition, the length and maximum width of each leaf was measured. It was shown that the mean leaf size did vary with location. However, the leaf size was remarkably constant within a location. Three locations were selected for further testing. A summary of these data is presented in Table VII.

TABLE VII

A SUMMARY OF LEAF MEASUREMENTS FROM GREAT DIVIDE, KICKING HORSE PASS, AND HERBERT LAKE (MEASUREMENTS IN CM.)

Regressions were calculated for the relationships between length, width, and area of leaf for each location. Simple correlations were calculated for the same relationships and the values were compared between locations. The values found for the correlations were extremely high for all locations with the over-all coefficient of correlation between area and length of leaf being 0.9856 (for 300 pairs of values). The coefficient between area of leaf and leaf width was 0.9884 (for 300 pairs of values).

The shape of the leaves varied to some extent as shown by a slight variation in the regression coefficients between length and width and area at the various locations. However, these variations were not significant (pooled χ^2 = $0.359:5\% = 1.635$ and the over-all regression value of 0.4166 was an adequate approximation of the individual regressions for area on length. Since it is easier to measure length of leaf than to measure width in the field, the length should be used as an index of area.

In order to correct for leaf size so that population data may be directly compared between locations where the sizes of leaves are different, the mean length of 100 leaves taken at random should be determined. Division of this mean leaf length by 0.4166 will give a value for the mean leaf area. This, multiplied by the number of leaves sampled (number of bunches times number of leaves per bunch), may be used to convert the population figure for the location from number of organisms per standard number of leaf bunches to number for an aggregate leaf area. The resulting datum could conveniently be reduced to number of organisms for a standard leaf area. An example of the application of such a correction will appear in the section which deals with the actual sampling technique.

Organisms Infesting Aspen

During the recording of the data on foliage distribution, a record was maintained of the insects and diseases infesting the leaves and terminal twigs of the trees. Identifications were not possible in the field, so the organisms were simply listed by appropriate descriptive names. The present identifications are all tentative as most of them are based on immature forms. When complete series of material have been obtained, positive identifications will be possible.

As recorded in the field, the population data were in the form of a simple record of the number of organisms on each leaf bunch. This number was then broken down in the field records into the number of organisms on each leaf of the bunch. This was done at the time because the utility of the leaf bunch had not been established as a sampling unit. The present analysis was done on the basis of number of organisms per thousand bunches. This conversion was done in order to permit easy calculation and comparison of results. The data on infestation of individual leaves may well be of considerable biological interest. However, from the viewpoint of the sampling problem, it is simpler to use the larger unit of the leaf bunch rather than the individual leaf.

IIENSON: SAMPLING SYSTEM 429

The sampling technique must be based on an understanding of the relationship between foliage and insect distributions. Therefore, no corrections for leaf size were applied for the purposes of this preliminary analysis.

It was shown that all but two of the organisms infesting aspen were distributed on the tree in the same manner as the foliage. That is, the density of infection showed no variation with either level or direction in the crown.

Slight regional differences from this general finding were detected. However, these differences were not significant in the over-all distribution of the organisms involved. Differences between locality and area densities which were detected were due to over-all density variations and not to changes in the patterns of distribution within the tree. This point was established by the finding that the interactions between level, and direction in the crown, and locality were in no case significant.

Two organisms were found to show preferences for one side of the tree. For sampling purposes, it was decided to oversample sufficiently that these directional differences in density could be ignored in the sample design. Of course, such oversampling would not result in an error in the figure obtained for population since the final expression of the population is to be in terms of over-all density.

The Design of the Sample

Because of the considerations discussed above, samples of poplar foliage taken in order to determine the densities of the organisms studied should be distributed in the same manner as the foliage within the tree. The largest sample is required for organisms that show the largest proportion of unexplained variance with respect to position in the tree. As the error term of the variance analyses which were carried out on organism distributions is made up of all unassigned and residual variance, the ratio of mean squares for error to total mean squares may be used as an index of the amount of unexplained variance in the distribution of each organism. A list of such indices for each organism studied is given in Table VIII.

TABLE VIII

THE INDEX OF UNEXPLAINED VARIANCE FOR EACH ORGANISM IN THE SAMPLING STUDY

All the identifications given in Table VIII are tentative. No identifications are given for the aphids or the galls because no mature forms have been recovered yet. Descriptive names are given simply for ease of reference.

The largest amount of unexplained variance is found in the distribution of leaf spot. This might be expected because this organism displays the distribution which varies most marltedly from that of the foliage. A sampling system which will yield sufficient data for leaf spot will automatically yield sufficient data (within the same limits) for all other organisms included in this study.

Though the distributions of leaf spot (and black spot) are different from that of the foliage, it was considered advantageous to distribute the sample according to the foliage pattern. Such a sample distribution imposes oversampling in a segment of the tree but this is acceptable for the sake of simplicity. The advantage of such a technique is that it makes possible the detection of position anomalies.

The size of the sample required to give sufficient data for leaf spot within 99% confidence at a precision of \pm 5% was calculated from the formula given by Oakland (5). The formula is:

$n = 2t^2$ ba/D^2

where t is the normal deviate corresponding to the level of significance in a two-tailed test (1), D is the confidence expressed as a decimal, ϕ is the probability of the event, and $q = 1-p$.

In this case, a value for ϕ was derived from the known over-all density of the organism: $p = 0.2959$ (based on the probability of an individual leaf bearing a single spot). With this value of p , the sample size required was calculated to be 1102 leaves. The equivalent number of bunches was 240.87.

A number of preliminary test samples were talcen with the aid of random number tables (6) from the field data. These samples all contained 240 leaf bunches stratified between the crown levels in the ratio of lower, 3: middle, 2: upper, 1, and at random with respect to direction within the crown. It was found that the best precision was obtained in these preliminary samples when the leaf bunches were talcen in equal numbers from 20 trees.

A further series of 10 test samples was talcen with the aid of random number tables. Each sample contained 240 leaf bunches distributed as 12 bunches from each of 20 trees, taken at random with respect to direction in the crown, and in the ratio of *3* : 2 : 1 between the lower, middle, and upper thirds of the crown.

The test samples all yielded figures for leaf spot density which were within acceptable limits of the true density found by complete counts of the trees. Pooled χ^2 was used to test the homogeneity of the samples and the result $(p = 0.99 +)$ showed that the sample results were satisfactory.

The density of black spot, the organism which showed the second greatest diversity, was calculated from the same test samples in order to test the

IIENSON: SAiMPLING SYSTEM **43 1**

applicability of the system to the sampling of other organisms. The samples yielded figures for black spot density which were both homogeneous and within acceptable limits of the true density ($p = 0.99 +$).

Discussion

A sampling system has been designed for a series of organisms which depends for the reliability of its results on the distributions of the organisms over the crowns of the infested trees.

The system was designed on the basis of data taken during one season only. Therefore, the distributions of the organisms are known for only one density and one climatic season. It is possible that at different densities, or under different climatic conditions, the organisms will not be distributed in the same manner. In addition, if the density of one of the various organisms increases to the point where considerable defoliation results, the distribution of the foliage of the tree might well be distorted. If this were to happen, the technique of sampling and also the manner in which the results were stated would have to be modified.' Some expression of absolute density xvould be required.

Because of these factors, it would be unwise to assume that the size and distribution of the sample which has been derived in this contribution would invariably give accurate results. However, the system may be tested and, if necessary, modified to suit field conditions. Testing should be done on the basis of complete counts of foliage and organisms on a series of trees. The number of trees which would have to be examined in this way xvould depend on whether or not it was found that the patterns in the first few diflered from those described herein. However, the field counts could be held to a minimum by a continuing analysis of the results which would allow the counts to be stopped as soon as the statistical requirements have been met.

A new sample could be designed in the manner set forth in this paper which would meet the requirements of a changed situation. Because the precision and confidence of the sample are set during the process of sample design, the results of sampling, even if the sample had to be redesigned from time to time, would be directly comparable. This condition would be met because any departure from the imposed limits would be in the direction of increased precision.

The Expression of Sampling Results

If a figure for absolute density is required for the estimation of the total population of the organism, the sampling data may be expressed in terms of number per tree, or, better still, number per acre. In order to convert to such terms, it would be necessary to derive a technique for the estimation of the amount of foliage in an acre. Such a determination would have to be made for each stand investigated. This aspect of the problem has not been examined.

For the purposes of the design of a sampling system, all figures on populations are considered in terms of density. For many purposes, such an expression is far from adequate. The amount of foliage which supports the organisms may show wide variation when it is measured only in terms of the number of leaves or the number of leaf bunches. So, even expressions of populations in terms of density per leaf bunch are open to considerable error.

In order to permit comparisons of density between localities, the figures must be based on comparable segments of the substrate. In this case, the most important source of variation is the size of the leaf.

It has been shown that a very close estimate of mean leaf size can be obtained from a measurement of mean leaf length. If figures for insect density are required, the expression of density per unit leaf area is probably the most satisfactory. The conversion from field data presents no difficulty.

For example, suppose a density of 200 organisms per 1000 leaf bunches has been found from field sampling. A measurement would be made of the length of 100 leaves, taken at random within the original sample. The mean length of the leaves would be calculated. Suppose this were 3.00 cm. The mean leaf area would be $3.00/0.4166 = 7.2011$ cm.² (where 0.4166 is the coefficient of regression of leaf length on leaf area). From this, the density of the organism would be: 200 \times 10,000/7.201 \times 4.562 \times 1000 = 60.881 organisms per $10,000$ cm.² leaf surface (where mean leaf area is 7.201 cm.² and mean number of leaves per bunch is 4.562). The unit area, 10,000 cm.², is chosen because it is a convenient figure in the same order of magnitude as the area of leaf surface per 1000 leaf bunches.

Summary

1. The foliage of *Populus tremuloides* Michx. appears as a remarkably constant number of leaves on each twig of the current year's growth. The statistics of these bunches of leaves are derived, and it is shown that the number of leaf bunches may be used as a measure of the number of leaves.

2. The distribution of foliage over the crown of poplar is described.

3. The characteristics of leaf area, length, and width are derived and a technique is described for correcting data based on leaf number, for size of leaf.

4. The distributions of a number of organisms infesting poplar are discussed. A comparison of the amount of unassigned variance in these distributions is made in order to detect the most variant organism for sampling purposes.

5. A sampling system with preimposed limits of precision and confidence is designed on the basis of the distribution of the foliage and of the most variant organism with respect to foliage.

6. The sampling system is tested both for the organism for which it was designed, and for another organism. It is shown that satisfactory figures for density for all organisms included in the study may be obtained in one sampling operation.

7. Various methods of expressing population data and the use of a correction factor for leaf size are discussed.

HENSON: SAMPLING SYSTEM 433

8. The possible need for continued modification of the sampling system to take into account changes in field conditions is discussed. The technique for such examination and for redesigning the sample is outlined.

Acknowledgments

The writer would like to express his appreciation of W. G. Wellington's direction and critical review of the work. G. B. Oakland suggested many of the statistical techniques and reviewed the contribution. Field assistance was given by R. F. Shepherd, and L. C. Weir assisted with the bulk of the computational work. W. C. McGuffin and V. J. Nordin helped with tentative identifications of field material.

References

1. FISHER, R.A. and YATES, F. Statistical tables. 3rd ed. Hafner Publishing Company, Inc., New York. 1948.

2. HENSON, W. R. Chinook winds and red belt injury to lodgepole pine in the Rocky Mountain National Parks of Canada. Forestry Chron. 28:62-64. 1952. Mountain National Parks of Canada. Forestry Chron. 28 : 62-64. 1952.

3. HENSON, W. R. Bioclimatological investigations of forest insects of Western Canada. Ann. Tech. Rept. Forest Insect Lab., Sault Ste. Marie, Ontario. R4imeographed Rept. 1952.

4. KENDALL, M. G. The advanced theory of statistics. Vol. II. Charles Griffin & Co., Ltd., London. 1948.

5. OAKLAND, G. B. Determining sample size. Can. Entomologist, 85:108-113. 1953.

- 6. SNEDECOR, G. W. Statistical methods. 4th ed. Iowa State College Press, Ames, Iowa. 1946.
- 7. WELLINGTON, W. G. Atmospheric circulation processes and insect ecology. Can. Entomologist, 86 : 312-333. 1954.