2nd World Eucalyptus Conf.
Aug. 1961
São Paulo, Brazil
Report & Documents

# Statistically Valid Planting Trials

C. B. Briscoe (*)

More than 100 million tree seedlings are planted each year in Latin America, and at least ten times that many should be planted (1). Rational control and development of a program of such magnitude require establishing and interpreting carefully planned trial plantings which will yield statistically valid answers to real and important questions.

Unfortunately, many forest tree plantations being established today in this region are conceived and executed by men completely without statistical background, or by men whose training in statistics took place in the distant past. Many of these men fully recognize the advantages to be derived from proper experimental design and would like to establish their trials in a manner permitting later valid statistical tests of the results obtained. In many cases, however, there is no qualified statistician whom they can consult; and, for one reason or another, they are unable themselves to obtain an adequate statistical background. It is for such men, men who want to establish properly designed trial plantings but who are completely unfamiliar with statistical theory, that this paper was prepared.

It should be emphasized that design of planting trials to obtain reliable information about complex relationships, with minimum cost, requires specialized techniques which cannot be discussed here. When a qualified statistician is available every effort should be made to consult him BEFORE field work is begun.

---

(*) Tropical Forest Research Center — Rio Piedras, Puerto Rico.

(1) Food and Agriculture Organization. 1959. Tree planting practice in Latin America. FAO Forestry Development Paper, draft. p. 79.

## STUDY DESIGN

The fundamentals of planning experiments are very few. First, restrict the study to answering a specific question. Second, study in replicate the range of conditions within the scope of the question. Third, locate plots at random to eliminate all possibility of personal or systematic bias.

That is all. If these rules are followed, anyone can design a planting trial which will yield statistically valid results. The costs may be somewhat higher than are strictly necessary, some side effects may be overlooked, some interactions may not be discernible or provable if discovered, but the basic question under study will be answered conclusively.

Now let us consider each of the three points in somewhat more detail.

*Restrict*

Deciding exactly what is to be determined is the most important step of all. A question must be asked before an answer can be found, and the question must make clear both what is included and what is excluded.

If, for example, eucalyptus are to be introduced into an area where they are not native, a number of questions immediately arise which must be answered: What species should be grown? What sites are favorable for the particular species? How fast will the species grow on the various sites available? The questions, as most of you know, go on forever; one study cannot answer them all. The narrower the question, the more complete the answer can be.

Perhaps the two most common questions of all are : «What species can be grown?» and «On what sites will this species surpass that one?»

Assume that through consideration of latitude, rainfall regime, elevation, and other characteristics of the planting area the selection of species has been narrowed to three : *Eucalyptus saligna* Sm., *E. robusta* Sm., and *E. citriodora* Hook. The sites within the planting area may be divided into three major topographic positions: valley, slope and ridge.

To be determined, then, is whether the *E. saligna, E. robusta,* or *E. citriodora* is best adapted to each of the three sites (2).

(2) Determining the optimum number of replications, species, plot size, and other details is being discussed in another paper at this meeting, but it may be worth noting that testing more species at a time than three is more efficient. The possible number of locations, of course, is limited only by the number of seedlings and the funds available.

*Replication*

To find out, all three species must be planted in replicate on each of the three sites. This is usually done in three ways which complement each other.

Location replication : obtained by establishing plots on each site at a minimum of at least two (2) locations. Valley sites might be sampled in Valley X, Valley Y, and Valley Z.

The locations to be tested are usually selected in one of three ways:

1. *Random.* That is, every valley suitable for testing may be assigned a number and sufficient numbers be drawn from a hat to provide the number of locations to be tested.

Randomly selected locations have an equal chance of falling anywhere within the entire are under study. This means that if the planting area were Puerto Rico, plots on valley sites would have en equal chance of falling in any valley in Puerto Rico. Such randomly selected locations will normally not be evenly distributed and may well be clumped to an undesirable degree. Especially when only a few locations are selected this can easily lead to an erroneous estimate. A clump of locations in arid southwestern Puerto Rico would suggest very different conclusions from those suggested by a clump of locations in the rainy eastern mountains, and neither will provide a good estimate for Puerto Rico as a whole.

In addition, complete randomization is expensive and inconvenient.

2. *Systematic.* Systematic location is a planned, uniform geographical dispersion over the planting area. It is more likely to include all possible conditions, so that a better picture of the whole is obtained.

3. *Stratified.* Instead of on a systematic geographical basis, location may be planned on the basis of rainfall, elevations, parent rock, past land use, or any other important and identifiable factor, or combination of factors, affecting tree growth.

For example, if we wished to stratify on the basis of rainfall is approximately 150 inches, another location in the contral mountain where the rainfall is approximately 100 inches, and another plot on a southern slops where the rainfall is approximately 50 inches. ?

On the other hand, if we wished to stratify on the basis of major soil types we might establish locations in the sandy loams, deep clays, and shallow clays.

The principal problem in stratifying location is to avoid confounding two or more factors. For example, if the location on the deep clay has 150 inches of rainfall, the location on the shallow clay has only 50 inches of rainfall, and the location of the sandy has 100 inches of rainfall, it would be impossible to determine whether the differences which appear in the forest planting were due to rainfall or soil.

The solution, when possible, is to select locations on all soil types or in regions of the same rainfall on each of the soil types. The second is usually easier to do in practice.

*Plot replication.* Obtained by establishing at least two plots of each species on each sits at each location. For example, in Valley X plant at least two plots of *E. saligna,* at least two plots of *E. robusta,* and at least two plots of *E. citriodora.* This is not to imply that two is the most desirable number of replications; two is the minimum.

It is worth pointing out that in adaptability studies (by which I mean the preliminary screening of many species about which little is known of their silvical characteristics or their site requirements) comparisons between individual trees are much more sensitive, and more economical to obtain, than comparisons between pure groups of trees.

*Time replication.* Obtained by establishing plantings in at least two different planting seasons. The importance of replication in time depends on the year variability in weather in the planting area. Almost anywhere the weather varies sufficiently between years to cast doubts on conclusions based on plantings made in one year only; in extreme cases the most casual observer may recognize that the results were influenced by abnormal weather conditions.

*Randomization*

The third essential step is randomization. This is usually accomplished by random assignment of the position of the individual plots on each site at each location.

Complete randomization of plots frequently results in an uneven distribution and misleading results; therefore, plots are normally grouped into blocks occupying a relatively uniform area, each block containing a fixed number of plots of each species.

For example, when a number of plots are to be situated on a slope it can usually be assumed that the lower slope is a more productive site than the upper slope. Thus one block containing randomly arranged plots of all species should be established on the lower slope, and another block of randomly assigned plots of all species should be placed on the upper slope.

The one thing to be carefully avoided at this stage is assigning plots on the basis of species requirements. That is, even if you feel sure that *E. saligna* demands a better site than *E. robusta,* never assign the good plots to the *E. saligna* and the poor plots to the *E. robusta.* Such action will completely prevent a true comparison of the adaptability or productivity of the two species, and thus will invalidate the study.

There are a number of systematic arrangements of plots within blocks in wide use. Unfortunately, they can be used safely only by a qualified statistician throughly familiar with local conditions. Therefore, they will not be discussed here.

## EXAMPLE

For clarification, let us consider the species introduction program in Puerto Rico.

### *Limits of the Problem*

We assume that several hundred thousand acres of deep mountain clays now supporting coffee plantations will be released from cultivation during the next two decades. Although many native species are present as individual trees or coffe shade, the natural forests have been virtually eliminated. Therefore, we are interested in learning what species are best adapted to this major site.

Because of high establishment cost we are only interested in planting species of super-vigor, super-quality, or which yield a product not now available.

Furthermore, because of the profuse and rapid growth of grass, weeds, and vines, and the relatively high cost of labor, we are interested only in species with very rapid initial growth. Our absolute minimum is a height of six feet in two years. Because such growth is believed to require full sunlight, we are primarily interested in species which can be established by planting in the open.

Considering all these things, then, our initial problem is this: what species of super-quality, super-vigor, or with a new product will make very rapid initial growth, will develop satisfactorily when planted in open fields, and will grow to merchantable size on the deep mountain clays?

For simplification, only four of the species under trial will be mentioned :

| | |
|---|---|
| Eucalyptus (*Eucalyptus patentinervis* R. T. Baker) | very high vigor, low marketability |
| Cadam (*Anthocephalus caddmba* Miq.) | high vigor, commercial veneer |
| Pine (*Pinus caribaea* Morelet) | high vigor, general utility plus long fiber |
| Primavera (*Cyibistax donnell-smithii* (*rose*) Seibert | high vigor, face veneer |

*Replications*

Literature and local experience suggest only very general ideas of the relative vigor and quality of each of the species over the range of rainfall and elevations within the deep mountain clay region; therefore, we must establish comparative planting trials.

*Location replication.* Within the region we recognize eight locations which cover the entire range of elevations, are geographically dispersed, cover the range of rainfall encountered, and represent certain other site factors, such as eroded soils and dry and moist microsites.

*Plot replication.* At each location, each year, we establish sixteen rows of plantings along the contour. Each of these rows may be considered a block, or data from two or more contiguous rows may be considered as one block at the time of analysis.

Each row is divided into three plots, each of which contains one seedling of each species under test. We never test fewer than three species in a plot and have not tested more than sixteen.

Since each plot has one tree of each species, there are three plots in a row, and there are sixteen rows, a total of three times sixteen, or 48 seedlings are required of each species at each location each year. For four species 48x4, or 192, seedlings are required.

*Replication in time.* When seed procurement exigencies permit we normally test a species in three calendar years.

*General Considerations*

Spacing is 2.5 x 2.5 meters; width of planting along the contour for four species is 4x3x2.5, or 30 meters, and length up and down the slope is 16x2.5, or 40 meters. This is a total of 30x40, or 1200 square meters per location.

For eight locations, number of seedlings is 192x8, or 1,536, and the total area occupied is 1200x8, or 9600 square meters. Thus it is obvious that this design permits us to test four species at eight location on less than one hectare of land.

This economy of area under test permits us to concentrate on intensive techniques which assure high initial survival and good early care, which would be impossible for us if we used large plots, each with a single species.

In addition, of course, we have a more sensitive comparison between species than could be obtained from many, many more seedlings planted in pure plots.

*Randomizations*

As noted previously, each plot contains one seedling of each species. We randomize by drawing numbered wooden blocks to decide species location within each plot.


EXPERIENCES DE PLANTATION DE VALEUR STATISTIQUE

*Résumé*

La note met l'accent sur les règles à suivre pour établir des plantations d'essai valables statistiquement. Ce sont essentiellement:

Restreindre les expériences à une série de conditions bien définies;
Répéter chaque ensemble de conditions qui doivant être expérimentées;
Planter au hasard, en vie d'éliminer une influence systématique ou involontaire.

EXPERIMENTOS DE PLANTACIÓN DE VALOR ESTADISTICO

*Resumen*

Este trabajo expone las reglas para establecer un estudio de plantaciones de valor estadístico, que son:

Restringir los experimentos a un conjunto de condiciones bien definidas;
Repetir cada conjunto de condiciones que deben experimentarse;
Plantar al azar, para eliminar la posibilidad de influencias sistemáticas o involuntárias.

EXPERIÊNCIAS DE PLANTAÇÃO DE VALOR ESTATÍSTICO

*Resumo*

O trabalho expõe as regras para estabelecer um estudo de plantações de valor estatístico, ou sejam:

Restringir as experiências a um conjunto de condições bem definidas;
Repetir cada conjunto de condições a serem experimentadas;
Plantar ao acaso, para eliminar a possibilidade de influências sistemáticas ou involuntárias.