MET Design Guide:

A Checklist

Ric Coe (r.coe@cgiar.org), Statistical Services Centre, University of Reading, UK and World Agroforestry Centre, Kenya.

7 April 2012

1. Introduction

In some ways, designing and planning a multi-environment trial (MET) in agriculture is like designing and planning any other experiment. There is a lot of literature available and a lot of collective experience in research organisations. And yet things often go a little wrong in the design of an MET, meaning that the trial turns out to be less useful than it should.

This document is organised as a checklist to take you through points to think about when designing an MET. As you plan, run through it from time to time to make sure you have thought of all the points. And then document them in protocol – see Section 17.

You will need to be familiar with the key concepts of METs before you try to design a specific trial. Start with Multi-Environment Trials: An Overview. Then get help from others, such as researchers who have done similar studies to yours, and a statistician with suitable interests and experience.

2. Objectives

- Have you written down clear and precise objectives for the trial? The whole of the rest of the design depends on the objectives so it is worth getting them right! Try to avoid objectives which read 'To compare treatments in different locations'. The treatments and locations are chosen to meet the objectives, not the other way around.
- Base objectives on your previous work, your knowledge of the problem and context and, particularly, the literature. There is no point in doing a trial to prove something that is already established, or to pursue something known to be impossible.
- Have you included hypotheses? Hypotheses help to keep focus. The hypotheses should refer to the environments or contexts as well as the options or treatments.
- Do your objectives and hypotheses refer to AEI levers and principles?
- Do your objectives lead to clear choices of both the treatments or options and the environments or contexts?
- Note that the objectives will probably be very different for different styles of MET. For example, breeders may be interested in selecting the 'best' lines for further development and use a controlled, on-station experiment. But farmers and those working with them may be interested in selecting a range of options suited to different purposes and conditions, and will need to make assessments on farm.
- Your MET will probably have multiple objectives. Are you sure they can all be addressed in a single trial? Often objectives concerned with biophysical responses and processes require a different trial from those concerned with farmers' assessments.
- Have you set the overall limits to your domain of interest? In one MET you will not be able to answer questions for the whole world. You need to be clear of the limits to the domain, or set of environments, that you want to get results for. This could be a geographical area (e.g. 'the lake zone of Uganda') or a geographical

area and some further biophysical, social or economic filters (e.g. ‘Farms in the lake zone of Uganda, of less than 3 hectares with bananas the main crop’).

3. Who

- Who is involved in the trial? List those people and organisations that will be involved in the design, management and analysis of the whole process. Do they all know they are involved? Have they been consulted?
- Who else has an interest in the research and its outcomes? This could include farmers, extension workers, traders, and consumers. How have they influenced your thinking? Have they had an opportunity to make suggestions on the aim, design, management, measurement, analysis and follow up to the trial?
- One feature of an MET is that it needs good coordination across sites. All methods at different locations must be the same or comparable. How will this be achieved? Who is responsible at each site? How will you ensure that they understand their responsibilities? Do you have agreements in writing? Written agreements are particularly important when the MET involves people from different organisations as these might have different standard practices and values.
- Will the MET involve farmers and communities? What is the process for community engagement – how will it be done and at what stage of the design? Don’t expect to be able to impose a design on farmers and get useful results if they have been involved in developing it!

4. Ethics

- There are ethical considerations in any research. The complexity of an MET means these might be deeper than for other trials. MET that involve human subjects (e.g. child feeding experiments) require by law in most countries) specific ethical clearance. But any MET has some ethical issues.
- Have you obtained prior informed consent from participants? If the experiment involves people (such as farmers) then you must. This is necessary even if they are only involved to the extent of loaning land for a season. ‘Prior informed consent’ means that before any commitment, participants have been informed exactly what will happen, have understood and explicitly agreed to it. They must be able to decline without any negative consequences. It includes understanding their rights and responsibilities at all stages of the trial. See, for example, the Code of Standards and Ethics for Survey Research (http://ethics.iit.edu/ecodes/node/3785)
- Are you going to collect any information about identifiable individuals? If so you have a duty to keep it confidential. Explain how you will do this.
- Do you have an explicit and written data sharing agreement in place? If the trial involves people from several organisations, or even several scientists within the same organisation, you need a written agreement on who has access to the data, when and for what purposes.
- Similarly you need an explicit agreement over writing up results and authorship. Get help drafting such agreements. The right time to do it is as you start design of the trial, not once a conflict arises.
- Do you have a common understanding among all involved on how information from the trial, including raw data, will be shared and made available to others?

Open access to results, data and protocols is becoming standard practice for publically funded research and requires explicit steps to make it happen. It will also require agreement from all partners.

5. Treatment or option choice

- The treatments or options are the things being compared in each environment. Have you been able to define them unambiguously, in a way that makes sense in each environment? They should be a direct consequence of your objectives.
- Are you familiar with the usual principles of treatment choice - contrast, control, factorial structure, quantitative levels? See for example GEAR Chapter 3.5.
- Are the treatments defined in a way that really makes comparison across sites possible? For example, if a treatment factor is level of N fertilizer input, they could be defined as (a) absolute levels in $kgha^{-1}$, (b) relative to the soil N content at each site, (c) relative to an estimated optimal rate at each site based on yield potential, or (d) based on previous available recommendations at each site. Which you choose should depend on exactly what it is you hope to find out.
- Variety trials often include a 'Local control'. How is that defined at each site? Should it be constant across sites or could it be a different variety at each location? Who chooses?
- In participatory trials, farmers sometimes have a choice of which subset of options they will test. Who is choosing treatments? If participants are choosing a subset of treatments, will you record reasons for the choice? If participants select treatments, then you may want to supplement those selected with your own suggestions.
- The set of treatments does not have to be identical in every environment. But deviations should be for a reason!

6. Environment or context choice

- The environments or contexts selected for the trial should be based on the objectives. It is unlikely that you will always be able to use the same selection of research sites. Think about what you really need, rather than limiting yourself to the usual places.
- How are you defining the environments? They should be defined based on objectives and hypotheses. For example, they could be defined by biophysical variables (e.g. length of growing season, soil P status), social variables (e.g. ethnic group, resource endowment) or economic variables (e.g. distance to market, level of off-farm income).
- How will you find out the range of values of these E variables within your domain of interest?
    - Some can be mapped (e.g. length of growing season, main ethnic group, and distance to market). Data layers probably exist and you can use GIS to show them on maps. Then, aim to select sites that cover the important gradients.
        - For continuous E dimensions (e.g. length of growing season), the same rules apply as for selecting quantitative treatments - aim for a

    few levels that span the range, and allow replication of sites at each.
    - If there are two or more E variables that can be mapped, then use the ideas of factorial experiments to be able to separate them. For example, if using both length of growing season and distance to market, you will probably want to select sites that have all combinations of short and long season with near and far markets.
  - Some cannot be mapped because they vary over short distances (e.g. soil P, farm resource endowment, and off-farm income), or mapped data of suitable resolution do not exist. Then you will need survey data (secondary data or your own baseline surveys) to understand the important ranges and where you might find sites or farms to use in your MET.
- Note that the idea of a site being 'representative' for some area is flawed unless you say in what way it is representative – in terms of rainfall pattern, for example. The only way you can be sure a set of sites is representative is by random sampling. That is often not feasible. But some studies do manage to use random selections, and can then make statistically valid statements about the extent to which results apply to larger areas.
- But, environments that are distinctly 'unrepresentative' should be avoided! For example, if selecting varieties for farmers who do not use fertilizer, there is generally no point selecting in a well-fertilized site.
- If looking for interactions with a particular stress you may get better information by taking some extreme rather than 'typical' or representative environments. For example, if looking for drought tolerance, then select an environment with appropriate levels of drought stress. The same concepts as those used in choosing levels of quantitative treatment factors apply.
- If you want to select environments that vary in a planned way (e.g. to have length of growing season of 80, 100 and 120 days), you need to decide whether to try to standardise for other characteristics. For example, if soils in the 80 day zone tend to be sandier than those in the 120 day zone, should we include those sandy sites (and confound growing season length with soil texture)? Or should we look for low-sand sites in the 80 day zone (and risk having a site that is atypical for the 80 day zone)? Again, the answer depends on specific objectives.

7. Farmers and groups who participate

- Is your MET a participatory trial that involves farmers? There is a large body of literature on farmer participatory trials that is relevant.
- Are you thinking of each farm as an environment? This will probably be necessary if the interactions that you are interested in are with farm- or farmer-level variables. For example, farmer's assessment of an option often depends on a farm typology based on resource endowment and farmer objective. If the farm is an environment, then all the points of Section 6 are relevant.
- Maybe it is more useful to think about the 'type' of farm or farmer as an environment. In the end, we want information about and for a larger population of farmers than those in the trial, rather than focusing on particular individuals. This comes back to objectives (Section 2) and has implications for replication (Section 8).

- When characterising farmers to invite as participants, or those that are self-selected, do so in terms which are relevant to your context and objectives. The usual variables of farmer age and years of education rarely have much to do with farmer x option interactions!
- How are farmers selected to take part in your MET? Self-selection is common and often necessary. For example, the project works with a community (perhaps through a community-based organization (CBO) or farmer organization). The MET is discussed and farmers ask to join in. Such self-selection will give a sample of farmers that might differ in important ways from a cross-section of the community. You cannot avoid that. But you can try to (a) identify the factors that might affect farmer management or assessment of the options, (b) find how those are distributed in the community, (c) check that you have a good representation among the participants and (d) actively seek participation to fill gaps. For example, we might hypothesize that older women farmers have a different opinion of desirable traits in a PVS trial. But only one volunteered to take part. So you would ask for more older women who might agree to join the activity. If that is impossible, you have to be open about the limitations to your results: they do not represent the situation for older women farmers.
- Have you prepared a detailed protocol for the way you will interact with farmers? This is necessary if you want (a) realistic results and (b) comparable results across sites. We know that the enthusiasm farmers bring to taking part in an experiment, and the way they assess results, is highly dependent on the interaction process. Be particularly careful of perks and benefits given to participating farmers.

8. Units, replication and layout

- What are the primary experimental units in your MET? Usually these will be plots in fields, but the MET design ideas discussed here are relevant to other types of investigation.
- Have you defined the size and shape? Usual guidelines apply for choice of size and shape (e.g. too large: expensive, cannot fit into homogeneous land area. Too small: cannot manage or measure in a realistic way, edge effects dominate, high between-plot variation).
- Sometimes an MET may involve other types of units – households, farms or community groups for example – see Multi-Environment Trials: An Overview. They still need defining carefully (e.g. Just what is a 'household'?) and in a way that makes sense across the environments. The primary unit is defined by the objectives and hypotheses. If the hypothesis is about communities, then the community is your unit, not people within the community.
- Most designs have a hierarchy of units such as plots within fields within farms within landscapes, or within communities within agroecozones. What is the hierarchy in your design? Do you have an objective way of selecting or defining the units that will occur at each level in the hierarchy?
- At each level in the hierarchy you also need to decide how many units to use. This is the question of replication. Are you sure you understand what is meant by replication and reasons for replicating? There are four of them. The most critical for a MET is the need to find effects that appear consistently or repeatedly. See (GEAR Chapter 3.5). Two examples:

    - A breeder hypothesises that shallow-rooted beans will perform better than deep rooted beans on low P acid soils, but the reverse will be true on high P soils. A number of deep and shallow rooted lines are to be tested. Do we need one low P and one high P site? Or do they need replicating? There will inevitably be many differences between the sites other than soil P. Hence to draw the conclusion that it is the soil P which is the important difference you need several low P and several high P sites that do not differ in any other systematic way. In other words, in this case the environments (low and high soil P) need replicating.
    - Farmers are evaluating how well six different legumes meet their requirements in an on-farm trial. The researcher expects favoured options to depend on farmer's resource endowment. Do we need replicates of each treatment within each farm? Each farmer is interested in results for their farm. They may choose to try the options in several different niches on their own farm. But researchers are interested in types of farm, and hence need to have several examples of each type to check for consistency of differences. It is farms within types that need replicating, not plots within farms.
- Are you able to manage the unwanted variation? Some methods available:
    - Exclude from the design any units (at any level in the hierarchy) that are clearly atypical and do not represent the populations or conditions you are interested in.
    - Manage all aspects of the trial in a uniform way, both within and across environments. For trials with farmers, this includes standardising the way you interact with farmers – for example, in the way you introduce the trial ideas to them.
    - Use blocking at any level in the hierarchy of units. The ideas of incomplete block designs are useful when there are large numbers of treatments.
    - Some designs (e.g. the 'babies' in mother and baby experiments) use a large number of treatments spread over several farms, so each farm gets a subset of the complete treatment set. Use incomplete block design ideas to choose the subsets. Get help if needed.

9. Randomisation

- Are you able to randomise treatments to units? You should use random allocation (within blocks, if using them). The reasons are the same as for any experiment (See GEAR Chapter 3.5).
- Each site or environment should have its own randomisation.
- The only time it may not be feasible to randomise is in farmer participatory trials in which the farmers choose where to plant each option they are testing. Explain the virtues of random choice.
- When collecting data from a trial in which farmers chose where to plant each option, include reasons for the allocation. This can make a very important difference to interpretation of results.

10. Management

- As in any experiment, much of the quality of an MET is determined by how well it is managed.
- For a crop experiment this means:
    - Ensuring all crop (etc…) management is appropriate – land preparation, planting, pest control, etc.
    - Ensuring management that should be uniform across plots and sites really is uniform.
- Have you decided which management practices will be constant for all environments and which ones will be adapted to be locally optimal? For example, it will rarely make sense for planting date to be the same calendar date at each site. Planting date should be determined by something objective that can be applied at each site – e.g. two days after the first significant rain, or when most farmers have started planting.
- Remember that if management varies by site, then site and management effects will be confounded. For example, groundnut agronomy recommends ridging in dry locations and flat planting elsewhere. If a variety trial is done at a dry site with ridging and a wetter site without, the climate and agronomic practice are confounded.
- For trials with farmers, discuss with them and reach a common understanding of which management factors will be uniform across farms, and which will be adapted by individuals to their own preference.

11. Monitoring

- How will you monitor the MET? There are two types of monitoring required:
    1. Routine monitoring for planned data collection and observation of compliance, decisions on management (e.g. pest control) and to keep farmers engaged.
    2. Monitoring to look for surprises.
- Looking for surprises has to be done by someone who understands the research, the crop and the context. That probably means…you! Look for:
    - Odd and unexpected behaviour of units or options.
    - High levels of variation.
    - Surprising comments and observations from farmers and field staff.
- Be prepared to initiate additional data collection to help understand unexpected patterns. For example, if you see very wide variation in performance of options on neighbouring farms then:
    - Come up with some suggestions (hypotheses) of the reasons. These should be based on your knowledge of the crop and on farmer's or field staff understanding.
    - Collected data that will help confirm the hypotheses.

12. Measurement

- Have you defined all the measurements you will take? The measurements are again defined by the objectives, but in an MET there may be a lot of things to measure. Try to list them systematically.

- Start by completing a measurement table, such as the one below, that lists all the quantities that need measuring at different levels in the hierarchy. As a minimum there will be two levels: the plot and environment. There may be more, as in this example:

Table 1: Organising think about measurements

	When to measure		
Unit	Before	During	At harvest
Plot	Soil texture	Weed level Disease occurrence Observations and notes	Crop yield Farmer's assessment of crop
Farm	Farm type	Observations and notes	
Site	Average rainfall	Actual rainfall Observations and notes	Results from group discussion on the trial

- Then for each measurement list:
    - Who will collect the data?
    - What measurement instrument and protocol will be used?
    - What equipment is needed?
    - What are the time and workload requirements?
- Carefully consider workload, workflow and logistics. Can the relevant staff get to all sites at the required time? How will they know a farmer's field is ready for harvest? How many samples can be dried or analysed in the lab at once? How long will it take to harvest and measure one plot? Etc. Is it all feasible?
- Remember the challenge with METs is to ensure the measurements are consistent and comparable across sites.
- 'Observations and notes' are listed to be recorded at each level during the MET. This means keeping careful records of what actually happens and what is observed, including observations by field staff and farmers. These could be the key to a good interpretation of the data.
- Will you record video and photos? It is cheap and easy to take photos of all plots. These can be an important record of what happened in the field. They should be a standard component of your 'observations and notes.' They are particularly important in METs because you, as the researcher, may not be able to visit all field plots across all sites during the trial.
- All measurement methods will need to be piloted and field staff trained in their use.

13. Data management

- Do you have well documented procedures for managing all data? See Data Management for Multi-Experiments Trials in Excel.

14. Analysis

- Do you have a clear plan of analyse methods and how to get analyses done? See the document An Introductory Analysis of Multi-Environment Trials (METs).

15. Reporting and feedback

- Do you have a reporting plan? This should include:
    - A complete report of what was done along with tabulated results.
    - A write up of your conclusions for the project, and next project steps.
    - One or more scientific papers.
    - A report for farmers and other participants. This must be delivered to them in a format that is acceptable, such as a verbal report at a meeting. Failure to feed results back to participants is a serious offence.

16. Quality assurance

- Throughout the experiment, pay attention to quality assurance. If, at the end of the trial, you present results and someone asks 'How do you know those numbers are correct?', you need to have an answer!

17. The protocol

- All this information should be written up into a protocol.
- The protocol should be drafted well in advance of the start of field work so that you have time to share it and get comments, as well as time to fully prepare for every aspect of your MET.
- The protocol should be shared for comment with:
    - Partners involved in implementing the trial at each sites
    - Scientists with experience in similar work
    - Your project leader or supervisor
    - A statistician or research methods specialist
- Keep the protocol up-to-date. As you start work on the experiment, there will be details that need changing. Document them carefully so that at the end of the experiment you have a record of exactly what was done.
- Store the protocol with the data. The data files are useless without it. Data analysis cannot be done without the protocol.

18. Further reading

GEAR – Graduate Environmental and Agricultural Research