UNIVERSITY OF COPENHAGEN

Københavns Universitet

Introduction to principles in design and evaluation of tree improvement Experiments

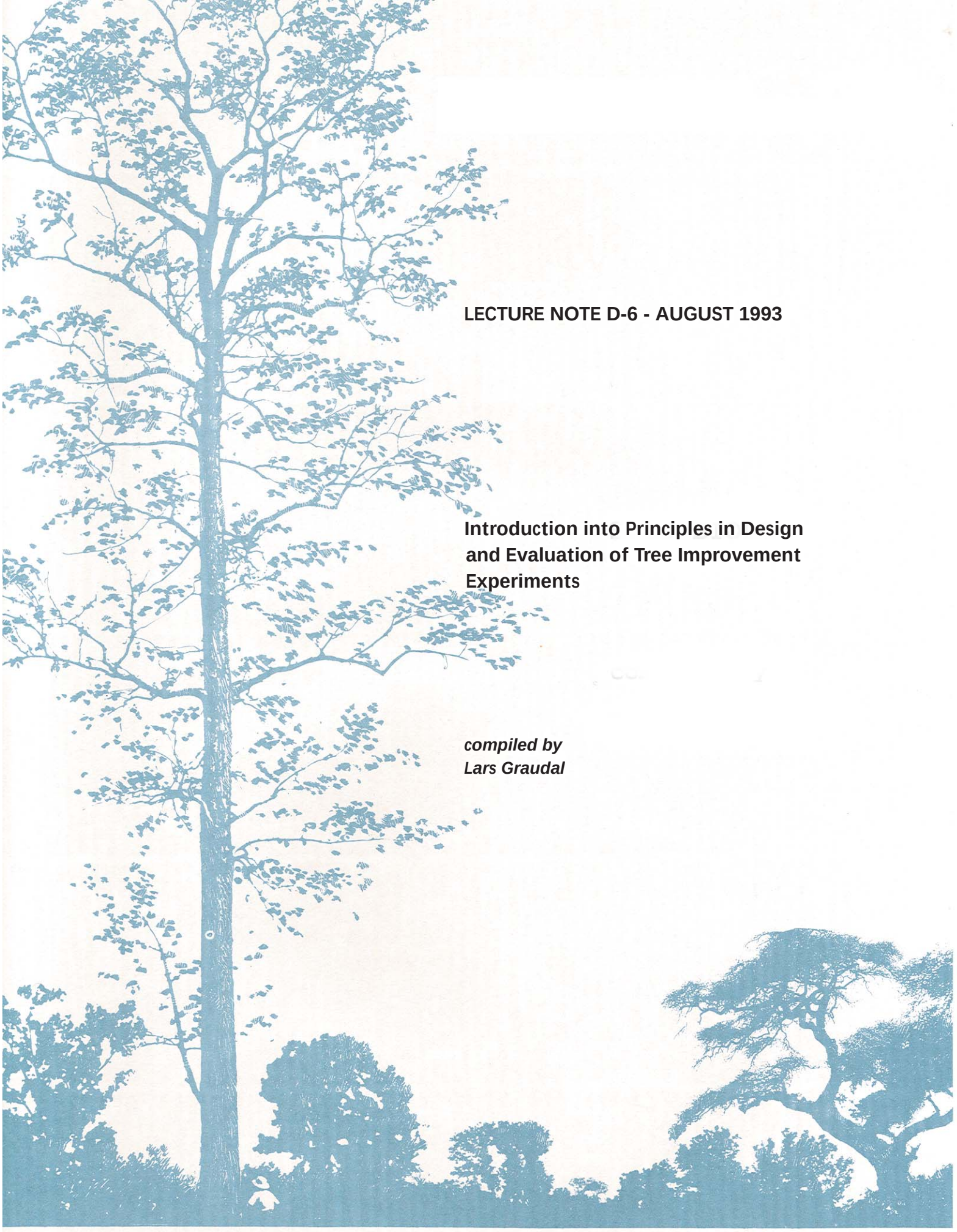
Graudal, Lars Ole Visti

Publication date:
1993

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Graudal, L. O. V. (1993). *Introduction to principles in design and evaluation of tree improvement Experiments*. Danida Forest Seed Centre. Lecture Note D-6

Download date: 17. Nov. 2026



LECTURE NOTE D-6 - AUGUST 1993

Introduction into Principles in Design and Evaluation of Tree Improvement Experiments

compiled by
Lars Graudal

DANIDA FOREST SEED CENTRE

Titel
Introduction to principles in design and evaluation of tree improvement experiments

Authors
Lars Graudal

Publisher
Danida Forest Seed Centre

Series - title and no.
Lecture Note D-6

DTP
Melita Jørgensen

Citation
Graudal, L. 1993. Introduction to principles in design and evaluation of tree improvement experiments. Lecture Note D-6. Danida Forest Seed Centre, Humlebaek, Denmark.

Citation allowed with clear source indication
Written permission is required if you wish to use *Forest & Landscape's* name and/or any part of this report for sales and advertising purposes.

The report is available free of charge
SL-International@kvl.dk

Electronic Version
www.SL.kvl.dk

CONTENTS

1. INTRODUCTION **1**

2. SOURCES OF VARIATION **2**

3. GENOTYPIC VALUES AND HERITABILITY **4**

4. PLANNING AND DESIGN OF EXPERIMENTS **5**

4.1 Planning of experiments 5

4.2 Elements of experimental design and common statistical terminology 5

4.3 Simple statistical models and field trials 7

4.4 More complicated statistical designs and models 9

4.5 Experimental layout 11

5. EVALUATION **11**

5.1 Assessment 11

5.2 Statistical analysis 12

6. LITERATURE AND COMPUTER SYSTEMS FOR DATA ANALYSIS **15**

7. REFERENCES **16**

APPENDIX:

Example of a combined species and provenance trial **19**

1. INTRODUCTION

The present Lecture Note is partly extracted from »A Manual on Species and Provenance Research with particular reference to the Tropics« compiled by Burley & Wood, 1976 (Chapters 4 and 6) and »Standard Research Methods for Multipurpose Trees and Shrubs« edited by MacDicken *et al*., 1991.

The purpose of this note is to highlight some of the basic principles in the design and evaluation of common experiments, such as species and provenance trials and progeny testing. Detailed guidelines for field practice and ana-lytic calculations should be sought in the sources already referred to or elsewhere (see section 6 »Literature and computer systems for data analysis«). The note can be read without any specific knowledge of statistical techniques but the application of the principles will require such knowledge at an elementary level (corresponding e.g. to the short course in Snedecor & Cochran, 7th edition 1980).

A comprehensive appendix gives an example of a combined species and provenance trial, which should help to clarify the various concepts and subjects dealt with in this note.

2. SOURCES OF VARIATION

Genetic improvement consists in the manipulation of existing natural variation.

Phenotypic variation may be a combination of:

(l) Developmental variation (e.g. a mature tree is taller than a young seedling of the same species).

(2) Environmental variation (e.g. a tree on a good site is taller and healthier than one of the same species and age on a poor site).

(3) Genetic variation (e.g. one tree may be taller or straighter than another of the same species and age growing on an identical site).

The task of the tree breeder is to identify the element of genetic variation, which will be passed on from parent to progeny, and to isolate it from the effects of development and environment.

The basis of tree improvement lies in the possibility of selecting trees which possess heritable, rather than non-heritable, superiority in desirable characteristics (ref. Lecture Note no. A-3).

The developmental variation is easily eliminated in an experiment by choosing plants of the same age in the layout. Developmental variation is therefore not discussed further in this note. (If, however, trials of different age are to be compared, it may be done through genetic correlation analysis (see Burdon 1977)). Environmental variation can be divided into three types defined by the level at which the variation is found:

- Variation between sites (e.g. the same provenance will usually behave differently at different latitudes and altitudes).
- Macro variation within a site (there may for example be a fertility gradient from one end of a field to the other).
- Micro variation within a site (induced by the fact that non-systematic differences in for example availability of nutrients often occur even within very small distances on the same site).

Thus the environmental variation cannot be eliminated completely, but only reduced by planting the trees to be compared on the same sites. This will be discussed in more detail later (see section 4.3).

In order to separate environmental and genetic variation it is necessary first to identify the variation in total. As it is impossible to have all the variation of e.g. a population represented in one trial, the researcher must select one or more samples of the total variation to represent the whole population both in terms of genetic composition and of the different environments where the species grow. By doing so, another source of variation is introduced:

(4) The sampling variation, which consists of both environmental and genetic variation.

Sampling is covered by e.g. Palmberg (1980) and Turnbull (1975). Initially, sampling is done throughout the range of the species. Taxonomic and botanical exploration should ideally precede collection, but exploration and preliminary collection will often have to be combined. The number of provenances (seed sources) to be sampled will depend on the extent and the heterogeneity of the distribution range and on the genetic diversity of the species. 10-30 well distributed provenances are considered preferable, including both provenances from areas with optimal development and from marginal sites. From each provenance, seed should be collected from 25-50 'average or not less than average' trees of dominant or co-dominant status in 'normal' stands.

Stands should be of such an age that seed is produced in large quantities, because of practical as well as genetic considerations.

In the case of introduced species, seed should whenever possible be collected from stands of known origin.

In order to avoid trees which are closely related genetically (or trees which have an abnormally high incidence of self-pollination), it is generally con-sidered that seed trees in natural stands should be between 100 and 300 m from each other. Adjacent trees in plantations are usually not closely related because the seed has been bulked and there need be no restrictions on spacing of seed trees.

A fifth source of variation is:

(5) The experimental error, arising from inaccuracies of assessment.

The sampling variation and the experimental error combined are usually termed the residual variation, i.e. the variation that is left when the genetic and the environmental variation have been identified (excluding of course environ-mental and genetic variation arising from sampling, cf. above).

A last source of variation is the systematic or »grave« error, which is that of undetected mistakes in the management of an experiment, such as planting trees of the wrong provenance in a given plot. This type of error cannot be taken into consideration by a statistical model and would be mixed up with the residual variation. Great effort should be made to avoid such errors through careful registration and control throughout the whole experimental process (ref. Lecture Note D-5).

3. GENOTYPIC VALUES AND HERITABILITY

In tree improvement experiments we observe and compare different characters, e.g. height growth. The observed value of a tree is the phenotypic value. The phenotypic value is composed of a genetic element and an environmental effect:

$$P = G + E,$$

where

P is the phenotypic value,
G is the genotypic value, and
E is the environmental deviation.

Likewise, the total variation in phenotypic values is composed of genetic variation and environmental variation:

$$V_P = V_G + V_E,$$

where

V_P is the phenotypic variance between individuals,
V_G is the genetic variance caused by genetic variation
V_E is the environmental variance caused by environmental variation.

Variance is a numeric measure of variability. Variance is defined as the sum of the squared deviations of each observation from the arithmetical mean divided by one less than the total number of observations.

The proportion of the total phenotypic variance caused by genetics is called the broad-sense heritability:

$$h^2_{BS} = \frac{V_G}{V_P} = \frac{V_G}{V_G + V_E}$$

(Ref. Lecture Note D-2 for more details on phenotypic and genotypic values, and heritability).

When comparing the productivity of e.g. different provenances in a tree improvement experiment, it can be done by calculating the phenotypic deviation of each provenance from the overall mean in the experiment. To achieve a more true picture of the productivity ranking of the provenances, the genotypic values can be estimated by multiplying the phenotypic deviations with the broad-sense heritability (see e.g. Keiding, Wellendorf and Lauridsen, 1986, pp 21-24 for an example).

These »more true« deviations are smaller, and is therefore also called the »shrinking« factor.

4. PLANNING AND DESIGN OF EXPERIMENTS

4.1 Planning of experiments

An experiment usually consists of various phases, for example:

- problem formulation and analysis
- defining objective(s)
- detailed planning of the experiment
- implementation (layout, assessment, management)
- statistical analysis
- conclusions

Identification and analysis of a problem should help us in defining the objective and the most suitable design of an experiment to achieve our objective and help to solve the problem.

If our problem is lack of fuelwood and fodder in a given area, we may want to identify fast growing species producing both fuelwood and fodder, and acceptable to the potential users. If good fast growing species are already known in our area, we may want to increase production by identifying the most productive provenances or even individuals. If, for instance, we know that *Acacia tortilis* is a good producer of both fuelwood and fodder, we may assume that some provenances are better than others. Such a tentative assumption is called a hypothesis. To test this hypothesis, we would plan our experiment with the objective of determining the relative productivity of total wood biomass, leaf and pod production by a number of provenances of *Acacia tortilis* on one or several sites under a specified management regime.

Defining the objective(s) is the most crucial step in any research process. It should clearly specify the limits of the population(s) being studied, and thus also the limits of the results being achieved. Care should be taken not to extrapolate the results to include e.g. age classes or site conditions which are not covered by the experiment(s).

4.2 Elements of experimental design and common statistical terminology

The statistical methods that are used in species, provenance, progeny and clone trials are the same as used in other technical fields. To understand the literature on tree improvement experiments it is necessary to be familiar with the elements of experimental design and the common statistical terminology, i.e. what is meant by concepts and words like treatment, experimental unit, comparative trials, variable, replication and randomization.

In experiments various treatments are compared. The word treatment should be understood in its broadest sense, i.e. comprising different silvicultural methods as well as different plant materials that are tested and compared. In a growth trial one may for example compare different types of manure or different provenances. In the first case the treatments are types of manure, and in the latter the provenances. In tree improvement trials the treatments are typically species, provenances, progenies or clones, or combinations like species/provenance or provenance/progeny.

To compare treatments, they are applied to what is called experimental units. An experimental unit may be a plot in a field trial. In a provenance trial, individuals belonging to the same provenance (i.e. »one treatment«) are planted in the same plot. In one experiment there is usually a fixed number of individuals per plot, and several plots receive the same treatment (replications).

In so-called comparative trials where different treatments are compared, these treatments are applied to experimental units, as identical as possible, to ensure that differences observed are due to the treatments and not to other factors.

The observed response or the observed variable values in a trial are the results of the assessment of the experimental units. A variable is a quantity or quality able to assume different values (e.g. height, straightness). In field trials the observed variable value in a plot will normally be the mean value of assessed individuals in the plot. A provenance may for example be represented by 16 individuals in each plot. When assessing e.g. the height of the provenance we may measure all 16 trees or a sample of them and calculate the mean value, which will be the observed variable value for height in the plot.

There will always, in biological experiments, be differences between the observed variables of experimental units that have received the same treatment (e.g. between four plots with the same provenance in one experiment). It is important that these differences which are caused by residual effects (ref. above) can be estimated. This requires replication of the treatments and a suitable distribution of the treatments on the experimental units. The most common way of making this distribution is by randomization (see below).

Increasing the number of plots (replication) of each treatment steadily reduces the residual variation associated with any comparison of different treatment means. Theoretically, the precision of such comparisons is proportional to the square root of the number of replications; in practice, however, there are other factors which slightly reduce the gain in precision. Chief among these is the fact that for any given plot size and number of treatments increased replication means a larger experiment and, consequently, a greater likelihood of site heterogeneity.

If it is assumed that residual plot-to-plot variation is random, there will always be relatively extreme plots. Some plots could be relatively fertile, perhaps in combination with a specific treatment, in particular in experiments comparing silvicultural treatments, while other plots are relatively infertile. Assignment of each treatment (e.g. provenance) to the plots at random is the best way to avoid bias in the estimation of differences between treatments (e.g. that a particular

provenance is planted on all »good« or all »bad« plots). In addition, random assignment of the treatment ensures an unbiased estimate of residual variation. Statistical methods based on independent random distri-butions may thus be used.

Randomization gives each treatment an equal chance of being alotted to any plot that may happen to give an unusually good or unusually poor response. Randomization can be done by giving each treatment a lot and then draw the lots at random, as in a lottery. Another more commonly used method is to use tables of random digits (see e.g. Snedecor and Cochran 1980).

Other important points concerning experiments are that they should be **simple** and **representative**. Simplicity promotes accuracy by minimising the possibility of mistakes at all stages of the experiment. Representation is decisive in how one may generalize the results, i.e. which conclusions that can be drawn from the experiment. In growth trials, it is thus important that the experimental sites are representative of intended plantation areas, and in provenance trials the plants must be representative and reflect as truly as possible provenances under test.

4.3 Simple statistical models and field trials

An experimental design is described by a model of variation. The model should represent the way in which an observed response or variable (e.g. height growth) is affected by each source of variation. For example, the response may be assumed to be an additive combination of genetic, environmental and residual effects:

Observed response	=	Common mean	+	Genetic effect	+	Environmental effect (Block)	+	Residual effect [1]

Models of this form constitute the basis of the analysis of experiments. If an experimental site was relatively uniform with only random variation from plot to plot, then a **completely randomized design** of a limited number of plots (replicates) of each of the treatments would be an efficient arrangement. The model would be:

Observed response	=	Common mean	+	Genetic effect	+	Residual effect [2]

Such an ideal site is rare. Field trials of forest trees cover a large area which inevitably includes variations in soil, micro-climate, topography, aspect, site history (i.e. previous vegetation and land use) and other factors.

The variation is often composed of (1) a gradient that is effective over larger areas and (2) local micro-variations. If there is a gradient for example as shown in below (a higher fertility towards east than towards west), the field should be divided into blocks that represent the different degrees of fertility. Blocks should be established to minimise within-block variation. In order that a blocking scheme can reduce experimental variability effectively, the field layout must suit the unique characteristics of the site.

[1] (or with symbols $X_{ij} = \mu + \alpha_i + ß_j + \varepsilon_{ij}$,
where $\{\varepsilon_{ij}\}$ is assumed to be normally distributed with zero mean (ε_{ij} N(0, σ^2)), i = 1,...,n_a and j = 1,...,n_b,
where n_a is the number of treatments and n_b is the number of blocks.

[2] (or with symbols: $X_{ij} = \mu + \alpha_i + \varepsilon_{ij}$, ε_{ij} N(0,σ^2), i = 1,..., n_a and j = 1,..., n_b,
where n_a is the number of treatments and n_b is the number of replications.

On a fertility gradient (as in the figure), the block is usually rectangular with the long axis along the contour of equal fertility. On level, apparently homogeneous ground, conditions usually vary with distance in any direction, so a block is square.

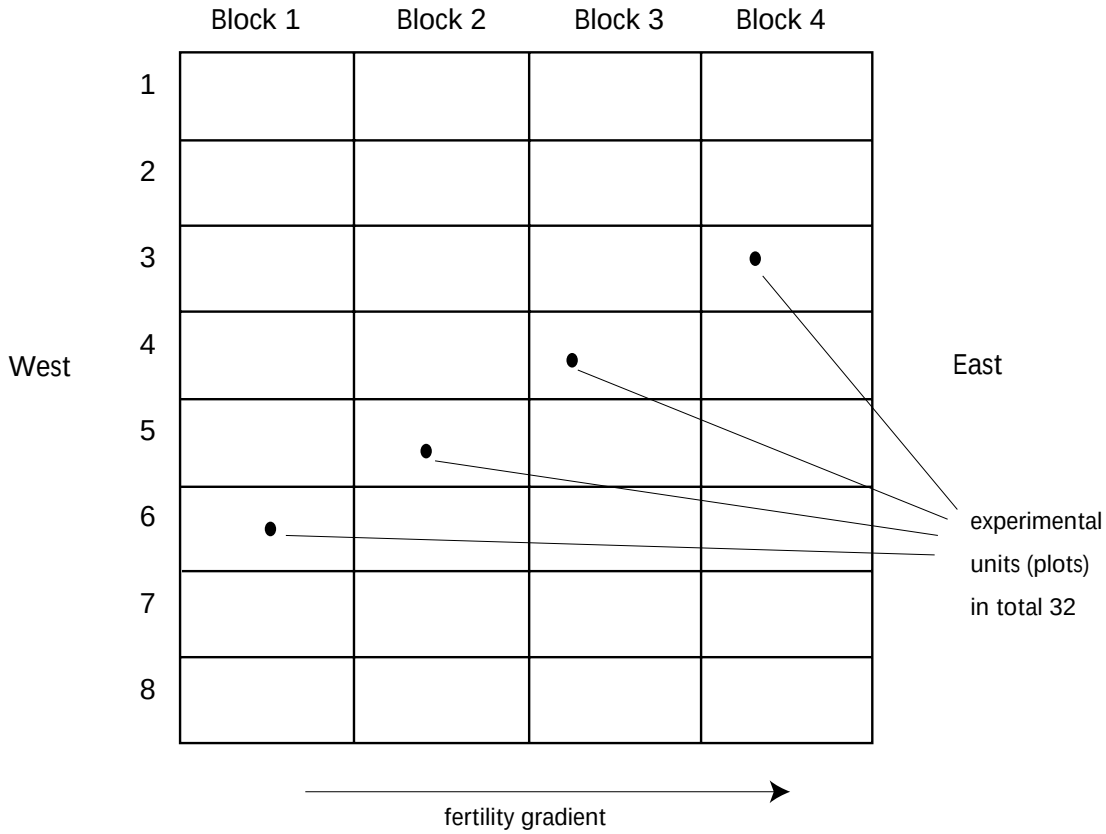


Figure 1. Example of field divided into blocks and plots

In the figure the experimental site is divided into a number of rectangular blocks of equal area, and each treatment (in the example 8) is represented once in each block (by a plot of one to several trees). Thus each block contains a complete replication of the treatments to be compared. The treatments are allocated at random to the plots within each block.

Such an arrangement is called a randomized complete block design, and the statistical model is the one presented at the beginning of this section. The environmental effect is more often called the block effect to make it clear that it represents only the macro-variation at site level (cf. section 2).

It should be noted that blocks are chosen to make the variation within blocks as small as possible, whereas plots are chosen to make the variation within plots as large as possible.

This is done to have as much of the macro site variation expressed by the block effect and as much of the micro site variation eliminated in the plot mean values.

Plot shape should be influenced by knowledge of soil gradients. If a gradient exists in one direction, it is sensible to use plots designed to maximize the variation

within the plot and thus minimize the variation between plots. When testing trees in blocks where soil variability is patchy, square plots are useful.

Choice of plot size depends on several aspects. As the plot size is increased within a certain range, the variability for any plot mean or total will be reduced as the small fluctuations are evened up. However, when the plots become too large, the site variation from plot to plot increases and the costs of establishing, maintaining, and monitoring the experiment will increase. In general, more replications are recommended for smaller plots.

Generally, several smaller plots will give the experiment better precision as long as the plot represents the system being tested. Very small plots may be too small to provide a realistic environment for the treatment being tested. Single-tree plots can be used when low competition is expected.

4.4 More complicated statistical designs and models

A field experiment is always more or less unique and the actual design should in each case be adapted to the problem to be solved, the planting material available, and the experimental sites. This may require more complex designs than discussed so far. Below is given some indication of more complex situations and possible solutions. For details refer to references (e.g. Matheson 1990, cf. section 6, p.15).

Incomplete block designs

The well known types of design described above are both simple and efficient where the number of treatments to be compared is small. If the number of treatments increases, these simpler designs often become less efficient be-cause both random and systematic variation within replicates tend to increase.
The main approach to reducing within block heterogeneity is to arrange plots in blocks which are smaller than a complete replicate of all treatments, leading to . A commonly used incomplete block design is the lattice design (see e.g. either Burley & Wood 1976 or Matheson 1990).

Factorial experiments

If the experiment should reveal the effects of more than one factor, it may complicate the model.

One common example is to have the trial repeated at several sites. The model would then typically be built as follows:

Observed response	=	Common mean	+	Genetic effect	+	Site effect	+	Block effect	+	Residual effect[3]

The analysis of an experiment based on this model would thus indicate differences between both treatments (genetic effect) and planting sites (variation between sites - cf. section 1, p.3). Further complication will occur if there is in-

[3] (or with symbols: $X_{ijk} = \mu + \alpha_i + ß_j + \gamma_k + \varepsilon_{ijk}$, ε_{ijk} $N(0, \sigma^2)$

teraction between treatment and site (i.e. the behaviour of provenances varies with site). The factors, treatment and site, in an experiment like this is said to be crossed because each treatment is repeated at each site.

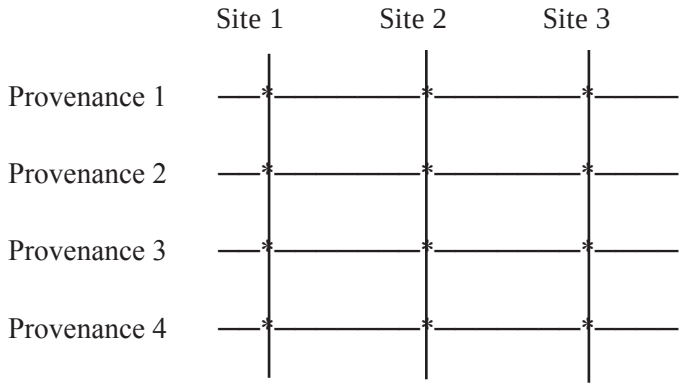


Figure 2. Example of crossed factors, provenance and site.

Another example is combined provenance/progeny trials, where each provenance is represented by several families. Here each family is a subsample of a specific provenance and the factors - provenance and family - are said to be hierarchal or nested, which will require yet another model, design and analysis.

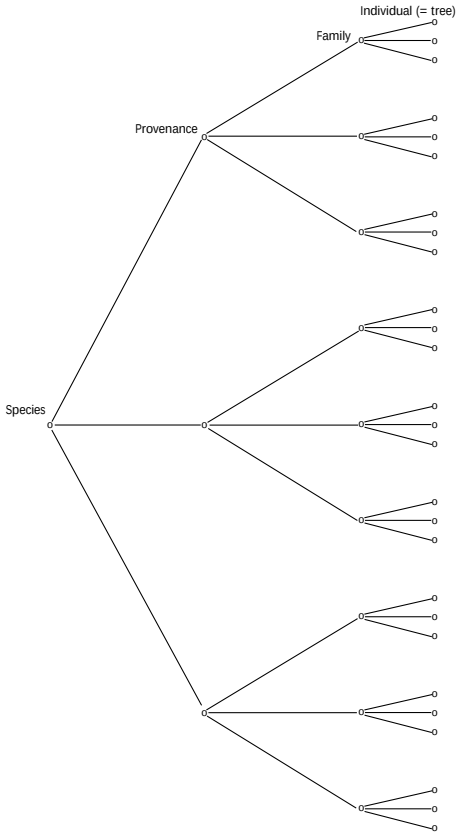


Figure 3. Example of different nested levels (provenance, family, individual). The factors 'provenance', 'family' and 'individual' show themselves with 3, 9 and 27 levels respectively. Within the factor 'provenance' three levels occur of the factor 'family'. Within the factor 'family', three levels occur of the factor 'individual'. The 'individual' is said to be nested within the 'family', 'family' within the 'provenance' and 'provenance' within the 'species'. Note that the word 'level' is used in two ways (cf. the figure): horizontally - species, provenance, family, and individual - and vertically -within each factor.

4.5 Experimental layout

Planning and implementing the layout of the trials should be given high priority. A trial should be laid out carefully and precisely according to a field plan designed in advance.

An experimental site should be closely examined before the experimental layout (field plan) is allocated to the area in order that each block should be as homogeneous as possible: soil type and depth, slope, aspect, exposure, previous land use and natural fertility.

The experimental site should be uniformly prepared according to management prescription and then the entire trial should be marked out before planting begins.

For details describing the practical aspects of field testing refer to Lecture Note D-5, Field Testing Practices of a Tree Improvement Programme.

5. EVALUATION

5.1 Assessment

Assessment of trials is often lengthy and expensive. The characters to be assessed will always depend on the objective of the trial. Usually, only those characters which are practically important and in which there is substantial experimental variation should be assessed. Quantitative traits are generally better than qualitative ones for subsequent analysis and interpretation and are often easier to assess objectively.

A detailed account of measurement options is given in MacDicken *et. al.* (1991).

Any measurement is subject to the possibility of errors. Errors may be attributed to the precision of the instrument, the care or the accuracy with which measurements are taken, and they may be either random or systematic.

Measurement instruments and precision

A measuring device should be chosen according to the trait to be assessed, the conditions under which it is used, and the required precision of assessment.

Precision refers to the exactness of measurement. Precision of measurement equals one-half of the unit of measurement. This is, if measurements are made to the nearest centimetre (cm), precision is one-half cm. If measurements are made to the nearest metre (m), precision is one-half m.

Measuring rods with a precision of one or two centimetres may be necessary in order to identify small differences in height among young trees. However, with large stems, where small differences have little practical significance, the convenience of a clinometer or similar optical device is much more suitable.

Assessment and recording should not exceed the real precision of an instrument (e.g. by unreliable interpolation).

Accuracy

Instruments should be used correctly in order to avoid bias or unnecessarily inaccurate observations.

Accuracy refers to correctness of measurement. It is determined by two factors: consistency and reality. If repeated measurements of the same object vary, they cannot all be accurate. As an example, with a pair of callipers it is possible to measure a tree's diameter at breast height (dbh) to 1/10 mm. But if the calliper is moved slightly, the new dbh measurement of the same tree will be different because of irregularities in stem roundness, bark thickness, and other variables. The precision is 1/10 mm divided by 2 = 1/20 mm, but as all the readings are not the same, not all can be considered equally accurate. In such a case, the precision may be illusory.

The second factor is the relationship to reality. If one needs to measure dbh of a tree, breast height must be determined. This is often and reliably done with a stick cut to a length of 1.3 m. But if the base of the stick is sliced at an angle, the sharp point may slide through loose litter, sink into soft ground, sit exactly on hard ground, or rest on a raised root or stone. In such a case, only the measurements taken when the stick rests on hard ground can be accurate -- and then only until the point of the stick wears down or breaks off.

Random and systematic errors

Accidental or careless errors of measurement are not included in the above discussion of accuracy and precision. Purely random errors tend to follow the rules of any random population. Positive and negative errors occur with equal frequency, and small errors occur more often than large ones. Thus, the errors tend to even up.

For an individual tree (the basic unit of genotype selection), errors from single mea-surements invariably produce wrong answers, whereas for a plot of many trees mean values will usually be more correct because the random errors tend to even up.

Many errors, however, are not random in the technical sense, but systematic. Foresters who fail to measure the location of dbh commonly estimate the location of 1.3 m too high in the morning and gradually lower as they tire during the day. These dbh measurements are too small for plots measured in the morning and too large for plots in the afternoon. These are not compensating errors.

5.2 Statistical analysis

Data from a trial should be analyzed as soon as possible. The result of each analysis may suggest changes in the planning of silvicultural treatments and assessment of the trial.

Many different types and stages of analysis are possible but the methods of analysis should be appropriate to the objectives of the experiment, to its design and to the traits being analyzed.

Most commonly in species and provenance trials, the analysis of variance is used to measure the relative importance of variation between treatments (genetic effects), systematic environmental variation and residual variation, and to estimate the accuracy with which treatment differences are calculated.

If we measure several traits of each treatment, we may wish to investigate how these traits vary together, e.g. to find out if there is linear relationship. This can be done by regression and correlation analysis, or more complicated so-called multivariate analysis, which can be used to investigate the ways in which the treatments differ in terms of all their measured traits, taking their interdependency into consideration.

There are also methods for analyzing the effects of genotype by environment interactions.

Many of these methods are complex and laborious. It is always desirable to carry out some simple, preliminary analysis of the data, such as calculating and comparing plot and treatment mean values, and make graphic presentations.

It is essential that, whatever form of analysis is carried out, errors are avoided which could lead to incorrect conclusions and, consequently, a wasted experiment or, even worse, misleading recommendations. All calculations should be carried out with sufficient accuracy to ensure that practical differences can be identified and to avoid rounding errors. Analyses carried out manually should always be carefully checked and print-outs of computer analyses should not be accepted without ensuring that the programme is reliable and the data are entered correctly.

Summary information about preliminary analysis and analysis of variance is given below. For details consult references, see section 6.

Preliminary analysis

The simplest form of analysis, which should be carried out as soon as possible, is the calculation of plot mean values. A plot mean is simply the arithmetic average, i.e. the sum of all measurements of a trait, such as tree height, divided by the number of measurements.

Plot means may be used to estimate treatment means and differences. These preliminary estimates show the way the crop is developing and should reveal any

substantial differences between plots and provide

- estimated differences between treatments and between environments.

Analysis of variance

The analysis of variance in individual traits is the most widely used method for interpreting the results of experiments. The analysis provides the following information:

(i) estimates of the relative magnitude or importance of each identifiable source of variation (e.g. provenance, site, block, residual);

(ii) tests of statistical significance of variances and differences;

(iii) indications of the accuracy of estimated differences by means of their standard errors.

re (i): Although the precise form of the analysis of variance varies according to the experimental design and the underlying statistical model, the essential principle is to estimate the variance attributable to each source (provenance, site, block, residual, etc.) of variation. The analysis of variance will thus tell us which of the sources are the more important, i.e. the more decisive for the differences observed.

re (ii): When analyzing our data, the primary question of interest will usually be whether there is an effect of the treatments. Using statistical methods, this question is answered by finding out how likely it is that our observed variables should have happened by chance. This is what we call a test of statistical significance. Such tests are based on theory of probability and sampling.

The objective of a provenance trial is to find out if some provenances are better than others. To do so we compare our experimental data summarized as a set of mean values, one for each provenance. As mentioned earlier there will always be differences between the observed variables in biological experiments.

If we assume that the experimental site is homogeneous, the question would in this case be whether these differences are due to genetic differences between the provenances or the result of the residual effect. To perform the statistical test, we assume that there are no significant differences between the mean values of the provenances. This is what we call a statistical hypothesis. We also assume that the mean values of the provenances follow a certain frequency distribution called the normal distribution.

If these assumptions are true, there will be a certain probability of having observed our experimental data by pure chance. If this probability is low,

we reject our hypothesis and conclude that there are statistically significant differences between the provenances. The lower the probability, the more certain is our rejection of the hypothesis.

re (iii): The standard error or the standard deviation is the square root of the variance around the mean (cf. section 3). Extraction of the square root of the variance recovers the original scale of measurement and provides an indication of the accuracy and statistical significance of estimated treatment (and/or environmental) differences.

6. LITERATURE AND COMPUTER SYSTEMS FOR DATA ANALYSIS

To understand and apply the experimental methods introduced in this note some basic knowledge of sampling techniques, probability distribution and tests of hypotheses is required. A classic introduction to these subjects and further statistical analysis is given by Snedecor & Cochran (1980,7.ed.).

A more advanced textbook on experimental design is Cochran & Cox (1957).

Concerning species and provenance research, the manual by Burley & Wood (1976) and the book by MacDicken *et al.* (1991) give details and practical guidelines.

Practical guidelines are also given in Lecture Note D-5, and basic textbooks in tree improvement also include more or less detailed accounts of conducting experiments (Wright 1976, FAO 1985, Zobel & Talbert 1984).

A brief but more in-depth introduction to designing tree improvement experiments will be found in Matheson (1990).

Today, many analyses are done by electronic data processing. A number of standard programmes for statistical analysis exists, both for advanced pocket calculators, personal computers and larger computers. Statistical Analysis System (SAS) is among the more well known programmes for PCs and larger computers, cf. e.g. PC SAS Introductory Guide, but many other systems are available (SPSS/PC+, Statgraphics).

7. REFERENCES

Burley, J. and P.J. Wood (eds.) 1976.
A Manual on Species and Provenance Research with particular reference to the tropics. Tropical Forestry Papers No. 10 and 10A. University of Oxford.

Burdon, R.D. 1977.
Genetic Correlation as a Concept for Studying Genotype-Environment Interaction in Forest Tree Breeding. Silvae Genetica 26, 5-6

Cochran, W.G. and G.M. Cox 1957.
Experimental Designs. Second edition. John Wiley & Sons.

FAO 1985
Forest Tree Improvement. Report of the FAO/DANIDA Training Course on Forest Tree Improvement, Merida, Venezuela, 1980. FAO Forestry Paper 20. Rome.

Keiding, H. 1992
Field Testing Practices of a Tree Improvement Programmme. Lecture Note No. D-5. Danida Forest Seed Centre.

Keiding, H., H. Wellendorf and E.B. Lauridsen 1986
Evaluation of an International Series of Teak Provenance Trials. Danida Forest Seed Centre.

MacDicken, K.G., G.V. Wolf and C.B. Briscoe (eds.) 1991
Standard Research Methods for Multipurpose Trees and Shrubs. Winrock International Institute for Agricultural Development.

Matheson, A.C. 1990
Designing Experiments for MPT Genotype Evaluations. In: Tree Improvement of Multipurpose Species, edited by Nancy Glover and Norma Adams. Multipurpose Tree Species Network Technical Series, Vol. 2. Winrock International Institute for Agricultural Development.

Palmberg, C. 1980
Sampling in Seed Collection. In: FAO, 1985.

SAS Institute Inc. 1985
SAS Introductory Guide for Personal Computers, Version 6 Edition. Cary, NC, USA.

Snedecor, G.W. and W.G. Cochran 1980
Statistical Methods, Seventh Edition. The Iowa State University Press

Turnbull, J.W. 1975
Seed Collection - Sampling Considerations and Collection Techniques. In: FAO/DANIDA Training Course on Forest Seed Collection and Handling, Chiang Mai, Thailand.

Wellendorf, H. and B. Ditlevsen 1992
Introduction to Forest Genetics. Lecture Note No. D-2. Danida Forest Seed Centre.

Willan, R.L., Kirsten Olesen and H. Barner 1989
Natural Variation as a Basis for Tree Improvement. Lecture Note No. A-3. Danida Forest Seed Centre.

Wright, J.W. 1976
Introduction to Forest Genetics. Academic Press.

Zobel, B.J. and J.T. Talbert. 1984.
Applied Forest Tree Improvement. John Wiley & Sons.

APPENDIX

Example of a Combined Species and Provenance Trial

CONTENTS

INTRODUCTION	18
THE TRIAL	19
PURPOSE OF THE TRIAL	19
TREATMENTS BEING TESTED	19
DESCRIPTION OF TRIAL SITE	19
TRIAL LAYOUT	21
CHARACTERS ASSESSED	21
PRELIMINARY ANALYSIS	24
STATISTICAL MODEL	31
STATISTICAL HYPOTHESES	40
ANALYSIS OF VARIANCE	41
CONCLUSIONS OF THE ANALYSIS OF VARIANCE	43
FURTHER ANALYSIS AND CALCULATION OF GENOTYPIC VALUES	44
FINAL REMARKS	47
REFERENCES	49

INTRODUCTION

The purpose of this appendix, with an elaborate example, is to provide a basic understanding of the interpretation of the results of statistical analysis of the same type of field trials as described in the Lecture Note, which in most presentations are given in a brief and concise form that require a thorough knowledge of statistical methods to be understood.

The appendix presents an example of an experiment including the purpose of the experiment, the layout in the field, the results of the assessment, the results of the statistical analysis, and the conclusions that can be drawn from the analysis. The statistical analyses comprise (i) preliminary analysis, (ii) the formulation of a statistical model, (iii) the formulation of statistical hyphotheses, and (iv) the results of the different analyses of variance and their interpretation. Finally, genotypic values are calculated.

The arithmetical operations and the theory behind the analyses are not presented. The appendix is thus not a 'cookery book recipe' on statistical analyses, but only a brief 'guide' to a better understanding of the concepts described in the Lecture Note and a help to reading and understanding results of field experiments as they are generally presented.

Paragraphs written with small print are more advanced and primarily meant for the reader with statistical knowledge and should be skipped by other readers. These paragraphs focus in particular on checking the assumptions of statis-tical models, a point which is often given too little emphasis. However, unless the basic assumptions of the models are sufficiently fulfilled, the analysis and conclusions may at best be worthless or at worst, give misleading recommendations. Another point of attention is the question of random versus fixed effects models. The small print paragraphs also provide references and details of analysis that will be of interest for readers using statistical tools themselves.

THE TRIAL

The trial is a combined species and provenance trial established by the Central Arid Zone Research Institute (CAZRI) in Jodhpur, India, in August 1984.

The trial has been regularly assessed by CAZRI. The trial is part of a series of species and provenance trials established on different sites in a number of countries within the framework of the FAO Project on Genetic Resources of Arid and Semi-arid Zone Arboreal Species (cf. e.g. FAO 1988). A selection of these trials has been assessed in collaboration between national institutes, FAO and DFSC. The trial presented is one of these.

The data is a sample from the assessment undertaken in collaboration between CAZRI, FRI (Forest Research Institute, Dehra Dun), FAO and DFSC in March 1991.

PURPOSE OF THE TRIAL

The purpose of the trial is to evaluate the survival and growth of some introduced and local seed sources of different species in the Thar Desert (CAZRI, Jodhpur, India).

TREATMENTS BEING TESTED

The seed sources comprise 8 provenances of *Acacia nilotica*, 3 provenances of *Acacia raddiana*, 2 provenances of *Acacia senegal*, and 3 provenances of *Prosopis cineraria*. There are thus two treatment levels, species and provenance, where provenance is nested within species. The seedlots are listed in table 1.

DESCRIPTION OF TRIAL SITE

Research Farm, Central Arid Zone Research Institute (CAZRI), Jodhpur, Thar Desert, India:

Latitude:	26°3'N	
Longitude	73°6'E	
Altitude:	app. 200 m	
Topography:	Flat, gentle	
Soil:	Sandy loam	
Rainfall:	Annual average 317 mm	1984: app. 240 mm
		1985: app. 210 mm
		1986: app. 250 mm
		1988: app. 240 mm

No.	Species	Seedlot Acc.no.	Country of origin	Provenance name
1	*Acacia nilotica ssp. indica var. cupressiformis*	1081/82	India	Bhiwani (Hissar), Haryana
2	*Acacia nilotica ssp. indica var. cupressiformis*	1083/82	India	Akola, Maharashtra
3	*Acacia nilotica ssp. indica var. vediana*	1084/82	India	Akola, Maharashtra
4	*Acacia nilotica ssp. indica var. jacquemontii*	1070/82	India	Pune, Maharashtra
5	*Acacia nilotica ssp. indica var. jacquemontii*	1069/82	India	Etawah (Mainpuri), U.P.
6	*Acacia nilotica ssp. nilotica*	1068/82	Sudan	Khartoum, White Nile Pune, Maharashtra
7	*Acacia nilotica ssp. indica var. vediana*	1071/82	India	Pune, Maharashtra
8	*Acacia nilotica ssp. indica var. cupressiformis*	1082/82	India	Rao
9	*Acacia raddiana*	1040/82 (82/622/ISRA/CNRF)	Senegal	Keur Mbaye
10	*Acacia raddiana*	1041/82 (82/621/ISRA/CNRF)	Senegal	Ein-Hazeva (Arava)
11	*Acacia raddiana*	1013/81	Israel	Windou Tiengoly
12	*Acacia senegal*	1036/82 (82/558/ISRA/CNRF)	Senegal	Namarel
13	*Acacia senegal*	1035/82 (92/559/ISRA/CNRF)	Senegal	Khanfar (El-Kod), Abyan, original source
14	*Prosopis cineraria*	1062/82	Yemen	Balhaf Bhiwani (Hissar), Haryana
15	*Prosopis cineraria*	1087/82	India	Trichy, (Trichunapally), Tamil Nadu
16	*Prosopis cineraria*	1089/82	India	

Table 1 Seedlots being tested. Provenances are usually identified by their name and/or seedlot accession number

TRIAL LAYOUT

Randomized complete block design consisting of 4 blocks (replications), 16 plots (experimental units) per block, each plot with 49 trees, spaced 3x3 m. Each seedlot (provenance) is thus represented with one plot of 49 trees in each block.

The trial area is rather flat and uniform, so square plots and a regular blocking scheme have been applied.

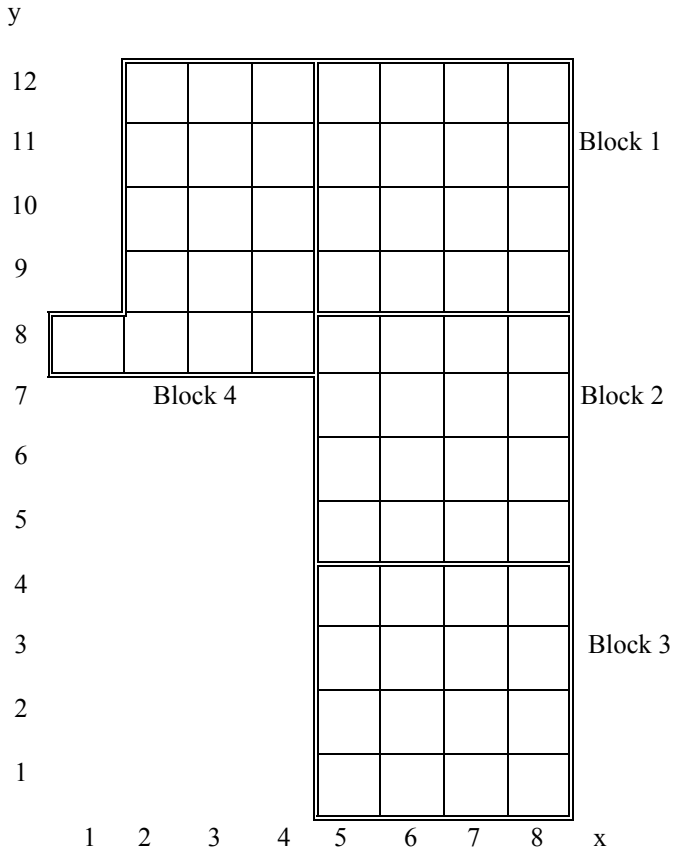
A sketch of the layout is given in figure 1. Plot and individual tree positions are defined by two coordinate systems as shown.

CHARACTERS ASSESSED

Survival and growth were registered in March 1991 (i.e. at age 6). Survival was counted among all the original 49 trees in each plot. Growth was assessed by measuring the height of the inner 25 trees in each plot. The mean value of the height of the trees measured were calculated for each plot. The resulting plot data set is shown in table 2[1].

[1] Other characters were assessed as well, but only survival and height are presented here.

Layout of blocks and plots in the field:



Individual tree positions in each plot:

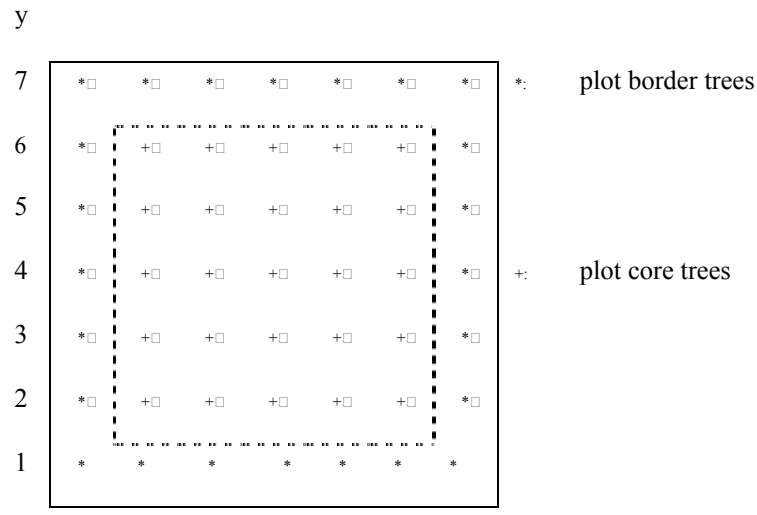


Figure 1: Sketchmap of the trial

Table 2. Plot data set for survival and vertical height (VHEIGHT). Survival is the number of surviving trees among the 49 planted in each plot. Vertical height is the plot mean value of the surviving trees (NVH) in the plot core area (cf. fig.1). Symbol (a to p) is used to identify individual provenances in graphic presentations (figures 8, 9, 12 and 13).

Arid zone species/provenance trials, India
CAZRI/FAO trial established 1984, A.ni/A.ra/A.se/P.ci
Plot data set, survival and height March 1991

BLOCK=1

SPECIES	PROVENANCE	SYMBOL	PLOTX	PLOTY	SURVIVAL	VHEIGHT	NVH
ACNI	1081/82	a	5	12	16	2.95857	7
ACNI	1083/82	b	5	11	6	1.62800	5
ACNI	1084/82	c	8	10	20	2.09500	10
ACNI	1070/82	d	8	9	20	2.04444	9
ACNI	1069/82	e	7	11	12	2.73143	7
ACNI	1068/82	f	7	12	10	2.04500	8
ACNI	1071/82	g	6	11	12	2.21000	8
ACNI	1082/82	h	6	9	3	2.71000	2
ACRA	1040/82	i	5	10	34	2.98059	17
ACRA	1041/82	j	7	10	34	3.05316	19
ACRA	1013/81	k	8	12	31	1.47143	14
ACSE	1036/82	l	7	9	31	1.81615	13
ACSE	1035/82	m	6	12	41	1.51391	23
PCIN	1062/82	n	6	10	0	.	0
PCIN	1087/82	o	8	11	1	0.20000	1
PCIN	1089/82	p	5	9	0	.	0

BLOCK=2

SPECIES	PROVENANCE	SYMBOL	PLOTX	PLOTY	SURVIVAL	VHEIGHT	NVH
ACNI	1081/82	a	8	5	28	3.03769	13
ACNI	1083/82	b	6	8	16	2.52667	6
ACNI	1084/82	c	5	5	5	1.22333	3
ACNI	1070/82	d	5	6	5	2.12750	4
ACNI	1069/82	e	6	7	13	2.86429	7
ACNI	1068/82	f	7	8	3	0.95000	1
ACNI	1071/82	g	7	7	19	2.04778	9
ACNI	1082/82	h	8	8	9	2.81600	5
ACRA	1040/82	i	5	8	23	2.73273	11
ACRA	1041/82	j	8	6	42	3.03526	19
ACRA	1013/81	k	6	6	31	2.16077	13
ACSE	1036/82	l	6	5	35	1.95444	18
ACSE	1035/82	m	7	6	33	1.71077	13
PCIN	1062/82	n	8	7	2	.	0
PCIN	1087/82	o	7	5	7	0.52500	2
PCIN	1089/82	p	5	7	1	.	0

BLOCK=3

SPECIES	PROVENANCE	SYMBOL	PLOTX	PLOTY	SURVIVAL	VHEIGHT	NVH
ACNI	1081/82	a	8	2	20	3.17462	13
ACNI	1083/82	b	6	1	25	1.76000	15
ACNI	1084/82	c	7	2	18	1.61833	12
ACNI	1070/82	d	7	1	15	2.04750	8
ACNI	1069/82	e	8	3	17	2.63182	11
ACNI	1068/82	f	6	3	8	1.02000	3
ACNI	1071/82	g	5	4	7	2.29400	5
ACNI	1082/82	h	5	3	7	2.18750	4
ACRA	1040/82	i	5	1	37	3.07300	20
ACRA	1041/82	j	8	4	35	3.05316	19
ACRA	1013/81	k	7	4	35	2.43600	20
ACSE	1036/82	l	5	2	19	1.14154	13
ACSE	1035/82	m	8	1	29	1.41105	19
PCIN	1062/82	n	6	2	3	0.20000	2
PCIN	1087/82	o	6	4	4	0.35000	4
PCIN	1089/82	p	7	3	0	.	0

BLOCK=4

SPECIES	PROVENANCE	SYMBOL	PLOTX	PLOTY	SURVIVAL	VHEIGHT	NVH
ACNI	1081/82	a	2	8	2	3.43000	1
ACNI	1083/82	b	4	11	7	1.65500	2
ACNI	1084/82	c	4	12	9	2.14400	5
ACNI	1070/82	d	3	11	5	2.49000	3
ACNI	1069/82	e	3	9	6	3.45000	4
ACNI	1068/82	f	4	9	5	1.97000	5
ACNI	1071/82	g	2	12	5	2.71500	2
ACNI	1082/82	h	4	8	2	2.98000	1
ACRA	1040/82	i	4	10	34	2.77941	17
ACRA	1041/82	j	3	12	28	2.76867	15
ACRA	1013/81	k	2	10	23	1.86846	13
ACSE	1036/82	l	2	11	28	1.56533	15
ACSE	1035/82	m	2	9	31	1.40000	15
PCIN	1062/82	n	3	10	0	.	0
PCIN	1087/82	o	3	8	0	.	0
PCIN	1089/82	p	1	8	0	.	0

PRELIMINARY ANALYSIS

From table 2 it is apparent that there are differences between the figures (our observed response variables) for both survival and height.

A graphic presentation of data is given in figure 2 and figure 3 for survival and height, respectively. The idea of these figures is to give a visual impression of the trial area and the data (their size and variation). The coordinate systems of figure 2 and figure 3 correspond to the coordinate system of the trial layout shown in figure 1. The differences observed do not seem to be related to the trial area in any systematic way. The type of systematic patterns that could be detected by such figures is e.g. a sloping crown canopy caused by a prevailing wind direction, which we then would have to take into consideration when analyzing our data. Usually the systematic variation should be taken into consideration by the blocking scheme applied.

Different mean values have been calculated and compared in order to get an impression of the causes of the differences.

Species mean values (across blocks) for survival and height are shown in figure 4. There would seem to be differences in both survival and height growth between the four species. Whether we can be more certain about the differences depends on the variation within the species. Taking e.g. height, we see from table 2, that maximum and minimum values for the four species are:

Species	Mean (m)	Maximum (m)	Minimum (m)
A. nilotica	2.3	3.5	1.0
A. raddiana	2.6	3.1	1.5
A. senegal	1.6	2.0	1.1
P. cineraria	0.3	0.5	0.0

It is apparent that the minimum-maximum range between the species overlap, and we can thus not be absolutely certain that the differences between the mean values are significant. This can be investigated by a statistical analysis of variance, see below.

Provenance mean values (across blocks) for survival and height are shown in figure 5 and figure 6 respectively. There also seem to be differences between provenances within the four species. Also here the figures are only indicative and we need a statistical analysis of the variance to test whether the differences can be considered significant.

Block mean values (across species) for survival and height are shown in figure 7. There do not seem to be large differences between the blocks (which is in accordance with the assumption made above that the trial area is uniform). However, survival in block 4 may be significantly lower than in the other blocks. Again we need a statistical analysis of the variance to test this.

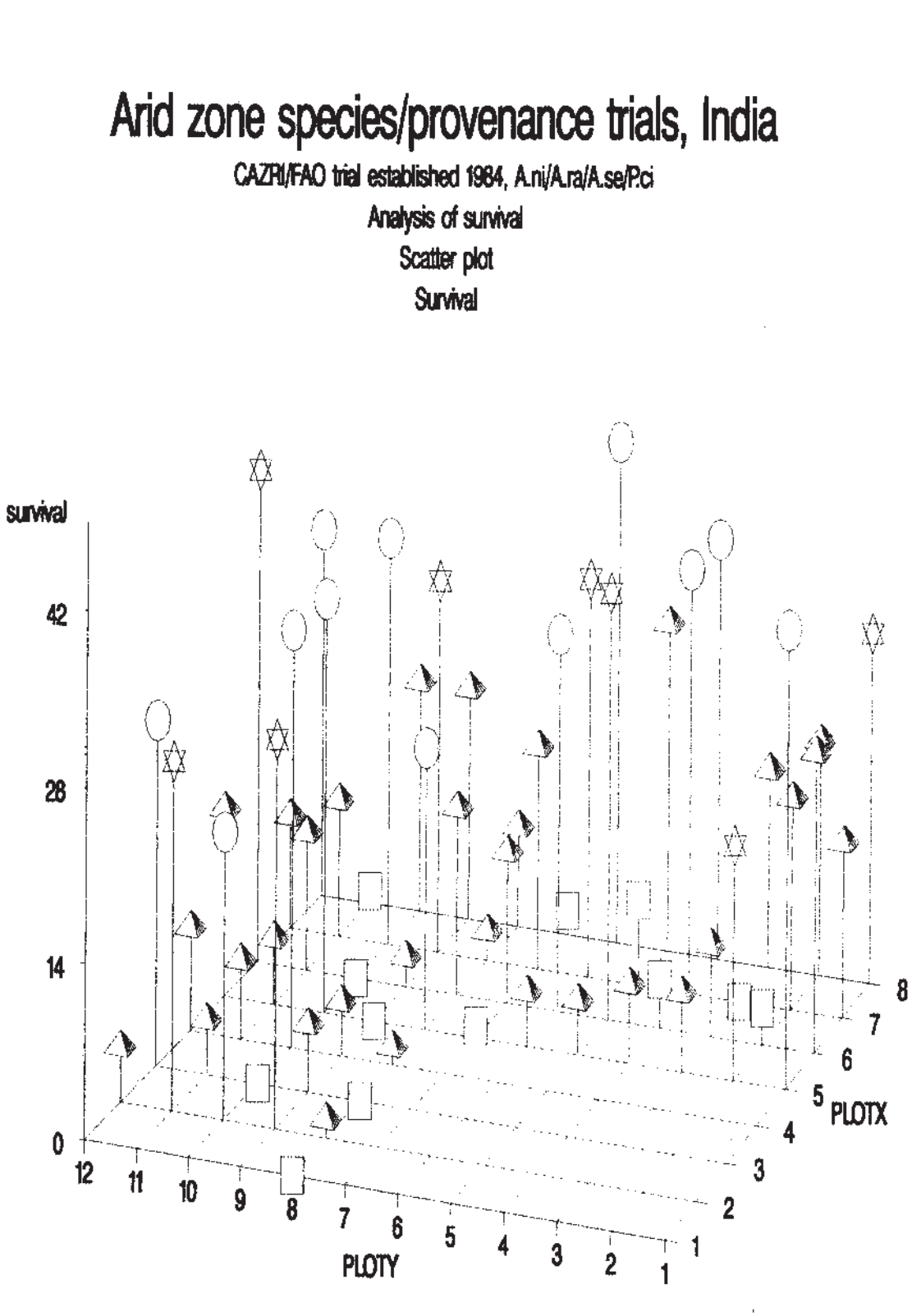


Figure 2: Illustration of survival in the trial. Eash 'needle' represents the plot value. The different symbols represent the different species. 'Balloon' is A. raddiana, 'pyramid' is A. nilotica, 'star' is A. senegal and 'square' is P. cineraria. Plot coordinates correspond to plot coordinates in fig. 1 and table 2.

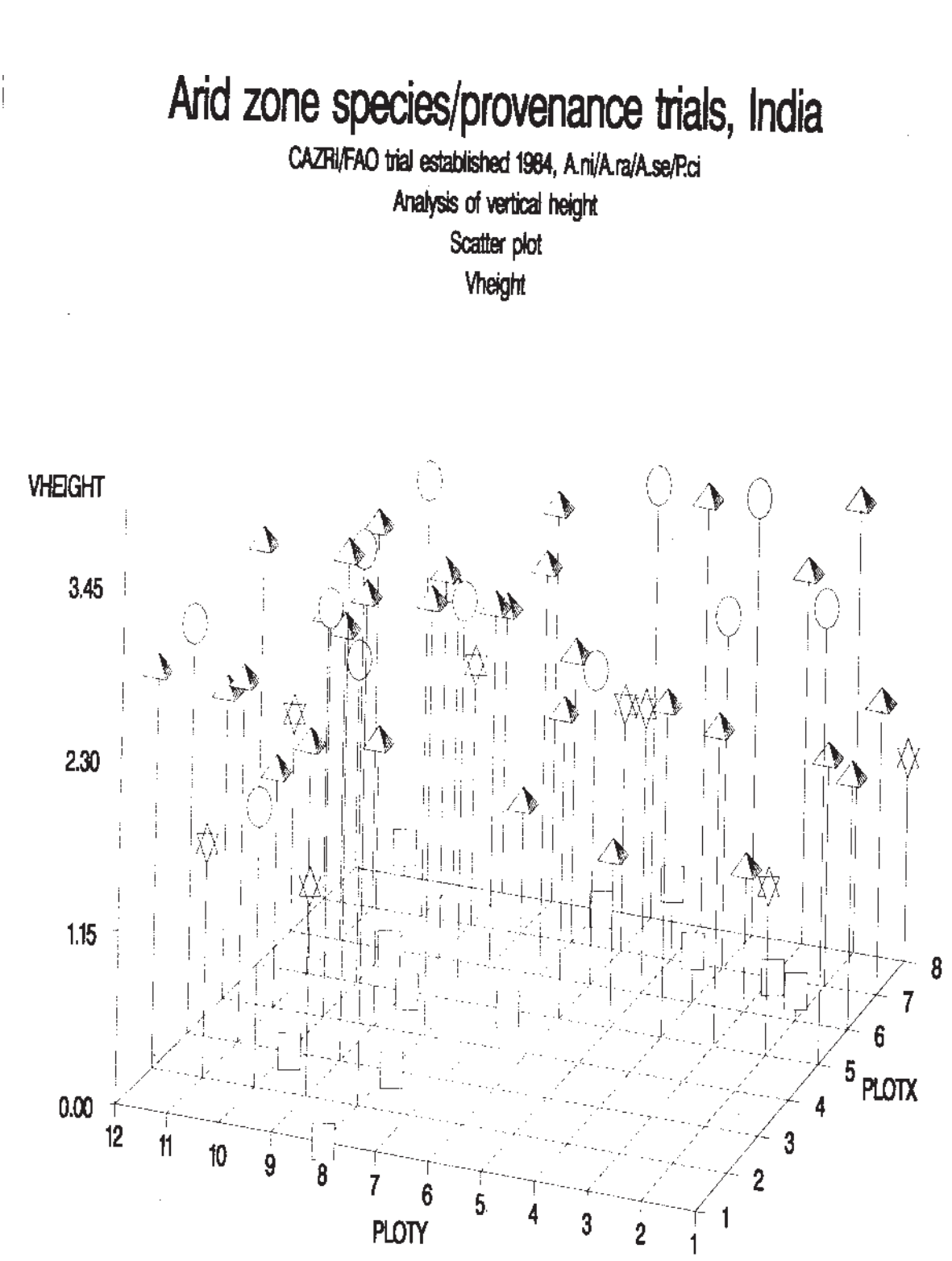
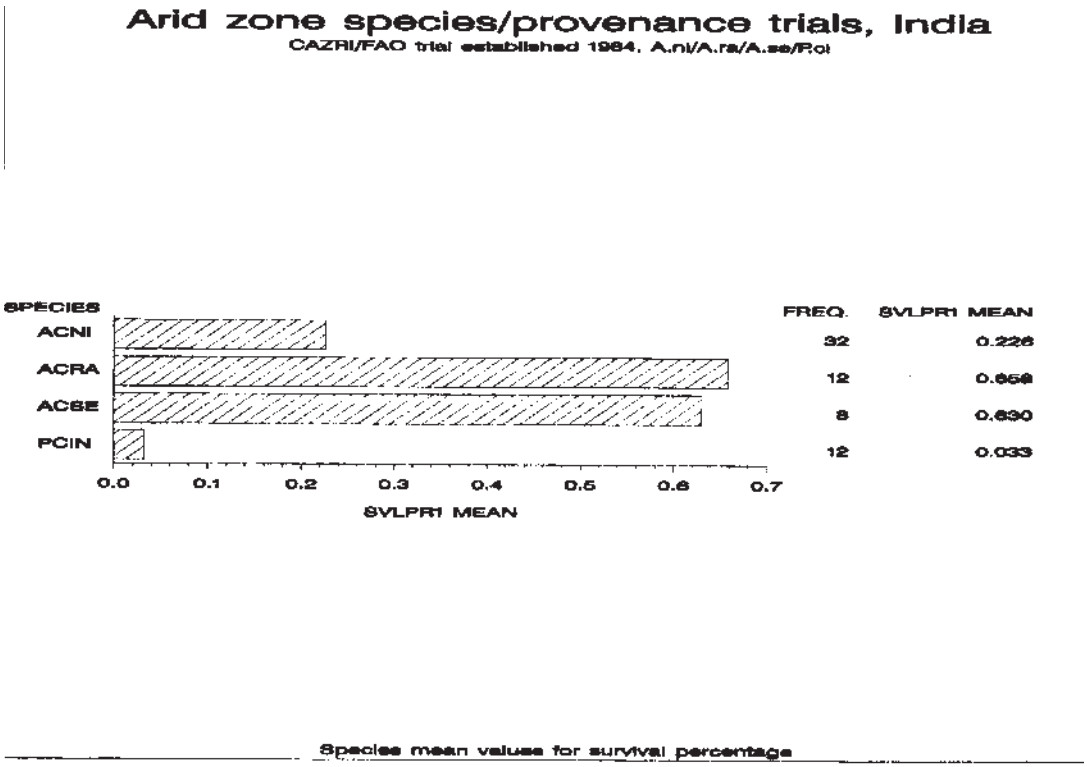
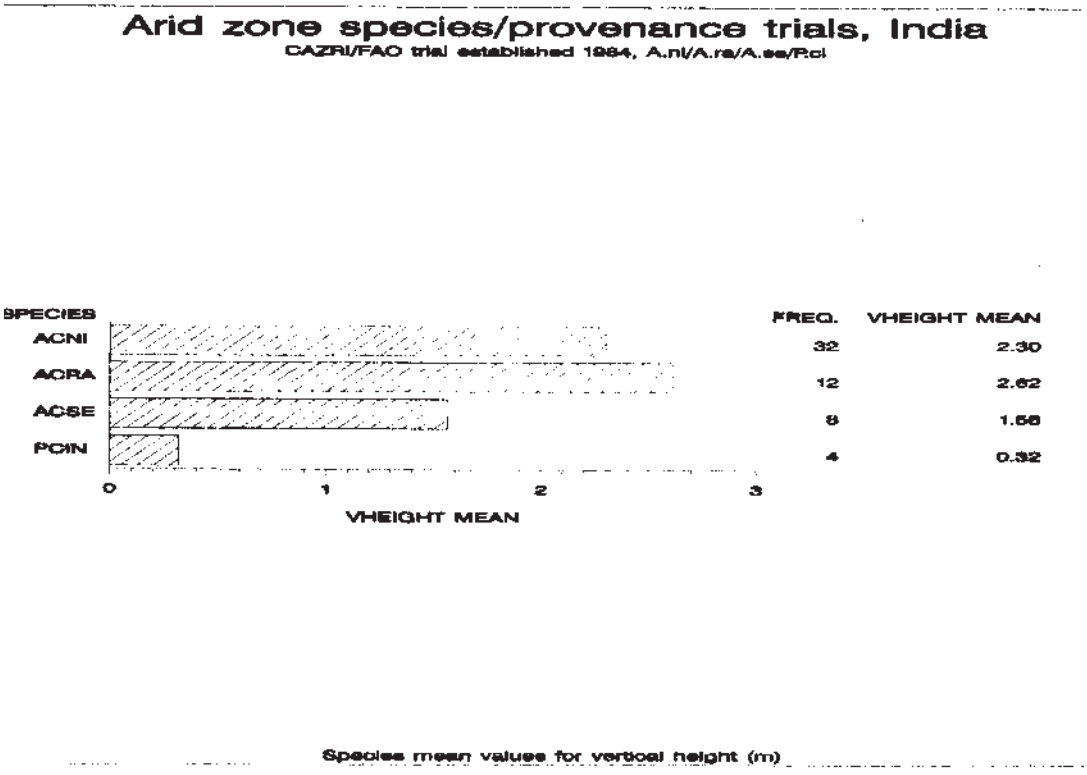


Figure 3: Illustration of vertical height in the trial. East 'needle' represents the plot value. The different symbols represent the different species. 'Balloon' is A. raddiana, 'pyramid' is A. nilotica, 'star' is A. senegal and 'square' is P. cineraria. Plot coordinates correspond to plot coordinates in fig. 1 and table 2.



Species mean values for survival percentage



Species mean values for vertical height (m)

Figure 4a - Species mean values for survival percentage

Figure 4b - Species mean values for vertical height (m)

Arid zone species/provenance trials, India

CAZRI/FAO trial established 1984, A.ni/A.ra/A.se/P.ci

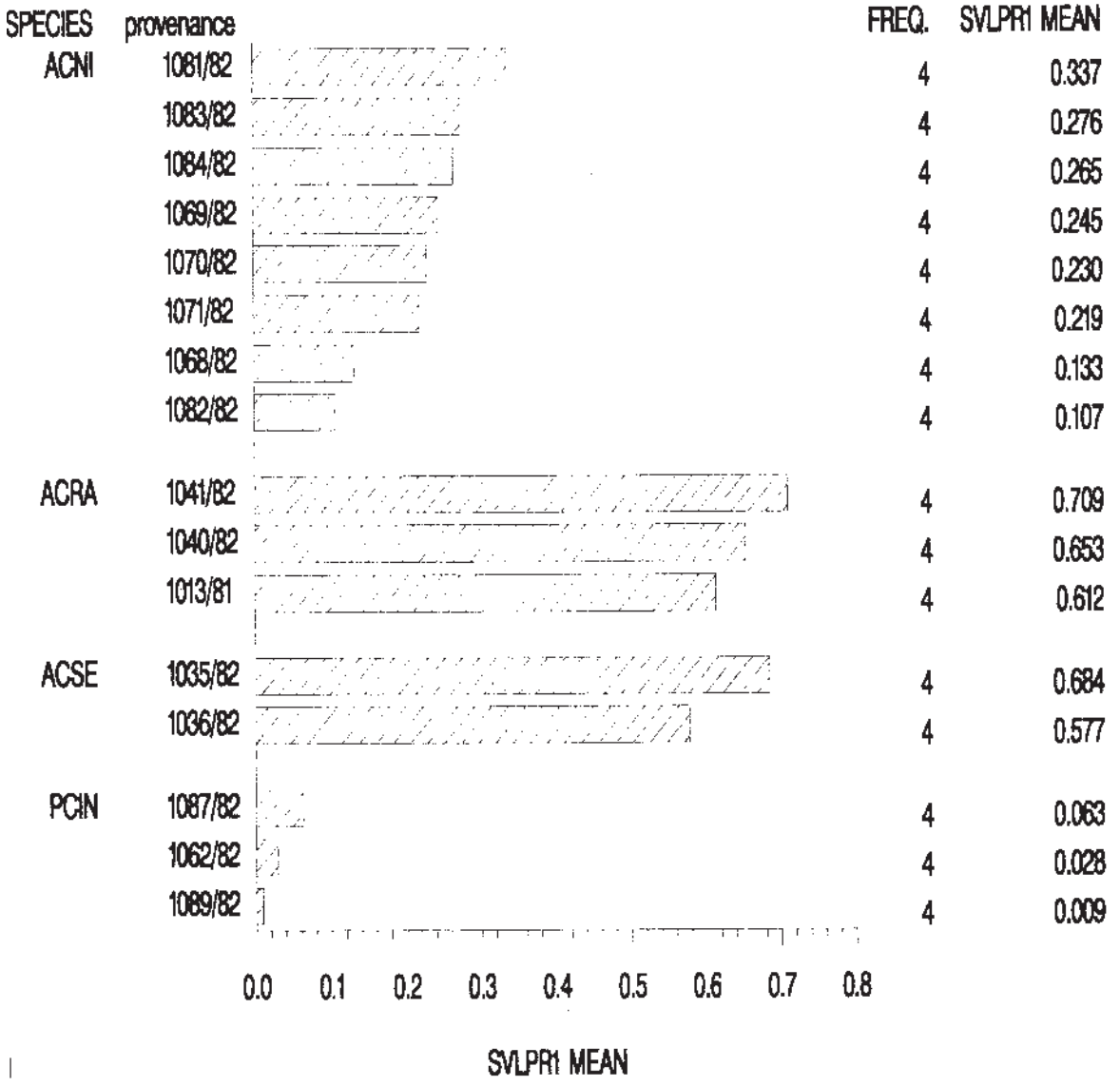


Figure 5: Provenance mean values for survival in descending order within species

Arid zone species/provenance trials, India

CAZRI/FAO trial established 1984, A.ni/A.ra/A.se/P.ci

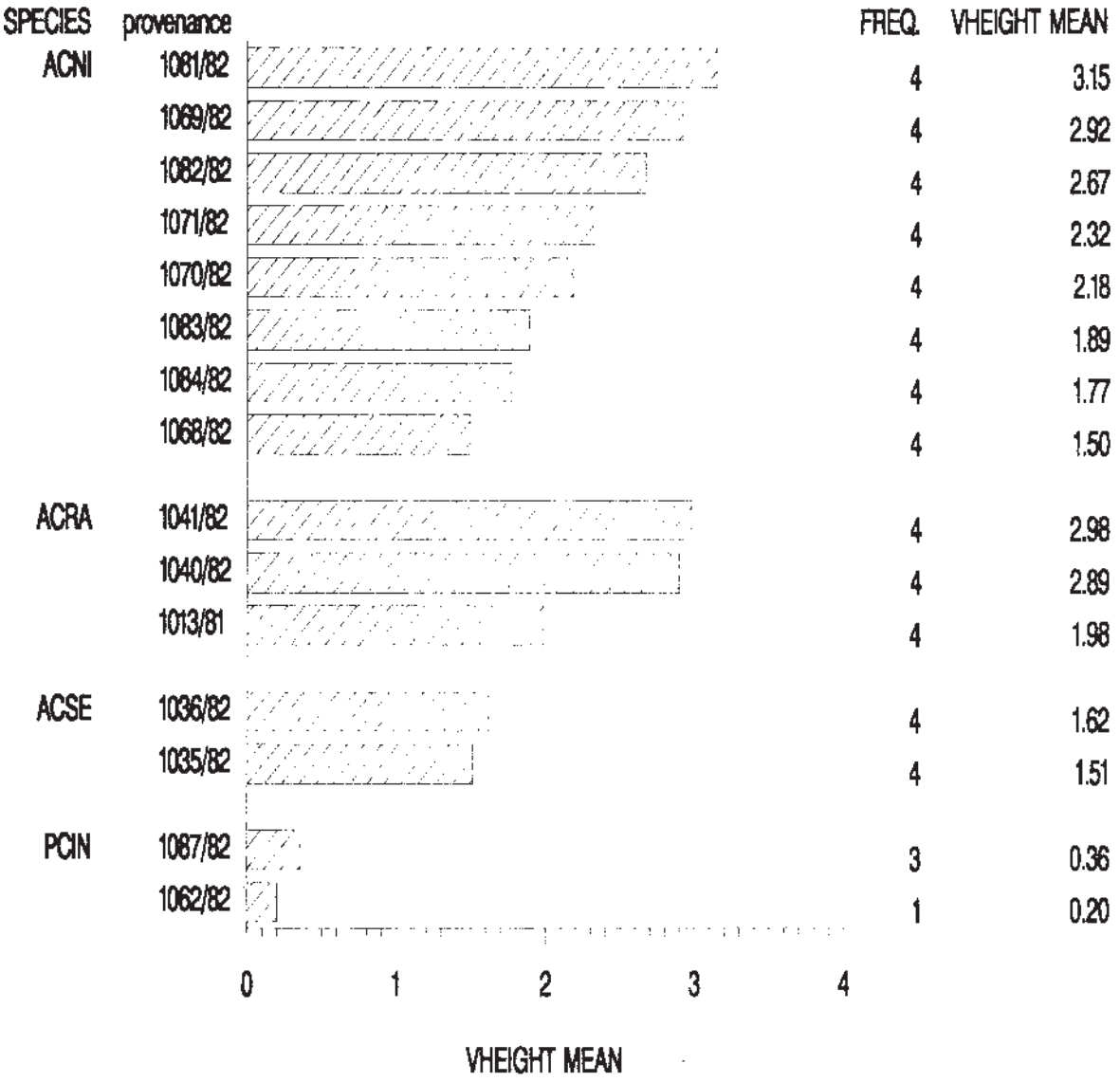
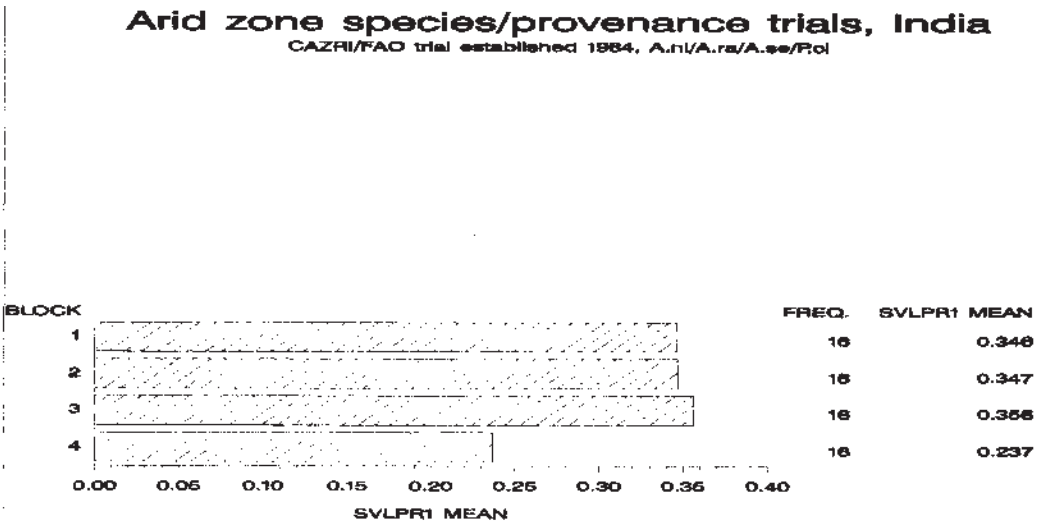


Figure 6: Provenance mean values for height in descending order within species.



Block mean values for survival percentage

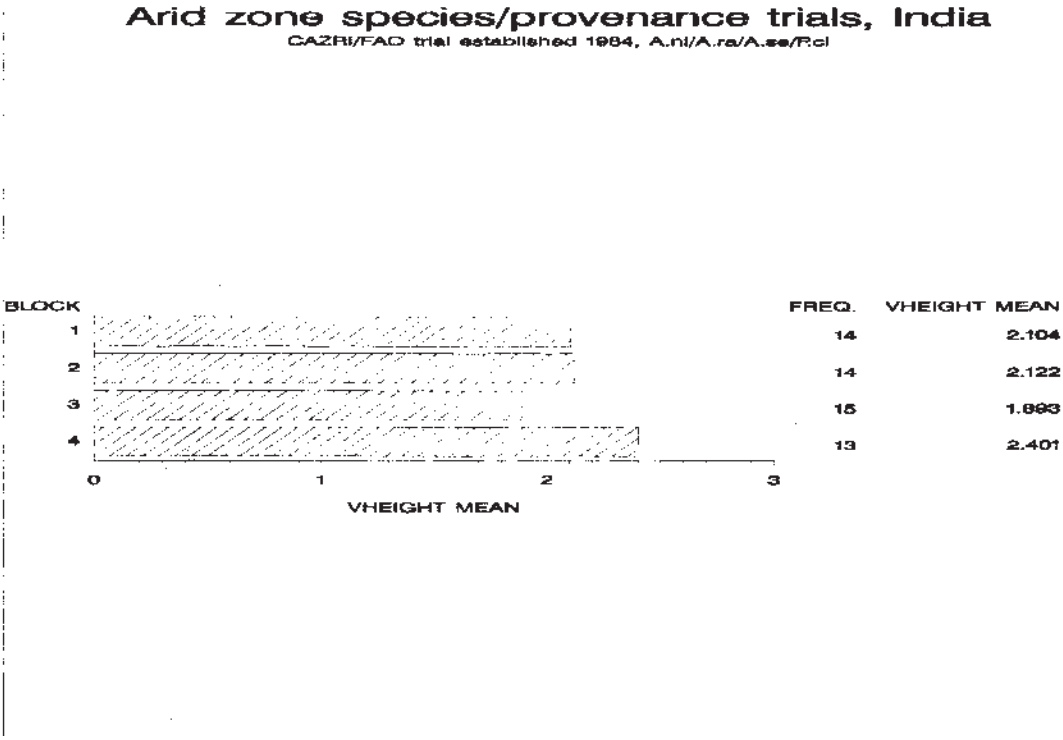


Figure7: Block mean values for survival and height

In summary we would guess that there are differences caused by genetics (species/provenances), but hardly by the environment (blocks).

> Block is here interpreted as environment. This does not mean that it includes effect of environment completely. The residual variation (plot-to-plot variation) is also (primarily) environmental, but of a random nature. Block variation thus constitutes the systematic part of the environmental effect that can be identified by the model.

To test whether the differences noted above are most likely to be due to genetics (species and/or provenance), environment (blocks), or the residual effect, we assume that the differences are insignificant and have occurred by pure chance (our statistical hypothesis, see below). To calculate the probability of this being true we formulate a statistical model. If the probability turns out to be low, we reject our hypothesis and conclude that there are statistically significant differences, otherwise we accept the hypothesis.

In summary, we (i) formulate a statistical hypothesis (typically 'no difference between treatments'), (ii) we calculate the probability of the hypothesis being true, (iii) if the probability is low, we reject our hypothesis and conclude that there are statistically significant differences.

STATISTICAL MODEL

The statistical model assumed is:

(1) $X_{ijk} = m + a_i + b_{j(i)} + c_k + e_{ijk}$,

where

X_{ijk} is the observed response, i.e. either survival or height,

m is the common mean,

a_i is the effect of species with index i, i = 1,2,3,4,

(a_1 = *Acacia nilotica*,
a_2 = *A. raddiana*,
a_3 = *A. senegal*, and
a_4 = *Prosopis cineraria*),

$b_{j(i)}$ is the effect of provenance with index j within species with index i, $j=1,...,n_i$, $n_1=8$, $n_2=3$, $n_3=2$, $n_4=3$,

($\{b_{1(1)},..., b_{8(1)}\}$ = *Acacia nilotica*,
$\{b_{1(2)},..., b_{3(2)}\}$ = *Acacia raddiana*,
$\{b_{1(3)}, b_{2(3)}\}$ = *Acacia senegal*,
$\{b_{1(4)},..., b_{3(4)}\}$ = *Prosopis cineraria*),

Table 3. Residual plot data for height. NVH, vertical height (ACTUAL VHEIGHT) and symbols (a-p) are repeated from table 2. PREDICTED VHEIGHT is the height estimated by the assumed model, and the residual is the difference between actual and predicted height.

Arid zone species/provenance trials, India
CAZRI/FAO trial established 1984, A.n1/A.ra/A.se/P.cf
Analysis of vertical height
Residual data (VHEIGHT=SPECIES PROV(SPECIES) BLOCK)

BLOCK=1

SPECIES	PROVENANCE	SYMBOL	NVH	ACTUAL VHEIGHT	PREDICTED VHEIGHT	RESIDUAL
ACNI	1081/82	a	7	2.95857	3.12700	-0.16843
ACNI	1083/82	b	5	1.62800	1.86919	-0.24119
ACNI	1084/82	c	10	2.09500	1.74694	0.34806
ACNI	1070/82	d	9	2.04444	2.15414	-0.10969
ACNI	1069/82	e	7	2.73143	2.89616	-0.16473
ACNI	1068/82	f	8	2.04500	1.47303	0.57197
ACNI	1071/82	g	8	2.21000	2.29347	-0.08347
ACNI	1082/82	h	2	2.71000	2.65015	0.05985
ACRA	1040/82	i	17	2.98059	2.86821	0.11238
ACRA	1041/82	j	19	3.05316	2.95434	0.09882
ACRA	1013/81	k	14	1.47143	1.96094	-0.48951
ACSE	1036/82	l	13	1.81615	1.59614	0.22001
ACSE	1035/82	m	23	1.51391	1.48571	0.02820
PCIN	1062/82	n	0	.	0.28994	.
PCIN	1087/82	o	1	0.20000	0.38225	-0.18225
PCIN	1089/82	p	0	.	.	.

BLOCK=2

SPECIES	PROVENANCE	SYMBOL	NVH	ACTUAL VHEIGHT	PREDICTED VHEIGHT	RESIDUAL
ACNI	1081/82	a	13	3.03769	3.14518	-0.10749
ACNI	1083/82	b	6	2.52667	1.88738	0.63929
ACNI	1084/82	c	3	1.22333	1.76513	-0.54179
ACNI	1070/82	d	4	2.12750	2.17232	-0.04482
ACNI	1069/82	e	7	2.86429	2.91434	-0.05006
ACNI	1068/82	f	1	0.95000	1.49121	-0.54121
ACNI	1071/82	g	9	2.04778	2.31165	-0.26388
ACNI	1082/82	h	5	2.81600	2.66833	0.14767
ACRA	1040/82	i	11	2.73273	2.88639	-0.15366
ACRA	1041/82	j	19	3.03526	2.97252	0.06274
ACRA	1013/81	k	13	2.16077	1.97912	0.18165
ACSE	1036/82	l	18	1.95444	1.61433	0.34012
ACSE	1035/82	m	13	1.71077	1.50389	0.20688
PCIN	1062/82	n	0	.	0.30812	.
PCIN	1087/82	o	2	0.52500	0.40043	0.12457
PCIN	1089/82	p	0	.	.	.

BLOCK=3

SPECIES	PROVENANCE	SYMBOL	NVH	ACTUAL VHEIGHT	PREDICTED VHEIGHT	RESIDUAL
ACNI	1081/82	a	13	3.17462	3.03706	0.13756
ACNI	1083/82	b	15	1.76000	1.77925	-0.01925
ACNI	1084/82	c	12	1.61833	1.65700	-0.03867
ACNI	1070/82	d	8	2.04750	2.06420	-0.01670
ACNI	1069/82	e	11	2.63182	2.80622	-0.17440
ACNI	1068/82	f	3	1.02000	1.38309	-0.36309
ACNI	1071/82	g	5	2.29400	2.20353	0.09047
ACNI	1082/82	h	4	2.18750	2.56021	-0.37271
ACRA	1040/82	i	20	3.07300	2.77827	0.29473
ACRA	1041/82	j	19	3.05316	2.86440	0.18876
ACRA	1013/81	k	20	2.43600	1.87100	0.56500
ACSE	1036/82	l	13	1.14154	1.50620	-0.36467
ACSE	1035/82	m	19	1.41105	1.39577	0.01528
PCIN	1062/82	n	2	0.20000	0.20000	-0.00000
PCIN	1087/82	o	4	0.35000	0.29231	0.05769
PCIN	1089/82	p	0	.	.	.

BLOCK=4

SPECIES	PROVENANCE	SYMBOL	NVH	ACTUAL VHEIGHT	PREDICTED VHEIGHT	RESIDUAL
ACNI	1081/82	a	1	3.43000	3.29165	0.13835
ACNI	1083/82	b	2	1.65500	2.03384	-0.37884
ACNI	1084/82	c	5	2.14400	1.91159	0.23241
ACNI	1070/82	d	3	2.49000	2.31879	0.17121
ACNI	1069/82	e	4	3.45000	3.06081	0.38919
ACNI	1068/82	f	5	1.97000	1.63768	0.33232
ACNI	1071/82	g	2	2.71500	2.45812	0.25688
ACNI	1082/82	h	1	2.98000	2.81480	0.16520
ACRA	1040/82	i	17	2.77941	3.03286	-0.25345
ACRA	1041/82	j	15	2.76867	3.11899	-0.35032
ACRA	1013/81	k	13	1.86846	2.12559	-0.25713
ACSE	1036/82	l	15	1.56533	1.76079	-0.19546
ACSE	1035/82	m	15	1.40000	1.65036	-0.25036
PCIN	1062/82	n	0	.	0.45459	.
PCIN	1087/82	o	0	.	0.54690	.
PCIN	1089/82	p	0	.	.	.

c_k is the effect of block with index k, k = 1,...,4 and

e_{ijk} is the residual,

$\{a_i\}$, $\{b_{j(i)}\}$, $\{e_{ijk}\}$ are assumed to be independent and normally distributed with zero mean and variances σ^2_a, σ^2_b and σ^2, respectively. Note that provenance is nested within species. This implies that the variance between provenances within the four species are assumed to be the same (σ^2_b).

In the model formulated, the effects of species and provenance are assumed to be random (opposed to a model with fixed effects, see e.g. Snedecor & Cochran 1980 p 238 ff). A random effects model is i.a. used to be able to estimate variance components, see below, 'Calculation of genotypic values'.

When the effect of a factor is assumed random, we consider the treatments drawn at random from a larger population (our variable is stochastic) to which we may extend our conclusions. When a factor is assumed fixed, we consider only the specific treatment levels included in the trial. Random and fixed effects are discussed in more detail below.

The basic assumptions of statistical models should generally be checked. The assumptions to be checked are the (i) expectation structure implied by the model (here additivity of factors included), (ii) independence, (iii) normality and (iv) variance homogeneity (for a discussion of these assumptions, see e.g. Snedecor & Cochran 1980 p 274 ff).

Re (i). The expectation structure implied by our model is that the observed response (survival or height) depends only on the factors species, provenance (nested within species) and block, and that the effects of the factors are additive. As each provenance occurs only once in each block, it is not possible to test whether these two factors interact. We may, however, get an impression of such interaction by plotting the two factors in a diagram. In figure 8 and figure 9 is plotted survival and height respectively as a function of provenance for each block. These figures may not, at first, look too easy to grasp. What is important is, however, to get an impression of whether the four lines each representing one block in the respective figures can be considered parallel except for possible random fluctuations. The degree of parallelism required cannot be stated in general. In our case the random fluctuations may seem to 'blurr' the figures completely. However, rather than looking for parallellism alone, one should also look for possible systematic deviations from parallelism. No such systematic deviations can be seen. The fluctuations seem to be random. The assumption of additivity in our model is thus likely to be a reasonable approximation.

Re (ii). It is only rarely possible to detect dependence between the residuals by looking at the data. It is more a question of considering circumstances that may have influenced the results. The most common cause of dependence between the residuals is that an influential factor have been 'forgotten' in the model or in the planning of the experiment. One could for instance imagine that a family relationship between some individuals within a provenance could influence the results decisively. It is therefore of utmost importance that all possible factors are considered when the trials are planned. In provenance trials the possible effect of family is eliminated by ensuring that seed is collected from a sufficient number of unrelated individuals. If the effect of family is being investigated, seed collections should be kept separate as single-tree collections. In general, randomization is an effective means against unknown or 'forgotten' factors.

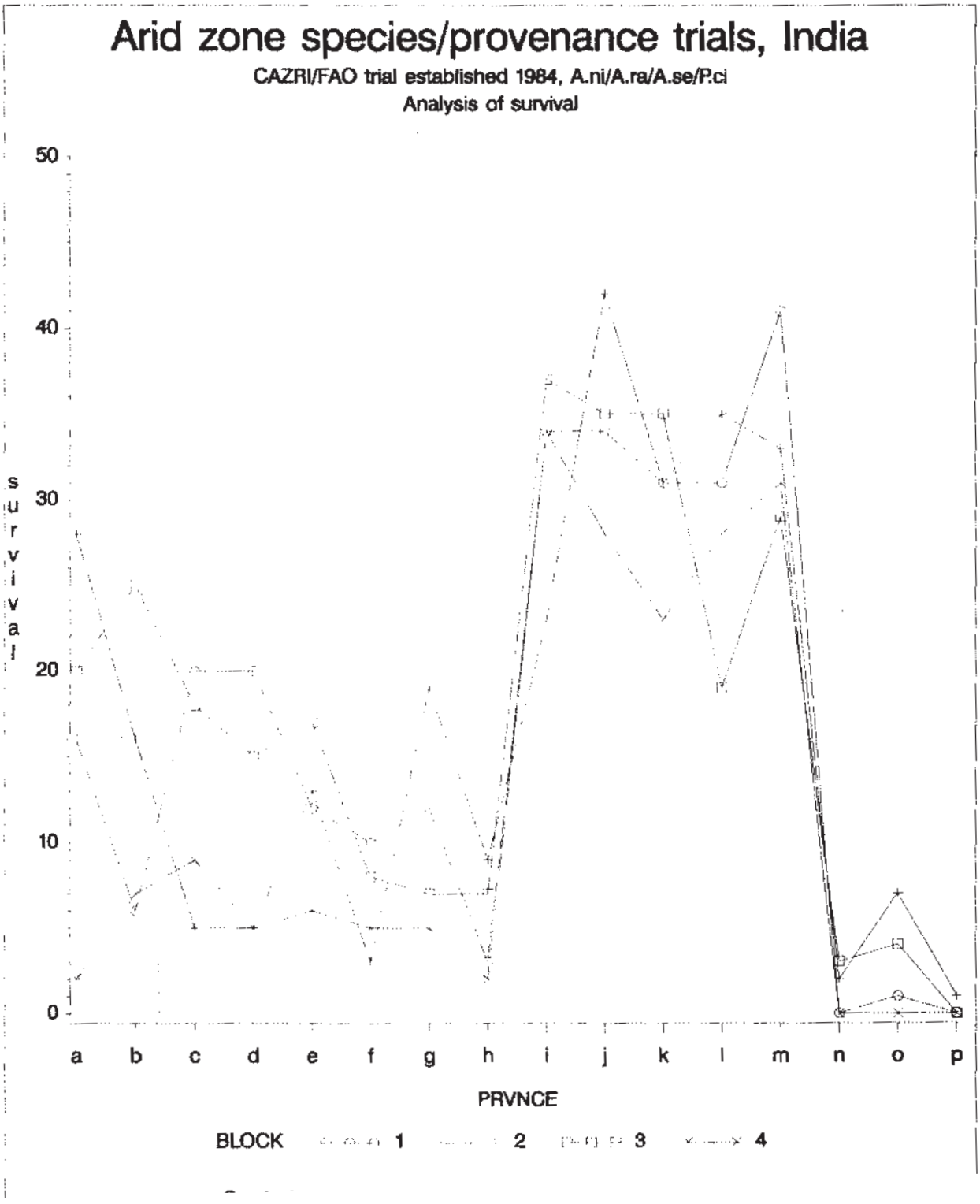


Figure 8: Survival as a function of provenance by block

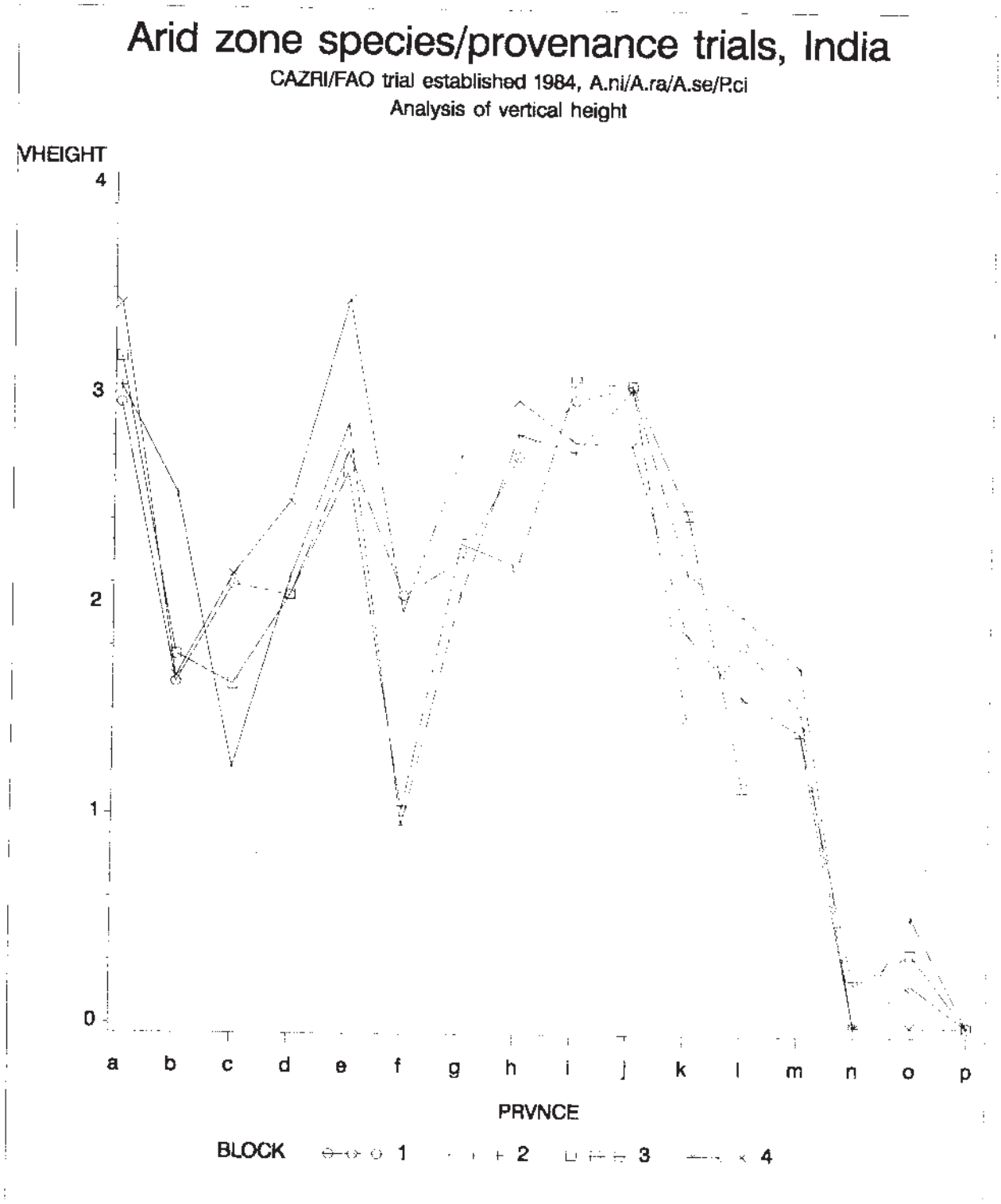


Figure 9: Vertical height as a function of provenance by block

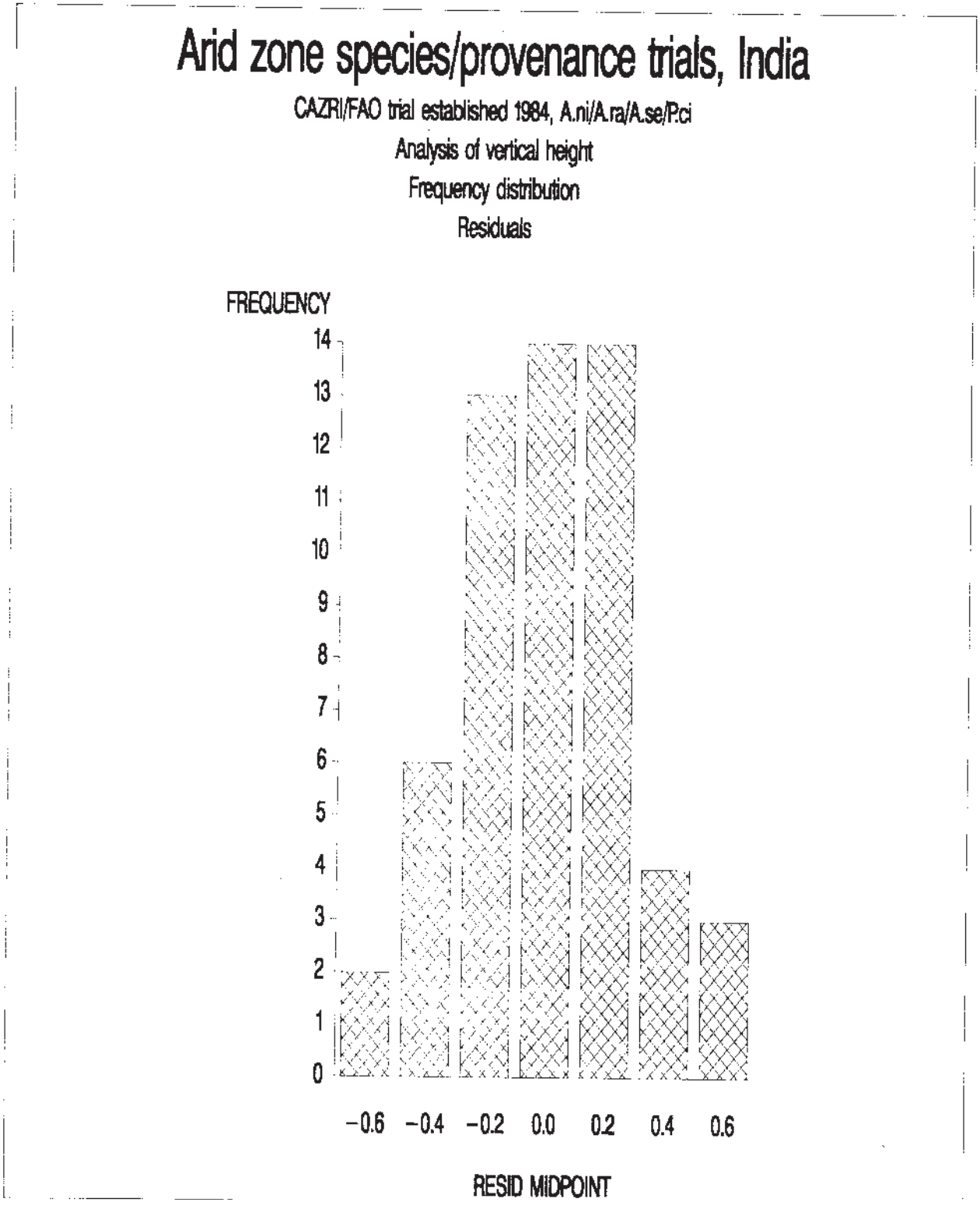


Figure 10: Frequency distribution of the residuals (RESID) from table 3. Residuals have been grouped in intervals according to the interval midpoints shown. Values halfway between two midpoints have been put in the higher interval

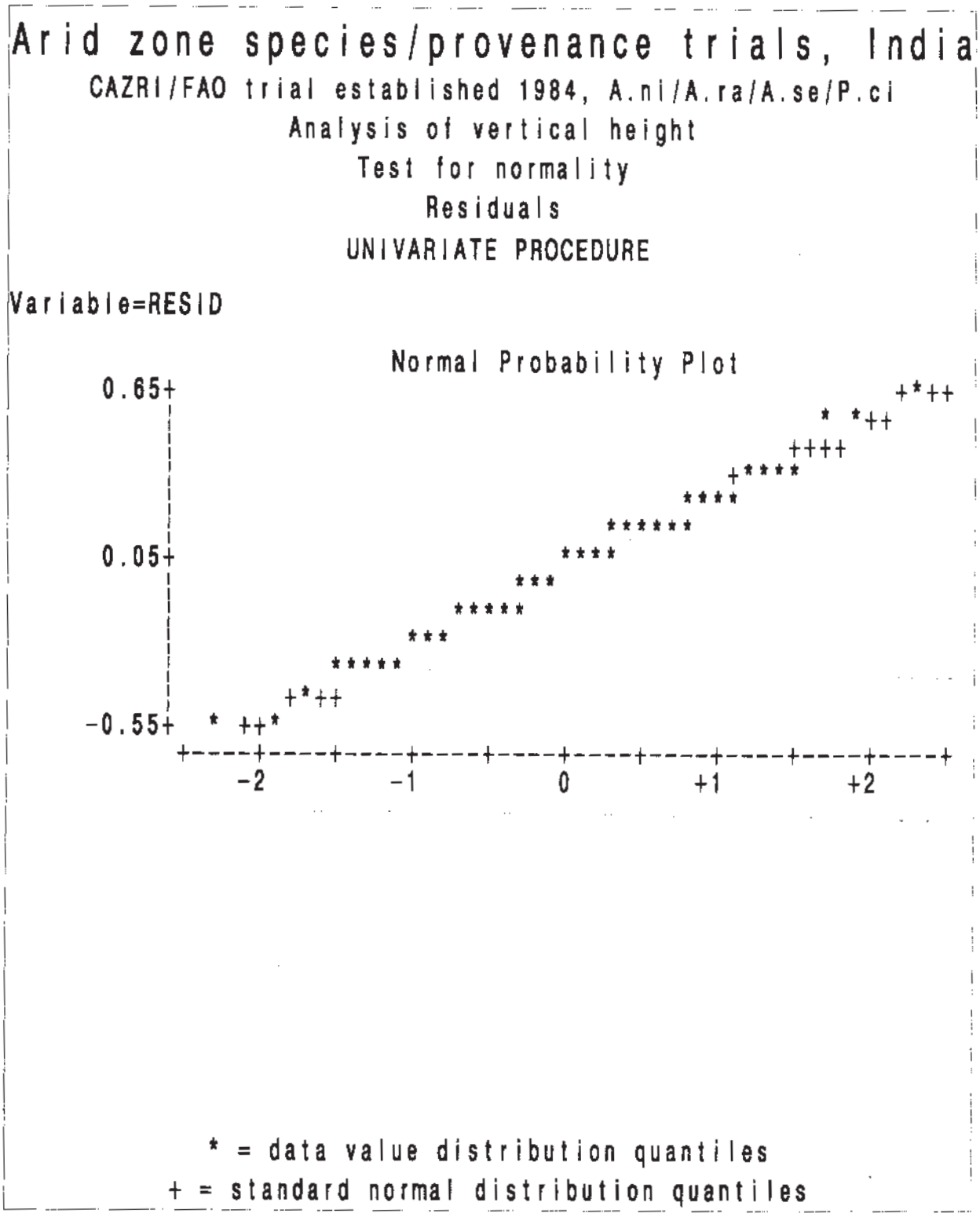


Figure 11: Normal probability plot of the residuals.

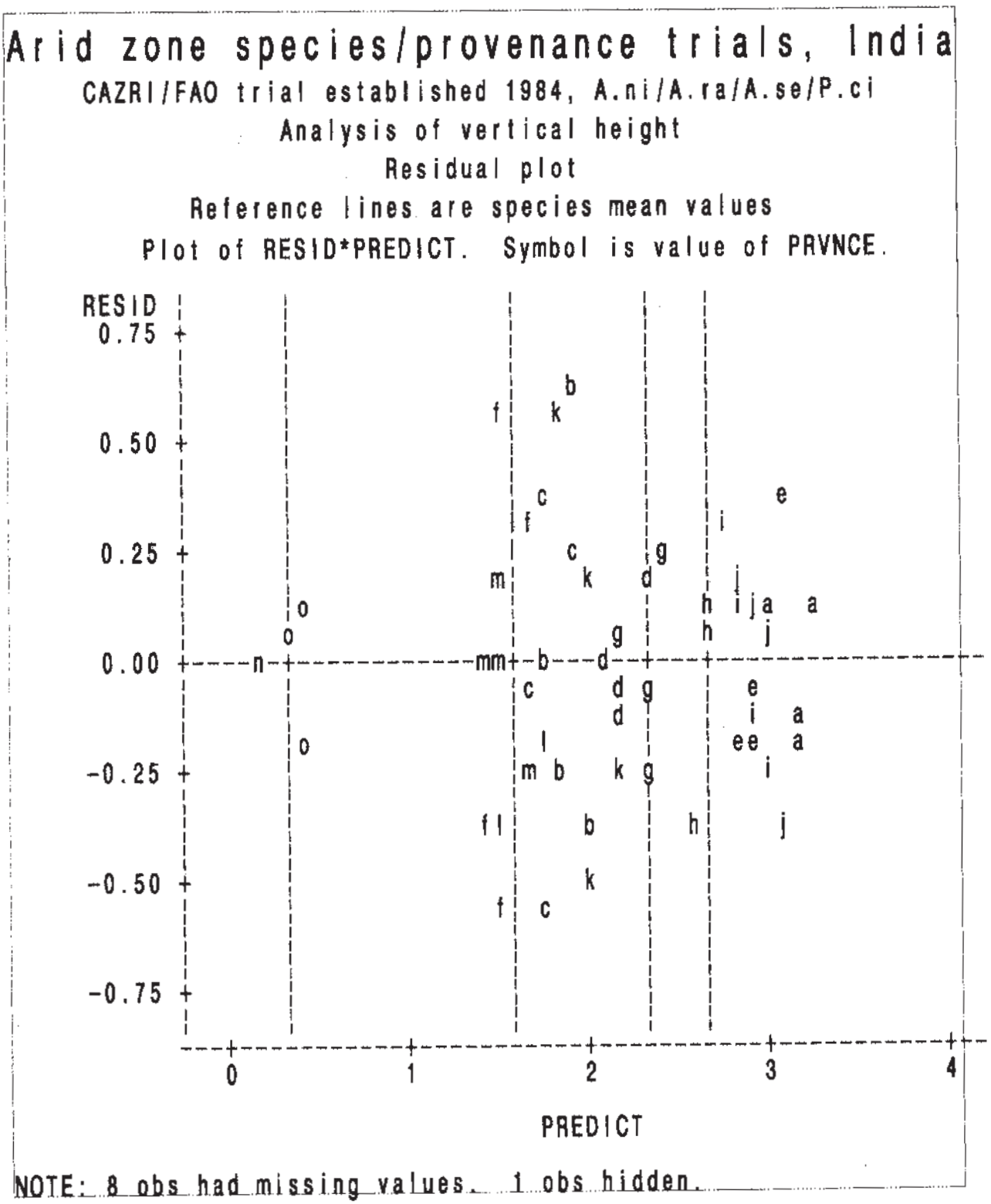


Figure 12: Residuals plotted against predicted vertical height. letter symbols represent different provenances, cfr. table 3.

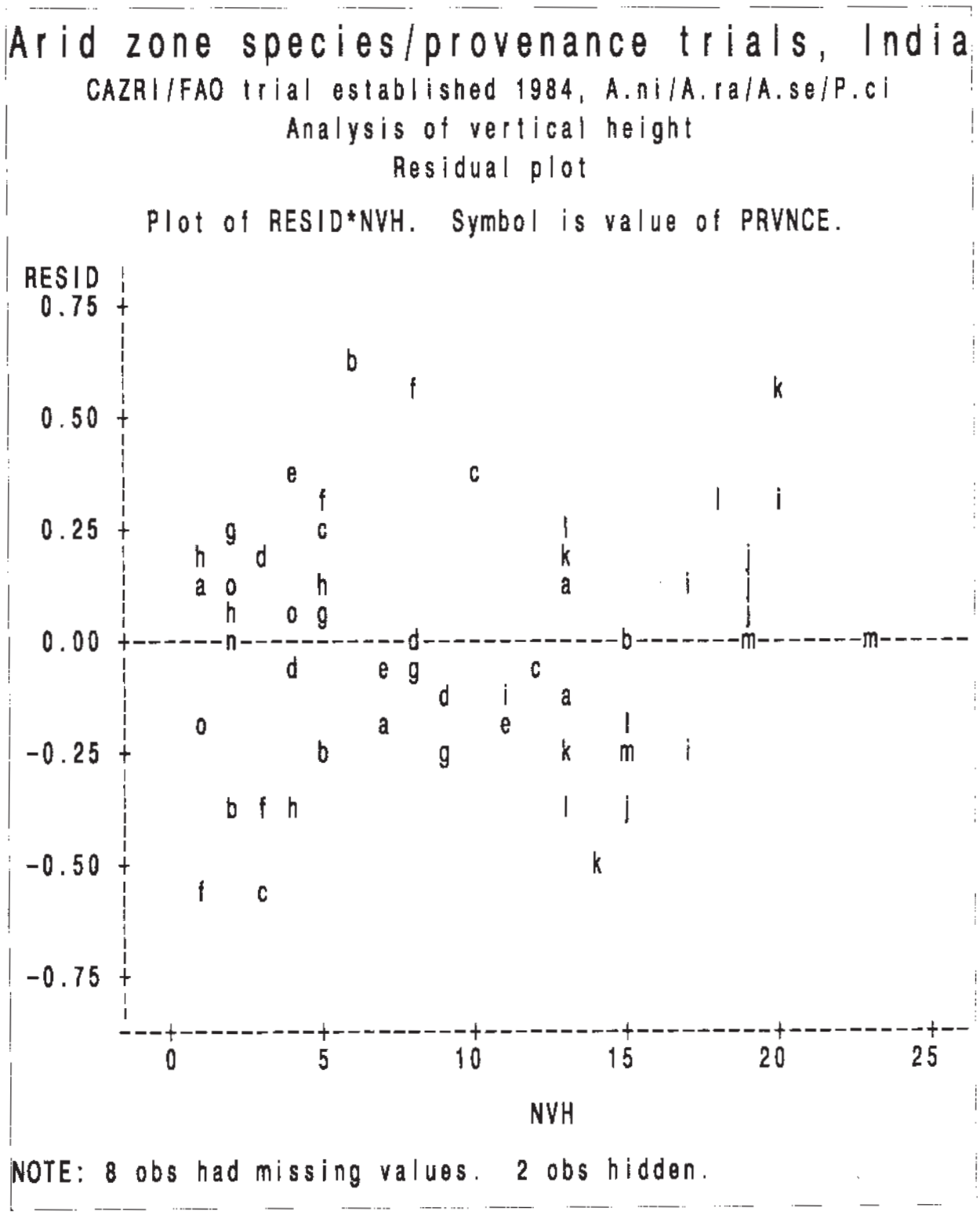


Figure: 13 Residuals plotted against number of trees assessed for height. Letter symbols represent different provenances, cr. table 3.

Re (iii) and (iv). Generally, assumptions of normality and variance homogeneity can be checked by a graphic analysis of the residuals, if the number of observations is sufficiently large. A good account of the examination of residuals (and 'standardized' or studentized residuals) is given by Draper & Smith 1981 p 141 ff.

Re (iii). Residual data is given in table 3. The frequency distribution of the residuals is shown in figure 10. It is seen that the distribution is approximately normal. Another way of checking normality, which can be used also with relatively few observations, is available in the SAS package using the UNIVARIATE Procedure providing a test-statistic as well as a normal probability plot and a number of other descriptive statistics for the distribution of a variable (SAS Procedures Guide). The normal probability plot for our residuals is shown in figure 11. Asterisks (*) mark the data values and plus signs (+) provide a reference straight line. If the distribution turns out not to be normal, it may be necessary to transform data to be able to perform a reliable statistical test (Snedecor & Cochran 1980 p 282).

Re (iv). The assumption of variance homogeneity, i.e. same variance for all treatments, can be checked by finding out whether there are systematic differences between the residuals, e.g. as a function of the values predicted by the model (cf. table 3). In figure 12 is shown a plot of the residuals against the predicted values of the vertical height (cf. table 3). There do not seem to be any particular relationship. If any, there may be a tendency that intermediate provenances are more variable than the extremes. The provenances labelled 'f', 'k', 'b' and 'c' would seem to be more variable than the others. From table 2 and 3 it is seen that our plot mean values for height are based on different numbers of individuals (NVH). The different plot values are thus not estimated with the same accuracy. A plot of the residuals against NVH, see figure 13, shows that the high variability of provenances 'b', 'c' and 'f' could be due to poor accuracy of the mean values because of few surviving individuals. For provenance 'k' this is, however, not the case, and in general there do not seem to be such a relationship, which could be accounted for by using NVH as a 'weight' in our analysis. Based on the residual plots, it would in our case hardly make any difference. A weighted analysis actually confirms this.

Different variance-stabilizing transformations are given by Snedecor & Cochran 1980 p 287 ff). Actual tests for variance homogeneity also exist. Often used is Bartlett's test (see e.g. Snedecor & Cochran 1980 p 252). It is, however, approximative and cannot be recommended for general use.

STATISTICAL HYPOTHESES

If no true differences exist, the following three statements would be valid:

H_{01}: $a_1 = a_2 = a_3 = a_4$ (i.e. no difference between the four species)

H_{02}: $b_{1(1)} = ... = b_{8(1)}$,
$b_{1(2)} = b_{2(2)} = b_{3(2)}$,
$b_{1(3)} = b_{2(3)}$,
$b_{1(4)} = b_{2(4)} = b_{3(4)}$
(i.e. no difference between provenances within each of the four species),

H_{03}: $g_1 = ... g_5$ (i.e. no difference between blocks),

The three hypotheses formulated are actually the hypotheses corresponding to a model with fixed effects. The effects of species and provenance were specified as random above. With a random effects model we assume that our treatments are representative samples from a defined population and the hypotheses of 'no difference' between species and provenances, respectively, are then correctly formulated:

H_{01}: $\sigma^2_a=0$.

H_{02}: $\sigma^2_b=0$.

From this formulation of the hypotheses, it may be clearer that the model used implies an assumption of equal variance between provenances within species. The discussion of variance homogeneity above referred to the variance of the individual treatments (i.e. provenances), but would also apply to the variance between provenances within species. An impression of this variation can be achieved by looking at figures 5 and 6. With less than 4 provenances in three of the four species, it is, however difficult to get any precise impression and it would not be possible to perform a reliable test of variance homogeneity either. A solution to this problem could be to make separate analyses by species (cf. below).

When conclusions are drawn from a random effects model, it is very important to specify the random effect populations correctly, otherwise conclusions may be too general. In our case it would not be reasonable to assume that the provenances included are representative samples of the total set of provenances for the respective species, but only for sub-sets of the provenances. In the conclusions given below we limit ourselves to the specific provenances being studied. Often it would then be considered more correct to use a fixed effects model, if it was not for our particular interest in estimating variance components.

Although we do limit ourselves to the specific provenances, it is, however, not the particular seedlots being tested that is of interest but rather the continuous seed supply from the stands of origin. A random effects model is also therefore the logical choice.

For each species, the particular stands (provenances) in which the seed was harvested can be considered a population. Each stand constitutes a stratum in the population. A seedlot harvested in one stand is thus a sample of this stratum and the samples from all the strata

constitute a sample from the different levels of the population. The population to which we extend our conclusions when using a model with random effects is thus the in principle infinite number of seedlots that can be harvested from the motherstands (or seedstands derived from these). This argument for using a random effects model is usually not considered in statistical textbooks discussing fixed and random effects in relation to different types of trials, including growth trials in the field. It is, however, in line with the intuitive thinking of tree breeders. As it is more difficult to achieve statistical significance with random than with fixed effects models (cf. below), one may consider the interpretation given here precautionary. With the variation in seed-set from year to year in a given stand both in terms of amount and of individuals (genotypes) contributing to the reproduction, it is no doubt more reasonable to use a random effects model.

We test the three hypotheses listed above by calculating the probability of achieving the observed response variables shown and illustrated graphically in figure 4, 5 and 6 by pure chance if the hypotheses are true and the assumptions of the model are fulfilled. To calculate these probabilities we make an analysis of variance.

ANALYSIS OF VARIANCE

The results of the analysis of variance for survival and height are summarized in table 4 and table 5, respectively.

Survival can attain one of two values, either dead or alive. It is a socalled binomial character following a binomial distribution, i.e. non-normality. To achieve normality survival percentages have been transformed, using the arc-sine transformation (see e.g. Burley & Wood 1976 p. 114).

The calculations for the analysis of variance were made by using a computer programme. For details on how to make such calculations see e.g. Burley & Wood, 1976 or Snedecor & Cochran, 1980.

The computer programme is Statistical Analysis System (SAS). The procedure used for the analysis was General Linear Models Procedure (GLM), see SAS/STAT Users Guide for details about this procedure. The GLM Procedure is a very powerful procedure providing i.a. analysis of variance, especially for unbalanced data, which is quite common in species and provenance trials. In the example given here, data is unbalanced only for the *Prosopis cineraria* provenances, as long as no weight statement is given. The random statement in GLM provides an option for testing each effect using appropriate F-test statistics as determined by the expected mean squares.

In this context attention should be given to the first and last column in the tables.

The first column of the tables shows the different sources of variation. The last column gives the probabilities of having achieved our observed responses by pure chance if the hypotheses above are true.

The second column shows the degrees of freedom for each source of variation. Degrees of freedom are the numbers of independent comparisons.

The third column shows the sum of squared deviations for each source of variation, which expresses the observed total variability of each source. The fourth column is the mean square, i.e. the sum of squared deviations divided by the degrees of freedom, which expresses the relative amount of total variability of each source. The fifth column is the test variable (labelled F), which is the ratio between the mean squares being compared. The size of the test variable indicates the probability of achieving the observations by pure chance if the hypotheses formulated are true.

The difference between random and fixed effect models is expressed in the expected mean square (EMS) column (see Table 7) and the F-test statistics column of the analysis of variance table, otherwise the analysis of variance is the same. Generally, random effect models allow us to draw wider conclusions than fixed effect models (we extend the conclusions to a larger population). However, at the same time, the likelihood of achieving statistical significance in a random effects model is usually smaller than in a fixed effects model, i.e. the F-test statistic for a source of variation will be smaller (or the same) when specified as random than when specified as fixed.

The probability values that can be derived from the F-test variable are shown in the sixth column.

Table 4: Analysis of variance for survival (arc-sine transformation). The combined species and provenance trial.
DF: degrees of freedom; SS: sum of squares; MS: mean square;
F_{obs}: observed F-test statistic;
$P(F \geq F_{obs})$: probability of achieving F_{obs} by pure chance;
***: $P(F \geq F_{obs}) \leq 0.1$ % (high significance);
**: 0.1 % < $P(F \geq F_{obs}) \leq 1$ % (1 % significance);
*: 1 % < $P(F \geq F_{obs}) \leq 5$ % (5 % significance);
NS: no significance $P(F \geq F_{obs}) > 5$ %).
The 'vane' in the F-test statistic column indicate how the F_{obs} values have been calculated from the MS values. See text for further explanation

Source of variation	DF	SS	MS	F_{obs}	$P(F \geq F_{obs})$	Significance level
Species	3	5.080	1.693	64.61	0.0001	***
Prov. (Species)	12	0.314	0.026	1.50	0.1584	NS
Block	3	0.279	0.093	5.33	0.0032	**
Error	45	0.784	0.017			

Table 5: Analysis of variance for vertical height. The combined species and provenance trial.
DF: degrees of freedom; SS: sum of squares; MS: mean square;
F_{obs}: observed F-test statistic;
$P(F \leq F_{obs})$: probability of achieving F_{obs} by pure chance;
***: $P(F \leq F_{obs}) \leq 0.1$ % (high significance);
**: 0.1 % < $P(F \geq F_{obs}) \leq 1$ % (1 % significance);
*: 1 % < $P(F \geq F_{obs}) \leq 5$ % (5 % significance);
NS: no significance $P(F \geq F_{obs}) > 5$ %).
The 'vane' in the F-test statistic column indicate how the F_{obs} values have been calculated from the MS values. See text for further explanation

Source of variation	DF	SS	MS	F_{obs}	$P(F \leq F_{obs})$	Significance level
Species	3	15.198	5.066	5.45	0.0145	*
Prov. (Species)	11	11.870	1.079	9.93	0.0001	***
Block	3	0.441	0.147	1.35	0.2723	NS
Error	38	4.130	0.330			

If the probability is low, we reject our hypothesis and conclude that there are statistically significant differences. Normally, a probability of 5 % or less is considered low. If the probability lies between 1 % and 5 %, the significance is usually labelled with one asterisk (*), between 0.1 % and 1 % with two (**), and less than 0.1 % with three asterisks (***).

One other statistic provided by the analysis of variance is the R^2 value (squares of the multiple correlation coefficient). R^2 measures how much of the variation in the dependent variable the model accounts for. R^2 is the ratio of the sum of squares for the model and the total sum of squares. R^2 ranges between 0 and 1. In general, the larger R^2, the better the model fits the data. R^2 may therefore be used to interprete how good the model is. It is, however, not the only measure of how good the model is, and it should not be used as the only measure because it may, by itself, be misleading. The assumptions of the model should be checked as described above, and the number of degrees of freedom used by the model should be compared with the total number of degrees of freedom available. It is always possible to get an R^2 arbitrarily close to 1 by including a large number of factors in the model. If the model takes up too many degrees of freedom, the results, however, tend to become meaningless. There is no 'magic' figure as to how many degrees of freedom may be used by the model, but a rough 'rule of thumb' is that at least half the total number of degrees of freedom should be used to estimate the residual error.

CONCLUSIONS OF THE ANALYSIS OF VARIANCE

From table 4 it is apparent that there is a significant difference in survival between the four species and between blocks, but no significant difference between provenances within species.

The survival of *Acacia raddiana* and *Acacia senegal* is relatively high (60-70 %), of *Acacia nilotica* low (10-30 %) and of *Prosopis cineraria* very low (cf. figure 4 and 5). From table 5, we see that there is a significant difference in height growth between species and between provenances within species. *Acacia raddiana* and *Acacia nilotica* are the better growing species (cf. figure 4 and 6).

With its higher survival, *Acacia raddiana* is thus likely to be the most suitable of the four species for wood production in the Thar Desert.

Within *Acacia raddiana* the two provenances from Senegal are better than the one from Israel (cf. figure 6). This is a point of particular interest as an early introduction of *Acacia raddiana* originating from Israel has performed well in Rajahstan. There may thus be scope for improvement through introduction of new material i.a. from Senegal.

> Above it was indicated that one of the assumptions of the model, equal variance between provenances within species, could be doubtful. If this assumption is false, we may have made a wrong conclusion concerning differences between provenances. Species-wise analyses do, however, confirm the results of the combined analysis.

Although the conclusions seem rather straightforward, it is important to note that we have not proven the unsuitability of any of the three other species.

Acacia senegal is slow growing, but we have only tested two provenances, and the primary product of this species within its natural distribution area is gum, rather than wood, so further testing may be relevant but probably within a different management regime, aiming at gum production. What should be further tested is rather gum production than growth.

Acacia nilotica has a low survival, but all the provenances originate from relatively high rainfall areas in comparison with Jodhpur. An exception is the Sudanese provenance, subspecies *nilotica*, which in Sudan requires annual inundations. The reasonably good growth of *Acacia nilotica* may thus call for further provenance testing.

Prosopis cineraria has been the most disappointing species. It is, however, quite common and popular in Rajahstan already. Why then the bad performance in the trial? There are possibly several reasons. Rainfall at the time of establishment was unusually low and none of the three provenances are local. The two Indian provenances originate from higher rainfall areas, and the Yemenite provenance, although from a very low rainfall area, from a different rainfall regime. Provenance testing of *Prosopis cineraria* should thus no doubt continue.

In summary the conclusions would be that there is scope for using/testing the Senegalese *Acacia raddiana* provenances under normal plantation conditions, whereas provenance testing of the three other species should continue (ref. Lecture Note D-3 concerning the various phases of species and provenance testing).

FURTHER ANALYSIS AND CALCULATION OF GENOTYPIC VALUES

For height growth we have found that there are significant differences between provenances within species. From figure 6, we see that these differences seem to be present within *Acacia nilotica* and *A. raddiana*. A more detailed statistical analysis than carried out above confirms this.

The analysis of variance is only the first step in studying the results. The next step is to examine the treatment means and the sizes of differences between them (Snedecor & Cochran 1980 p 224, 233). Above this was done simply by looking at the relevant graphic presentations. The more detailed analysis - which is not reproduced here - was in our example made using multiple comparison options in the SAS GLM Procedure, referred to above.

Above we concluded that the Senegalese provenances of *Acacia raddiana* were probably the best. But how much can we actually expect to gain by selecting the best provenance?

To answer this question we may estimate the genotypic values and the genetic gain.

The genotypic value is the value we can expect from growing the provenance on the test site. Each provenance is, however, only represented on parts of the test site, and only with samples. The observed phenotypic response may thus deviate from the genotypic value, because of random plotwise variation in environmental factors within blocks, and because the planting material in each plot represents a sample of the provenance that may deviate from the average genotype of the provenance.

In trials where provenance mean values are estimated with a small variance (trial error), there will hardly be any difference between phenotypic and genotypic values. With increasing trial error, there will, however, be an increasing tendency that large phenotypic values become larger than the genotypic and that small phenotypic values become smaller than the genotypic. The phenotypic differences observed between provenances are thus larger than can actually be expected from growing the different provenances. In the following it is shown how genotypic values can be calculated using the three *Acacia raddiana* provenances as examples.

A separate analysis of variance for *Acacia raddiana* is shown in table 6. The statistical model used for this separate analysis is:

(2) $X_{ij} = m + a_i + b_j + e_{ij}$,

where

X_{ij} is the observed response, i.e. vertical height,

m is the common mean,

a_i is the effect of provenance with index i, i = 1,2,3,
(a_1 = 1041/82 Keur Mbaye,
a_2 = 1040/82 Rao, and
a_3 = 1013/81 Ein-Hazeva),

b_j is the effect of block with index j, j=1,2,3,4,

e_{ij} is the residual,

$\{a_i\}$, $\{e_{ij}\}$ are assumed to be independent and normally distributed with zero mean and variances σ^2_a and σ^2, respectively.

The phenotypic value of provenance with index i, P_i, is

$P_i = m + a_i$, i = 1,2,3

The genotypic value G_i of provenance with index i is:

$G_i = m + h^2_{BS}(P_i - m)$,

where h^2_{BS} is the broad sense heritability of the provenance mean value (or, shorter, the provenance broad sense heritability).

The genetic gain G_i of selecting provenance with index i among the provenances being tested is:

$G_i = h^2_{BS}(P_i - m)$.

Often it is more relevant to calculate the gain compared with a standard provenance already being used in the area ('local control') instead of the common mean. In this trial, there is no local provenance included. It would be tempting to consider the Israeli provenance in the trial as the 'local control', because the already introduced material being used in Rajahstan originates from Israel. That would, however, be an example of drawing too wide a conclusion. The originally introduced material is possibly from another Israeli seed source, and it is no longer a first generation introduction.

Table 6: Analysis of variance for vertical height. Separate analysis for *Acacia raddiana*.
DF: degrees of freedom; SS: sum of squares; MS: mean square;
F_{obs}: observed F-test statistic;
$P(F \geq F_{obs})$: probability of achieving F_{obs} by pure chance;
***: $P(F \geq F_{obs}) \leq 0.1$ % (high significance);
**: 0.1 % < $P(F \geq F_{obs}) \leq 1$ % (1 % significance);
*: 1 % < $P(F \leq F_{obs}) \leq 5$ % (5 % significance);
NS: no significance $P(F \geq F_{obs}) > 5$ %).
See text for further explanation

Source of variation	DF	SS	MS	F_{obs}	$P(F \geq F_{obs})$	Significance level
Species	2	2.423	1.212	19.37	0.0024	**
Block	3	0.273	0.091	1.46	0.3173	NS
Error	6	0.375	0.063			

Table 7: Variance components and provenance heritability, vertical height, *Acacia raddiana.*
DF: degrees of freedom; SS: sum of squares; MS: mean square;
EMS: Expanded mean square

Source of variation	**DF**	**MS**	**EMS**	**Variance components**
Provenance	2	1.212	$\sigma^2 + 4\sigma_a^2$	$s_a^2 = 0.29$
Block	3	0.091	$\sigma^2 + 3\sigma_b^2$	
Error	6	0.063	s^2	$s^2 = 0.063$

$$h^2_{BS}\,(provenance) = \frac{\sigma^2_a}{\sigma^2_a + \sigma^2/4} \approx \frac{0.29}{0.29+0.063/4} = 0.95$$

Table 8 Vertical height of the three provenances of *Acacia raddiana*. Phenotypic and genotypic values, and genetic gain (_G) from selecting respective provenances among the three.

Provenance (seedlot nr.)	**Phenotypic value (m)**	**Genotypic value (m)**	**_ G**	
			m	%
1041/82	2.98	2.96	0.34	+ 13
1040/82	2.89	2.88	0.26	+ 10
1013/81	1.98	2.01	- 0.61	- 23

The provenance broad sense heritability (h^2_{BS}) is a value between 0 and 1. Compared with P_i, G_i is thus 'pushed' towards the common mean. The differences between the provenances are in other words 'shrunken', and h^2_{BS} is therefore also called the 'shrinking' factor (cf. e.g. Kung 1979).

The provenance broad sense heritability is calculated as follows (cf. e.g. Burley & Wood 1976 p. 124):

$$h^2_{BS} = V_G/V_P = s_a^2/(s_a^2 + s^2/4),$$

where s denotes estimates of σ.

The residual variation (s^2) in the nominator is divided by the number of replications of each provenance (i.e. 4) to achieve the broad sense heritability of the provenance mean value (cf. Nanson 1970). The variation between provenances (s_a^2) is the variation within a set of provenance mean values and may thus be considered the genetic variation (V_G). The residual variation (s^2) is the environmental plot to plot variation within a block. Each provenance is replicated 4 times and the environmental variance component (V_E) of the phenotypic variation between provenances is therefore one fourth of the residual variation.

Burley & Wood 1976 provide a word of caution in the treatment and application of heritability values, and point out that, normally, heritability values are not computed for provenance studies. With the increasing computational capacity that has become generally available it has, however, become more and more common to estimate broad sense heritability in provenance studies. Still, it should be done with some caution, as the accuracy of the estimates is often low, especially in trials with few provenances, and heritability estimates should probably be limited to cases where the analysis of variance shows

The variance components s_a^2 and s^2 are derived from the analysis of variance, cf. table 6.

Each mean square (MS) in the analysis of variance is made up of one or more of the components caused by the different identifiable sources of variation. The composition of the mean square values are shown in the expected mean square (EMS) column, see table 7. From the set of MS-EMS equations, the different variance components can be estimated.

The different genotypic and phenotypic values of the three provenances are summarized in table 8.

The differences in genetic gain presented in table 8 illustrate the importance of selecting the right provenance.

FINAL REMARKS

The appendix has provided an example from real life. It shows the quite considerable amount of analytic work required to interprete field trials, from which the results are often presented in a few lines.

Very often the analysis is more complex than here, e.g. when assessing more characters and investigating how they may be correlated or when comparing several trials at different sites. In correlation analyses it is important not to misinterprete from empirical to causal relationship (think e.g. of how survival may be positively as well as negatively correlated with growth).

It should always be kept in mind that any trial is unique and that the best way of analyzing data in general is to commence by simply "looking at them", as done with the various illustrations presented (see figure 2-7).

Such illustrations are useful when formulating models and hypotheses and interpreting results, because they may reveal causal relationships. They provide preliminary analyses and can be made also by the researcher with limited knowledge of statistical analysis. They will usually be sufficient for drawing preliminary conclusions. How certain they are, can to some extent be seen by looking at the variation in terms of the ranges of minimum-maximum values. The higher the degree of overlap between ranges of different treatments or blocks, the less statistically certain are observed differences between mean values. One should thus not refrain from trial establishment and assessment because of lack of more advanced statistical know-how. Simple analyses and common sense will often provide the main results. Statistical methods are needed to quantify their reliability and statistical know how should thus as a general rule be consulted.

REFERENCES

Burley, J. and P.J. Wood (eds.) 1976
A Manual on Species and Provenance Research with particular reference to the tropics. Tropical Forestry Papers No. 10 and 10A. University of Oxford.

Draper, N. and H. Smith 1981
Applied regression analysis, Second Edition, 709 pp. John Wiley, New York.

FAO. 1988
Final Statement on the FAO/IBPGR/UNEP Project on Genetic Resources of Arid and Semi-arid Zone arboreal Species for the Improvement of Rural Living, 1979-1987. Forest Genetic Resources Information No. 16.

Kung, F.H. 1979
Improved Estimators for Provenance Breeding Values. Silvae Genetica 2-3, 114-116. J.D. Sauerländer's Verlag, Frankfurt a.M.

Madsen, S.Fl. 1989
[Stand Progenies of Danish Norway Spruce, Reports of the Danish Forest Experiment Station, vol.42, 3, 147-21, Copenhagen].

Nanson, A. 1970
L'Héritabilité et le gain d'origine génétique dans quelques types d'éxperiences. Silvae Genetica 19, 114-121. J.D. Sauerländer's Verlag, Frankfurt a.M.

Pedersen, A.P., Kirsten Olesen and Lars Graudal 1993
Tree Improvement at Species and Provenance Level, Lecture Note D-3, Danida Forest Seed Centre.

SAS Procedures Guide 1988
Release 6.03 Edition. Cary, NC: Sas Institute Inc., 441 pp.

SAS/STAT User's Guide 1988.
Release 6.03 Edition. Cary, NC: Sas Institute Inc., 1028 pp.

Snedecor, G.W. and W.G. Cochran 1980
Statistical Methods, Seventh Edition. The Iowa State University Press.