On-Farm Trials –
Some Biometric Guidelines

March 1998

University of Reading
Statistical Services Centre

Biometrics Advisory and
Support Service to DFID

Contents

1. Introduction

Farmer-participatory research trials have rapidly gained popularity in the past few years with due consideration being given to the knowledge, problems and priorities of farming families. The move towards participatory on-farm research means that many researchers, such as breeders and agronomists, who have been trained in techniques of on-station research, are now under pressure to move on-farm. It is therefore perhaps not surprising that the design of such trials has, too often, resulted in miniaturised research institute experiments. Conditions for on-farm trials are typically less controlled than those for research institute fields, and this means that more thoughtful designs are needed.

This booklet concentrates primarily on providing guidelines on aspects of the design and analysis of on-farm trials that are different from on-station research. It is concerned primarily with experiments where the farmer has considerable involvement – and not situations where the farmer's only participation is by providing land. Equally, it does not include experiments that are really just demonstrations. You should therefore ask yourself whether your experiment (a) involves the farmer in design, management or assessment, (b) seeks to address unanswered questions and (c) is an experiment, i.e. it involves planned changes.

2. Types of on-farm experiment

The common distinction is whether an experiment is both designed and managed by the research team, or researcher-designed and farmer-managed or, to some extent, farmer-designed as well as farmer-managed. Let us consider first the trials that are researcher designed and managed. These are trials where the farmers' fields are effectively borrowed by the research team. They become temporarily a part of the research institute. This type of trial is important in the same way as there remains an important place for on-station trials. Furthermore, bringing the research institute, to some extent, onto farms can broaden the range of soils, pests and diseases that are encountered as well as encourage interaction with farmers. The design of such trials is broadly the same as on-station trials, so the information given in the guide called *Concepts Underlying the Design of Experiments* can be used. For researcher-managed, on-farm trials, the present booklet is relevant on problems of site selection and on collaboration with farmers concerning site characterisation and blocking etc. However, the main emphasis of this booklet is on guidelines for trials that are, at least to some extent, managed or designed by the farmers.

From a biometric perspective, a key point in participatory on-farm trials is that their design needs some ideas that are normally associated with the design of a survey, together with concepts from the design of an experiment. Thus we collect some data at the plot level (as in an experiment) and other data at the farmer level (as in a survey). This latter component is new to some scientists, who are used to on-station trials. It implies that concepts of survey design need to be considered, in particular that of stratification. We give guidelines for these two components in sections 4 and 5.

3. Specifying objectives

The initial stimulus for organising experiments on farmers' land was to broaden the range of validity of conclusions beyond the narrow confines of a research institute setting. This is still a valid reason for conducting on-farm trials, but it is now recognised that farmer participation is important and that successful programmes must incorporate farmers' abilities to experiment and innovate.

As with any scientific investigation, it is crucial to specify the objectives of the study clearly. Time must be allowed for this phase and the objectives need to be re-assessed during the planning of the trial, to see whether they need to be revised. This is particularly challenging in on-farm trials, where researcher and farmer are now working together, often with extension staff and NGOs. It is important that the objectives are clearly identified from all perspectives.

Trials should be designed to resolve specific research questions, and researchers need to be impartial to the perception that donors expect to see the words "on-farm" and "participatory", before they will consider supplying funds. Usually a careful assessment of the gaps in the current knowledge will show that a series of initiatives is needed. These may include a small survey, plus a number of trials, some on-station and others that are on-farm, possibly some researcher- and some farmer-managed.

In defining the objectives it is important to check that there remains some genuine research, i.e. some hypotheses to be tested. If the major objective is to encourage adoption of a new technology by farmers, then this may be important extension work, but it is not research. Check also that there are not too many objectives for a single study. For example, objectives relating to adoptability and profitability of different technologies often imply different levels of farmer participation and hence may be better considered in separate studies.

4. Choice of farms and villages

The selection of farms must be closely related to the objectives of the research, and in turn to the recommendation domain for which results are intended. The large variation that generally exists between farms means they must be selected with care to ensure that conclusions will apply to the appropriate group of farmers. An initial survey is valuable in identifying how farms may be grouped, for example according to their socio-economic characteristics and environmental conditions. Decisions have then to be made whether research results will be relevant to all groupings or only to a subset of such farming groups. A representative (usually a random) sample of farms is then selected from the relevant group(s) of farms. Enough farms have to be used to have a reasonable estimate of between-farm variability. Stratified sampling may be recommended, to ensure that a wide range of farms is included in the sample.

A multistage sampling scheme is often used, with village as the primary unit and farming household as secondary units. The *Guidelines for Planning Effective Surveys* booklet gives more guidance on this part of the design. The sample of farmers must be large enough for a valid analysis when split into different groups, for example by soil type, tenants and owners, access to credit or not. Where resources seriously limit the number of farms in the study, the objectives of the study may have to be re-examined. For example, for a new topic, the first year may become a pilot study, from which ideas and objectives are refined for the following year's research.

When selecting villages, consideration must be given to how long a village remains associated with a research institute. Repeated use of the same villages, or use by different organisations is simple, but such villages may become less representative of the region. When selecting farms any restriction of the sampling scheme so that only "good" farmers are included will restrict the recommendation domain in the same way. One justification for using "good" farmers is that they set an example for their neighbours. Here this argument is at best weakly relevant because we are concerned with research not demonstration.

5. Choice of treatments and units

As noted in Section 4, the decision on the treatments and their layout within the farms, depends on the objectives of the study.

5.1 Choice of treatments

The same concepts of treatment structure are needed in participatory on-farm as in on-station trials:

- Treatments may be unstructured, e.g. genotypes;
- There may be a need for one, or more, control (or baseline) treatments;
- Factorial treatment structure remains important;
- The number and levels of quantitative factors have to be determined.

Below we concentrate on two points, where guidelines may be different for on-farm studies.

In participatory experiments the farmers may choose some of the treatments themselves. For example, varieties may be chosen from a village-level nursery, or from open days at a research institute for an experiment in the following season. This is sometimes done on a group basis, to arrive at a consensus for the trial. Alternatively it may be done on an individual basis, with a design that then has some treatments that differ from farm to farm. It may also result in some farms having more treatments than others. We would encourage this flexibility: the extent to which it is permitted will depend on the objectives of the trial. If the main objectives relate to yield differences, then some recommended varieties may be included in all fields, to which the individual farmers add further plots as they wish. Other objectives might imply greater, or less, freedom for individual farmers.

The second topic is that of the control treatments. These must be justified, as treatments in their own right. In on-farm trials the control is often the farmer's normal practice. Since this is likely to be different for each farmer, it cannot be regarded in the usual sense of "control", i.e. as a baseline treatment for the experiment as a whole, against which other treatment are compared. The farmer's normal practice will be useful as a baseline for each farmer, but the researcher may also wish to have a common baseline in addition.

In participatory trials, farmers may wish to use the concept of "controls" for their evaluation in a way that is broader than researchers are familiar with. For example, in a soil fertility experiment, they may request fertiliser on the control, on the grounds that the new technology should be as good as fertiliser. Or the enhanced treatment may

be planted on a poor part of the field, because it should bring the yield to the level of the rest of the field. The extent to which such suggestions are accepted depends on the agreed objectives of the research.

5.2 How many treatments?

We would not wish to see any rules prescribed here. The minimum may be one, when a single new variety is distributed in an adoption study. However, normally there are at least two treatments being compared.

We do not agree with the frequently-made statement that four treatments is in any sense a maximum number. The statement may be related to the general view that many on-station trials have only about eight to ten treatments, and that participatory on-farm experiments should be simpler than on-station trials. This logic is flawed for a number of reasons. The first is that often on-station trials could usefully have more treatments. The second is that experiments are time-consuming and costly, and it would often be wasteful to go to the effort of an on-farm trial with just three or four treatments. We suspect that, as real participation develops, the researcher/farmer group will often suggest more than four treatments for experimentation.

Set against the inclination to have a large number of treatments is the recognition that many treatments per farm usually implies complexity of the design, which may lead to partial failure of the trial. Where there is no simple solution, the design team should reassess the objectives. They could also consider splitting a complex study into simpler related experiments that may differ in their level of farmer participation.

5.3 How many treatments per farm?

In animal experiments, where there is only one animal per farm, or in fish farming, where farmers have only a single pond, it is only possible to study one treatment per farm. For further discussion of this problem see section 5.7. Otherwise there are usually at least two treatments per farm.

Questions which experimenters often ask include, "What do I do when there are more treatments to be investigated than there are plots in each farm?" "What if some farms have more plots available than others?" These are practical realities in on-farm experimentation, and the answer need not be to cut down the number of treatments, or restrict the experiment only to farmers who have a certain amount of land available to them. Care in the allocation of treatments within farms, at the design stage, can ensure a successful experiment is carried out in such circumstances.

5.4 Replication and resources

In designing an on-farm trial the researchers need to consider their resources carefully. For precise treatment comparisons there needs to be sufficient replication – but at what level? It is usually preferable to have more farms and fewer repeats of the same treatment per farm, rather than fewer farms and more replication within a farm. Consequently, in on-farm experiments, it is frequently the case that there are many farmers but each farmer has only one replicate of each treatment.

The problem with having no within-farm replication, is that the farmer-treatment interaction is then normally used as the random (or residual) variation. However, the treatment effects may really be different for the different farmers and understanding this interaction, e.g. which treatments are most effective for which types of farmers, may be an objective of the research. In such cases one would like to distinguish between the interaction and the residual, and having some within-farm replication is the obvious way to do this. This does not necessarily mean that there should be complete replication of all treatments within each farm, which would be wasteful of resources. We suggest instead that consideration be given to a design where each farmer repeats a single treatment. If there is a reasonable number of farms in the experiment then this should allow a valid subsequent analysis of the data. The replication should also be sufficient if the data are to be split into two or three subsets for analysis. The choice of which treatment is to be repeated is not critical. It could always be the same, most important, treatment, which is then estimated with greater precision. Alternatively, several (ten or more) farmers could each repeat one treatment, not necessarily the same one throughout.

If within-farm replication proves impossible, then it is still possible to carry out some investigation of the farmer-treatment interaction, provided there is information on the characteristics of the farms (see section 7).

5.5 Crop Experiments: Plot size

It is often assumed that the plot size should be larger for on-farm than for on-station trials. This is the result of past on-farm trials usually having been at the validation stage of the research. There is no general justification, on statistical grounds, for preferring large plots in on-farm trials. The most efficient use of a given area, or of a given number of trees, is normally achieved with more small plots rather than fewer larger plots, unless there is considerable lateral interference, as is the case e.g. with some agroforestry systems, or areas with nutrient or water movement. Normally, there is a balance between the preference of farmer and researcher for larger plots on the

basis of realism, or ease of treatment application, and the statistical benefit of improved precision from more, smaller plots.

The cases for realism when seeking farmer opinion regarding treatments, or when comparing treatments with regard to labour requirements are examples of compelling reasons for using large plots. However, the case for large plots should be made in relation to the objectives of the experiment; merely stating that on-farm experiments require large plots is not enough.

5.6 Crop Experiments: Plot layout

Layout of plots within each farm will primarily be guided by perceived or known variation within the farming area. The farmers' knowledge about the variation in their fields should be used to determine the location of the plots and any blocking scheme, and to avoid using particular patches of the field where necessary.

It is important to ensure that farmers and researchers are using the same criteria to define suitable locations. Researchers normally strive for homogeneity, while farmers may have particular parts of their field where they would like to try some treatments. For example, they may feel that addition of crop residues is most appropriate on degraded patches. Where large sections of the field are degraded, this can be accommodated within the design by putting all treatments on this type of land. Otherwise the liberty given to farmers will depend on the objectives of the trial. If farmers' opinions are of paramount importance, then the loss of randomness in the allocation of treatments to plots is of minor concern. The important sampling is at a higher level, namely in the choice of farmers. On the other hand, if a comparison of yields is an important part of the trial, then it is important to allocate treatments "fairly" (i.e. with some element of randomness) to the plots. In such a case, use of the degraded patches could be in addition to a replicate of the treatment on ordinary parts of the field.

Many practical considerations need to be taken into account when considering block and plot layout. In an on-station trial, for instance, a split-plot experiment may be carried out because it is convenient to plant one large area at one time, whilst the application of different levels of fertiliser can be on smaller areas. In on-farm trials these considerations may still apply. Another important practical aspect is the interview process. For instance if a farmer is to give an assessment of different varieties, where fertility is a secondary factor, then it may be convenient if the varieties are grouped together.

5.7 Livestock Studies: Units and replication

Most of the principles discussed earlier with respect to crop experiments also apply to on-farm livestock experiments. However there are particularities of livestock studies that require further consideration.

First is the definition of the experimental unit. In livestock experiments the experimental unit is most likely to be an individual animal, although there will be some instances where it is a group of animals. When designing an experiment, it is important to be clear on what constitutes the experimental unit, to ensure sufficient replication of the units. For instance in a vaccine trial where each animal is injected with a dose of vaccine, the experimental unit is the animal. In a poultry feeding trial where a brood of chicks eats from the same handful of feed, the experimental unit is the brood. The fact that measurements might then be made on each chick does not increase the amount of replication.

The second issue is that of blocking. In crop trials the natural blocking unit is the farmer. With livestock trials, breed, parity and age of the animals are also well defined blocking units. Some of these blocking units often need to be incorporated in the design, to construct a satisfactory experiment.

Linked to the issue of blocking is that of resources. Farmers will have different numbers of animals and hence, if an individual animal is the unit, this can involve a different number of treatments per farm. In addition, many farms may have only a single animal. The value of these farms, in an experiment, depends largely on the use of the other blocking factors, such as age and breed of the animal.

Finally, it is sometimes possible to investigate more than one treatment on an animal. Because of the variation that exists between animals, an appropriate design might be a cross-over design where each animal receives successive treatments over a period of time, and so acts as its own "control". In such designs, the order of the different treatments is varied for different groups of animals so the conclusions are unrelated to the sequence of treatments. Here the experimental unit is an animal for the period of time in which a single treatment is applied.

With livestock experiments there can therefore be several different levels of variation. There is variation between farms, between animals within a farm, and between periods within animals. Depending on the particular situation, treatments can be allocated at one or more of these different levels. Researchers who wish to do such studies, and who are in doubt on how best to design their experiment, should consult a statistician for advice.

6. Measurements

In participatory trials, we can distinguish between three types of measurement.

(i) Measurement of the type that are taken in on-station trials. These are usually yield components, time to flowering, milk yields, disease scores, etc.

(ii) Measurements of concomitant variables. These can be at a plot level, for example problems of waterlogging, or at a farm level, for example rainfall or soil type. Some variables, such as dates of sowing and weeding, and other management practices may be at either level.

(iii) Measurements of the farmers' opinions. These are from informal discussions or from questionnaires.

In on-farm trials, there is often still too great an emphasis on the first type of measurement, because the implicit assumption is that the methods of analysis will be the same as on-station. Whilst these data may still be of interest, we suggest that more attention be given to the collection of measurement types (ii) and (iii). The main reason for devoting time to the concomitant information is that we still need to try to understand the causes of as much of the variability as possible. In on-station trials the plots may be smaller and are likely to be more homogeneous. In on-farm trials there may be more variation within a farm than on-station and in addition there is variation between farms. In on-station trials there is a consistent management structure, whereas here there can be large differences in management practice between the farms. As a general guide, what is not controlled should be measured, both at the plot level and at the farm level, if it is of direct interest or if it might explain some of the variation in the data.

In general, the objectives of the trial determine what is to be measured. Thus the direct and concomitant measurements to be taken are normally decided at the planning stage. Often too much data is collected that is never analysed. Our encouragement to measure potential concomitant variables is not intended as support for the measurement of all possible data, just in case they may be useful.

In some trials, where farmers have chosen where they will apply particular treatments, there may be little reason for measuring yields. What is needed, instead, are the farmers' reasons for choosing a particular plot and their reactions at the end of the season. In less extreme situations there may still be little reason to devote much time to the detailed measurements of yield components. A quick assessment of yields using "number of bundles", plus some idea of harvest index, will often be sufficient.

A participatory trial is not really participatory, unless a record is kept of the important contributions made by the farmers. These may be the actual farmers who use the land, or others who view the fields. There are many ways of recording this type of data: please see the booklet *Guidelines for Planning Effective Surveys*.

7. Analysis

As with the design, the analysis of the data will use a mixture of methods that are appropriate for the analysis of experimental and survey data. The analysis can be viewed in three stages:

(i) Analysis of questionnaire-type data, resulting from interviews and other observations. This information is normally at the farmer level, though some questions can relate to particular plots.

(ii) Analysis of yield type data. This information is mainly at the plot level, though with some observations at the farm level.

(iii) Combination of (i) and (ii) above, using the results from interviews to understand the variation in yield type data.

The type of trial will dictate the proportion of time spent at each stage. One extreme might be a farmer designed and managed trial, within which, the main objectives relate to their choices and opinions. Most of the analysis effort would therefore be on (i) above. In some researcher-designed and managed trials the yield data is of particular importance, in which case most of the time is spent on (ii).

Experiments with sufficient within-site replication and detailed measurements of yield response can have separate within-site analyses initially, then a combined analysis. This is usually only the case for researcher-designed and managed trials. Others will use the data within a single analysis. However there are two main differences between on-station and on-farm trials that have a bearing on the analysis. One is that with on-farm trials we expect a farm by treatment interaction, and one of the objectives of the trial is often to explore this interaction. The other difference is that there is now variation at different levels - there is variation between farms because of characteristics such as different agroclimatic conditions, management practices, as well as variation between plots within farms. As always, any analysis should try to explain as much of the variation as possible.

Approaches used can range from some simple analyses on different subgroups of the data to more sophisticated modelling of the whole data set. The analysis is often to evaluate relationships between biophysical responses and environmental, management and social variables. The data are also used to understand reasons for farmer assessments. These may be turned into decision trees for farmers or maps of recommendation domains. In analysing on-farm trials data we should be ready to:

(i) split the data up into subsets, e.g. groups of similar farms;

(ii) omit particular plots, e.g. the farmer's own treatment; or particular farms;

(iii) pay close attention to comments made about individual plots, e.g. "crop eaten by animals" may mean that a recorded yield of zero should be treated as a missing value;

(iv) use additional information, both at the farmer level and at the plot level, e.g. farmers may be classified as wealthy or poor, or plots may have information about pest damage;

(v) in the absence of within-farm replication, use treatment contrasts at the farm level to investigate the farm by treatment interaction, or investigate the interaction using the additional farmer information (as in (iv) above);

(vi) report on, and possibly follow up on, particular farmers who show interesting results.

Data recorded on questionnaires, such as preference for varieties, can be summarised in two-way (or n-way) tables of response by farm type. Percentages can be presented if the total number of farmers is large enough. Provided there are sufficient data, models can also be fitted to these tabulated data to explore how the responses vary across different farm types.

8. Further topics

The analysis is not the end of a study, but it does conclude those aspects that we feel merit discussion as different from other research exercises. The results have then to be reported: our guide called *Informative Presentation of Tables, Graphs and Statistics* may help. Reporting must be to all interested parties, with the collaborating farmers having high priority.

The details of the study and the data must then be archived: the booklet *Data Management Guidelines for Experimental Projects* gives some advice.

The Statistical Services Centre is attached to the Department of Applied Statistics at The University of Reading, UK, and undertakes training and consultancy work on a non-profit-making basis for clients outside the University.

These statistical guides were originally written as part of a contract with DFID to give guidance to research and support staff working on DFID Natural Resources projects.

The available titles are listed below.

- *Statistical Guidelines for Natural Resources Projects*
- *On-Farm Trials – Some Biometric Guidelines*
- *Data Management Guidelines for Experimental Projects*
- *Guidelines for Planning Effective Surveys*
- *Project Data Archiving – Lessons from a Case Study*
- *Informative Presentation of Tables, Graphs and Statistics*
- *Concepts Underlying the Design of Experiments*
- *One Animal per Farm?*
- *Disciplined Use of Spreadsheets for Data Entry*
- *The Role of a Database Package for Research Projects*
- *Excel for Statistics: Tips and Warnings*
- *The Statistical Background to ANOVA*
- *Moving on from MSTAT (to Genstat)*
- *Some Basic Ideas of Sampling*
- *Modern Methods of Analysis*
- *Confidence & Significance: Key Concepts of Inferential Statistics*
- *Modern Approaches to the Analysis of Experimental Data*
- *Approaches to the Analysis of Survey Data*
- *Mixed Models and Multilevel Data Structures in Agriculture*

The guides are available in both printed and computer-readable form. For copies or for further information about the SSC, please use the contact details given below.

Statistical Services Centre, University of Reading
P.O. Box 240, Reading, RG6 6FN United Kingdom

tel: SSC Administration +44 118 378 8025
fax: +44 118 378 8458
e-mail: statistics@lists.reading.ac.uk
web: http://www.reading.ac.uk/ssc/