An Introductory Analysis of Multi-Environment Trials (METs) and Large N Trials in Agricultural Research

Originally written by:
Ric Coe (r.coe@stats4sd.org), University of Reading, UK and World Agroforestry Centre, Kenya. 5 April 2012

Updated 31 July 2017:

Sam Dumble (s.dumble@stats4sd.org), Statistics for Sustainable Development, Reading, UK

Update Note – July 2017
This document provides an overview of various techniques to make the most of your data so that you can go beyond presenting simple averages and p-values from ANOVA tables.
This document was originally produced purely to reference multi-environment field trials (METs). There are many overlaps between the analysis required for METs and the analysis required from 'Large N' trials. Indeed most 'Large N' designs would be considered a form of a multi-environment trial.

There is no detailed methodological discussion around any of the examples given, but references are provided if you want to learn more. Whenever you see a graph or table with a numbered reference this indicates that a practical explanation of how to replicate this analysis is provided in the appropriately numbered appendices for either Microsoft Excel or R.

Introduction

Analysing data from an MET or large N trial requires the same general strategies, and is at risk from the same problems and protected by the same good practice as analysis of other data. Some useful guides are listed below. However, METs also have some specific analysis challenges. They can be complex and have their own specific objectives for which analysis methods have been developed. This brief overview is designed to explain some of these special concerns and means of addressing them, and refers to the following documents in this series:

- Multi-Environment Trials: An Overview
- Multi-Environment Trials: Data Quality Guide
- Data Management for Multi-Experiments Trials in Excel
- An Easy Introduction to Biplots for Multi-Environment Trials
- MET Design Guide: A Checklist

The hardest part of analysis can be thinking through the logic of what you are trying to do. What information are you trying to extract? What will the result look like? Can you draw up outlines of the tables or graphs that would tell you what you want to know before you start looking at the numbers? Does the data you have really contain this information? Once you have these ideas clear, then the statistical analysis is often not that complex. You can always get help with the statistics, but the logic and outcome of the analysis is your responsibility as a scientist.

METs are useful in a wide range of contexts. They may have social or biophysical objectives. They may be participatory in different ways. An MET could be everything from a very controlled on-station variety selection trial, up to a participatory Large N trial. But they will all require the same basic approaches and methods for analysis. This is even true if the data are qualitative though some of the methods may look different.

A successful analysis meets the objectives, is valid and provides insights; it will be tuned to the specifics of your study and what you want to find out. There is (almost) never just one analysis that can be done for any dataset. Unless the trial was small and simple, there will be many questions that can be answered with the data, each with its own analysis approach. Rich data sets will not only be analysed and reported once,

but a good analysis will also be an iterative and learning process. While you should start with a clear goal, you also need to be alert to surprising patterns or new ideas that emerge from the data. Some step of analysis may suggest that your original plans could be improved, and you go back and try again. The richness of data from an MET that can be gleaned through repeated analysis means the data and documentation have to be carefully archived. See the MET data management guide

- *Data Management for Multi-Experiments Trials in Excel.*

The nature of analysis of METs or large N trials means it should be conducted by the scientists who planned and are responsible for the work. It cannot usually be delegated to a technician or statistician, neither of whom usually have the scientific background needed. You should probably be getting expert assistance from someone with specialist statistics skills. Then the analysis should proceed as a dialogue between you as a scientist and the statistician. If there are others directly involved in the trial and interested in the results, then it could be a three-way dialogue. This can include dialogue with farmers. Farmers who managed a trial on their own land or provided you with data are probably the best ones to give you insights into surprising patterns (and if there are no surprising results, have you really learnt anything from the trial?). Notes taken from formal interviews, or even informal discussions, with farmers themselves provide a valuable source of information and should be consideration when reporting the results. They can provide explanations that can then be confirmed with further analysis.

A good analysis of data from these trials will almost certainly require you to use statistical software. Excel alone is not adequate; although is still a vital tool for initial data exploration! Many general statistical analysis systems will be suitable. In accompanying documents we use just two other software packages, R and QGIS. These two are selected for good reasons:

1. They are well suited to their tasks.
2. They are both completely free and open source, so are unlikely to ever transition into commercial software (as has happened in the past with Genstat Discovery)
3. There is good documentation and training material available.

You may also find some specialist statistical software useful, though we will not use any here. Once you get beyond the statistical analysis to using and interpreting the results, you may find other software useful, such as crop models.

The rest of this guide is presented as bullet points for easy reading, and can be used as a checklist or reminder to pay attention to as you analyse your data.

1. Objectives

- Every part of the analysis depends on your objectives. List them clearly and precisely.
- The objectives will be based on:
    - The objectives that you set when designing the study; and
    - Ideas you you developed since designing the study, particularly from observations made in the field during data collection.

- Get specific. The more detail you can give the easier the analysis becomes. Try to sketch out the key tables and graphs that will give you the information you need.
- Set objectives that answer research questions (e.g. 'Find out if any genotypes are widely adapted') rather than objectives that specify a type of analysis (e.g. 'Do a stability analysis').
- Many objectives concern interaction between options and context, or concern risk. Make sure you understand these concepts. See the Overview document Multi-Environment Trials: An Overview.
- Remember there may be multiple objectives, with the objective driven by your scientific interests perhaps needing to be integrated with farmers' objectives. For example, farmers may not be interested in the 'the best' but a range of options with a range of properties.

2. Prepare the data

- Before analysis can start you will have to prepare the data, which involves getting data files checked, complied and formatted for analysis.
- In principle, this step is the same for as for any other experiment. But the added complexity of the multiple environments and large quantity of data can make it much harder in practice.
- The combination across sites often means a combination of data collected, entered, organised and 'owned' by different people. Hence, there is considerable opportunity for the process to be complex and messy.

Refer to the guide in managing data from METs
Data Management for Multi-Experiments Trials in Excel

- for suggestions on good practice and avoiding problems.

There is a standard 'long format' for data from multiple sites that is being prepared in Excel for analysis. Make sure you understand what that is, and that your data comply.
See the note on Excel for METs
Data Management for Multi-Experiments Trials in Excel

- When all the data have been compiled, check datasets carefully. Do you have what you would expect? Do you have the right number of observations for each treatment and from each site?

3. Farmers and analysis

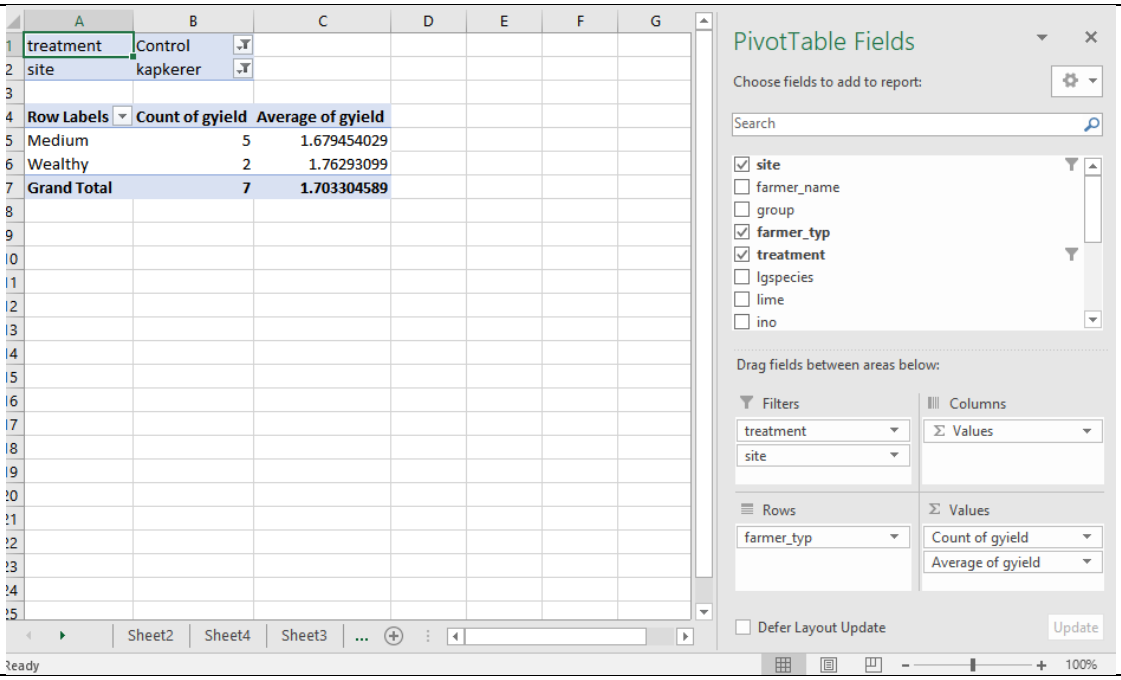
- We tend to think of analysis of data as strictly being a task for scientists. But any participants, including farmers, can have a role in analysis.
- Farmers will probably not be interested in or able to contribute much to the formal, numerical and statistical analysis. But they can certainly contribute to some of the aims. For example, they can help:
    - Explain why the main patterns are as they are – for example, explaining that option A suffered more from drought while option B was prone to pests.
    - Explain apparent discrepancies or oddities in the data. For example, they might tell you why plots were neglected on some farms, why

crops did very well in some niches, or why women rate varieties differed from men.

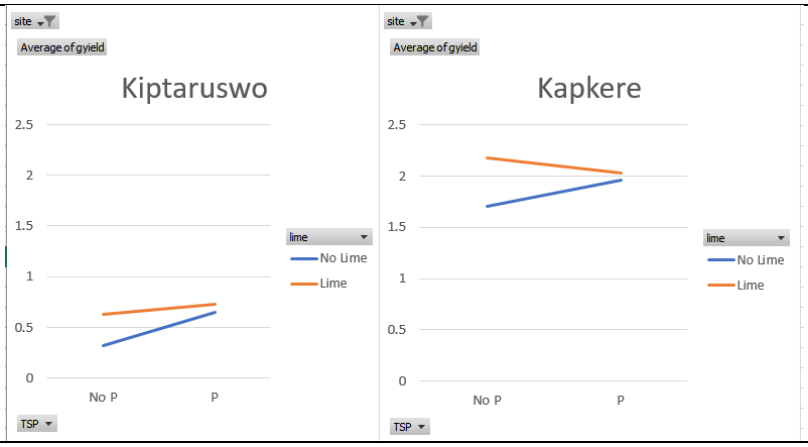
Your challenge is to find accessible ways of presenting collective results that farmers can understand. This will allow them to put the experience on their farm into the broader context of what others experienced.

4. Exploratory analysis

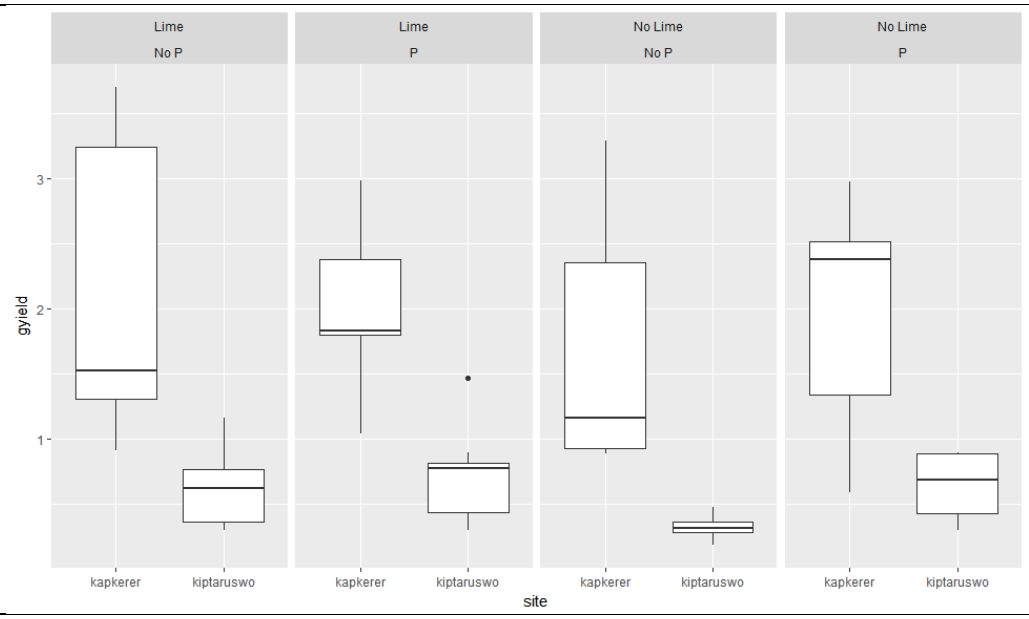
- In the exploratory stage of analysis, calculate the summaries and present them in tables and graphs that will start helping you answer your questions and meet objectives.
- This step is less dependent on statistical methods as on your imagination in thinking of the displays or presentations that will answer your questions and linking them to the data available.
- Design them based on what you want to know, not on what you know how to generate.
- A few simple example are shown here:



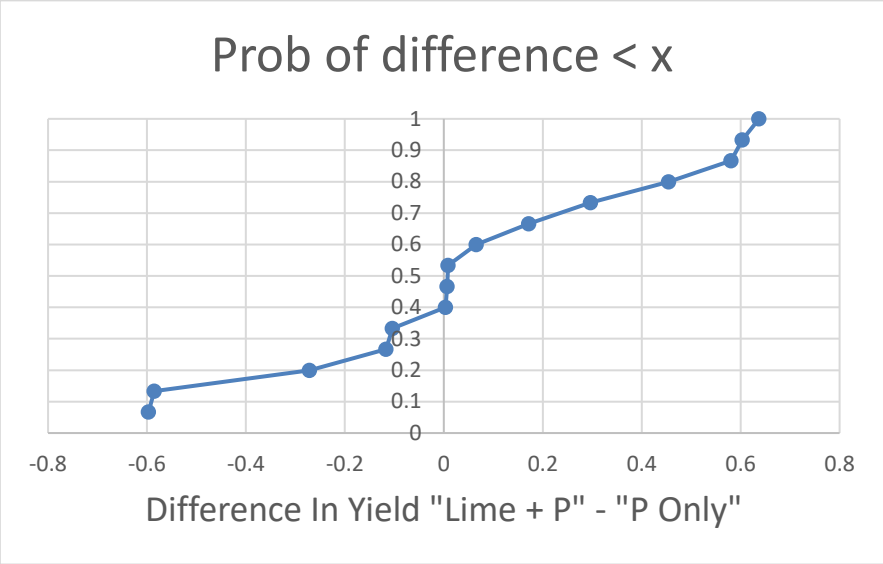
PivotTables in Excel are an easy way to interactively explore and understand relationships between variables with your data. In this case an experiment was done with some farmers, with each farmer using the same 4 different fertiliser treatments, including a control group. Looking at the relationship between an assessment of a farmer's wealth status and crop yields for just the control plots within a single location, does not show a big difference in the average values of yield when split by wealth. Although there are only a small number of farmers included in this trial. Using the filters, and the row and column structure can allow us to investigate many different relationships within our data.



The graph presents the interaction plot at each location. The x-axis is for ±P, y-axis is crop yield, orange and blue lines are ± lime and each frame is a site. It shows there is an interaction at both sites, but it is not the expected interaction and it is not quite the same at both sites. P fertilizer is less effective with lime than without, the opposite of what was expected. And in the second site Lime + P is worse on average than Lime Only. This plot also helps to emphasise how different the yields in the two sites.



Boxplots are a way of exploring the variability in the data as well as the averages; using the data from the same example as the line plots shows extremely large amounts of variability in the Kapkere site relative to the Kiptaruswo site. The plot also clearly shows the benefits of using either lime or P in Kiptaruswo, given all farmers are getting similar, low, yields (tightly packed box at bottom of scale) where neither is used. The yields where either lime or P or both is used look to be higher for all farmers. The benefits are much less clear for Kapkere; some farmers are already getting good yields without any fertiliser.



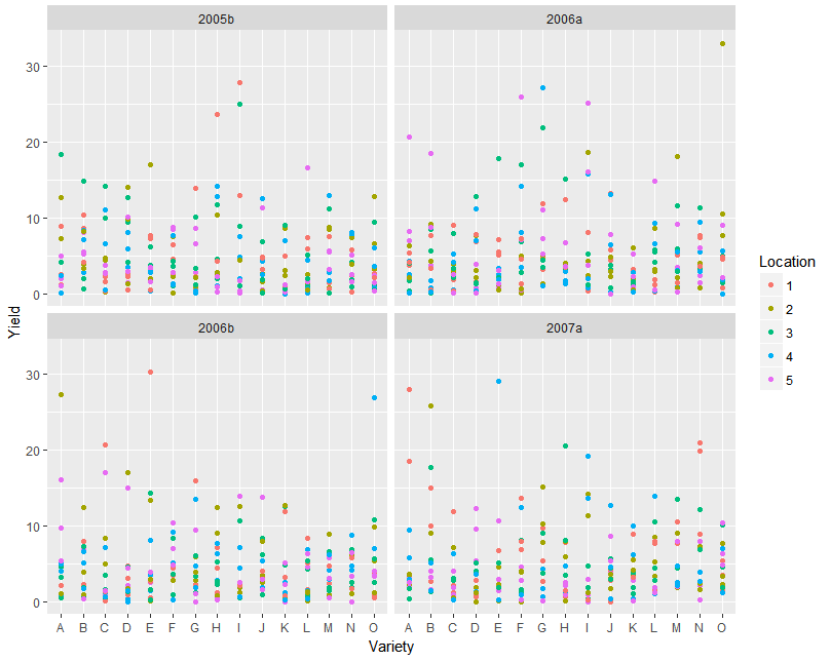
From the same trial – looking at the comparison between Lime + P and P Only from 15 farmers. The cumulative distribution is plotted, showing about 40% of farmers had a yield loss with adding lime to P. For some farmers the difference was almost exactly zero; for others it resulted in a yield gain of 0.6 t/ha; whilst for others a yield reduction of 0.6 t/ha. The next step is to explore what might explain this wide variation

Plotting the data from each treatment in a scatter plot can also help to explain variability. Each point represents one farmer with the x value being the yield from the P-Only treatment and the y value being the yield from the Lime + P treatment. The blue line is a smoother fit to help show the trends. In this case we might hypothesise that adding lime on top of P is effective if yield was <2 t/ha; and is a hindrance if yield was >2 t/ha. However, in this case the number of farmers is still quite small, so caution is needed when interpreting these results. But this approach may start to help generate future hypotheses.

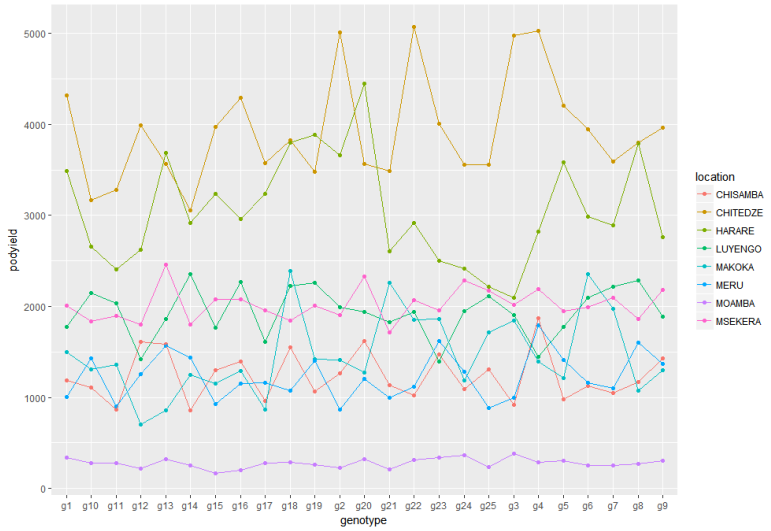
- Exploratory analysis will reveal:
    - Odd observations that need investigation. Check for errors and refer to field notes for explanations. Correct errors. Omit odd data points if they you know why they are odd and that they do not present what you

wanted to measure. Do not simply omit any data point that fails to meet your expectations.

- Note that spotting odd observations will require some of your plots to show original data rather than just means.
- Unexpected patterns that might make you go back and rethink your analysis objectives. See the first example below.

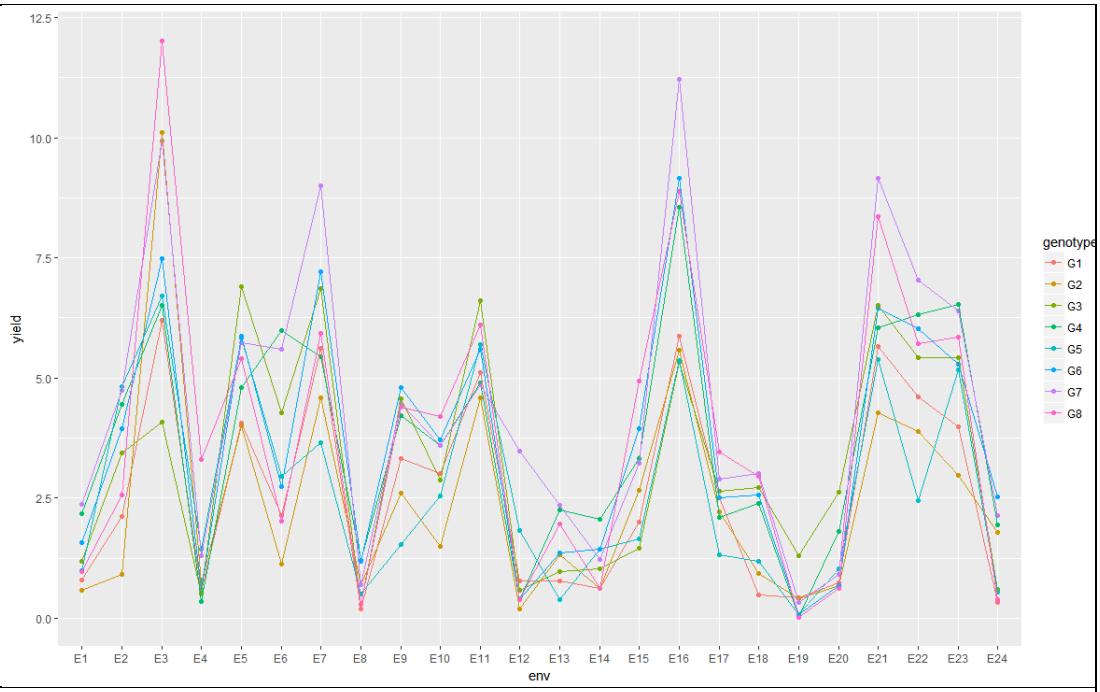


A plot of raw data from a variety trial showing yield (y-axis) for a number of varieties (x-axis) for four seasons (separate graphs) and 5 locations (colours). A large amount of raw data is shown in a compact style. The key observation is the very high variation in yield between multiple observations of the same clone in the same site and season. First step: check the data. If it is confirmed, then analysis should focus on explaining the source of that variation.



Yields (y-axis) of 25 varieties (x-axis) across 8 locations (colours). One location (in purple) has such consistently low yields for all varieties that it probably does not make sense to include it in a GxE analysis with the other sites.

- Sometimes exploratory analysis will show that results in one (or more) environments are so different from the rest that including it in a common analysis does not make sense. See the second example above.
- Sometimes simple summaries, tables and graphs cannot pull out the main results, and more sophisticated exploratory methods are needed. See the example below and the section on GxE.



Performance of 8 genotypes of a crop across 24 environments. In principle the GxE interactions can be inferred from the graph. But the number of points and lines makes it hard to spot patterns such as which environments are similar or whether some genotypes are particularly well adapted to a group of environments.

5. Confirmatory analysis

- By 'confirmatory analysis' we mean the formal statistical procedures that allow you to confirm that patterns revealed by exploratory analysis are 'real'.
- By 'real' we mean they are repeatable, or large compared to the uncertainty. This is shown by calculating standard errors or confidence intervals for critical quantities. Results of statistical hypothesis tests may also be useful sometimes, though these are often overused and incorrectly interpreted. See Confidence and Significance. [link].
- You do need to be familiar with statistical concepts and methods to do this part of the analysis, and you will need appropriate statistical software.
- There is no single correct analysis for an MET. The methods to use will depend on:
    - Your specific objectives;
    - the design; and
    - the nature of the data.

- For experiments with replicated designs in each environment, it is usually wise to start with a separate analysis first to check for oddities and look for common patterns. Looking for oddities often means looking at the residuals from a statistical model. You need to get familiar with interpretation of residuals.
- For balanced experimental designs, you will probably need to be familiar with the methods of cross-environment analysis of variance. For the purposes here, a balanced design is one in which (a) the same treatments occur in the same sort of layout in each environment, and (b) the layout in each environment a completely randomised or randomised block design.

```
Error: location:rep
          Df Sum Sq Mean Sq F value Pr(>F)
location   7  885.3  126.47   148.2 <2e-16 ***
Residuals 23   19.6    0.85
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: Within
                   Df Sum Sq Mean Sq F value   Pr(>F)
genotype           24  12.85  0.5356   1.654   0.0268 *
location:genotype 168  88.86  0.5289   1.633 1.92e-05 ***
Residuals         552 178.79  0.3239
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Basic cross-location ANOVA for a trial involving 25 genotypes evaluated at 8 locations.

- For unbalanced and less regular designs, then you will need to be familiar with the use of mixed models and REML.

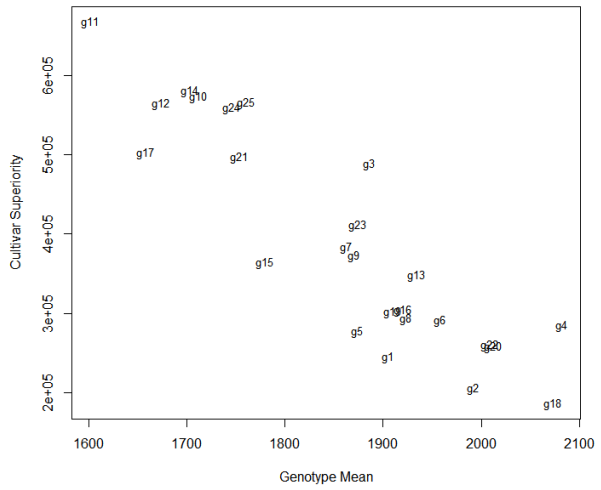
```
> anova(mod4,ddf="Kenward-Roger",type=1)
Analysis of Variance Table of type I  with  Kenward-Roger
approximation for degrees of freedom
                  Sum Sq Mean Sq NumDF  DenDF F.value    Pr(>F)
location          321.87  45.981     7  23.00 148.242 < 2.2e-16 ***
genotype           11.65   0.485    24 537.10   1.564   0.04347 *
location:genotype  82.70   0.492   168 508.21   1.587 6.624e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>

Random effects:
 Groups             Name        Variance Std.Dev.
 location:rep:block (Intercept) 0.01651  0.1285
 location:rep       (Intercept) 0.01842  0.1357
 Residual                       0.31017  0.5569
Number of obs: 775, groups:  location:rep:block, 155; location:rep, 31
```

Using a mixed model and REML to analyse the previous example. The design at each site was a lattice, with small BLOCKs within each REP.

- Both of these methods have tricks and complications to get familiar with, for example:
    - Generalisations such as those to deal with data measured on other scales, such as yes/no data or scores.
    - Tools for checking the assumptions and changing approach if needed. This includes analysis of residuals and looking for spatial pattern.
    - Methods for extracting and presenting means and their measures of precision.
- It will be necessary to iterate between exploratory and confirmatory analyses.
- Sometimes the statistical tools of confirmatory analysis are needed to extract exploratory information. For example, if there is a hypothesis that 4 factors all affect the performance of a technology on farms, statistical modelling will be needed to estimate their effects. It is usual not possible to find simple summaries and plots that will show the effect of multiple factors on a response.

6. GxE or OxC interaction

- The MET Overview document [link] has explained the concept of interaction in objectives or METs. Traditionally this is genotype by environment (GxE) interaction, but the same ideas and methods can be used to look at other situations that we described generally as options by context (OxC) interaction.
- In the simplest cases, good graphs will suggest whether there is interesting GxE interaction, and ANOVA or mixed models will provide the confirmation and measures of precision. However, we usually need to use some specifically adapted tools to pull out clear conclusions from large or complex studies.
- The GxE interaction may be so striking that it simply needs common sense to describe it. For example, if there are useful genotype differences at one site but at another everything fails, this is GxE. But it is handled simply by looking at the two sites separately, and trying to understand why there was complete failure at one of them.
- After exploratory analysis, try to start with an appropriate ANOVA or mixed model that includes a GxE term. This will show if there is any interaction to understand and explain.
- Be aware that a conclusion of 'no significant GxE interaction' in an ANOVA or mixed model does not necessarily mean there is none! In large problems, some interesting pattern can be lost in a lot of noise. Be alert for this if the degrees of freedom for GxE is large.
- The concept of stability v local (or niche-specific) adaptation is important. An option is stable if its performance does not change much with environment. An option is locally adapted, or adapted to a specific niche, if it performs well in some well-defined contexts even though it may not perform well generally. The ideas are explained well by (Paolo Annicchiarico 2002).
- For many years, agricultural research sought stability and statisticians developed measures to help identify stable options. Currently there is increasing recognition of the benefits of matching options to niches, so stability might be a less important concept.
- When measuring stability, distinguish between static and dynamic stability.
    - An option or genotype is statically stable if its performance does not change with context or environment.
    - An option or genotype is dynamically stable if its performance does change with context or environment, but in line with the mean for that environment. Thus a dynamically stable genotype will have low yield in poor sites and high yield in good sites, but its performance relative to other genotypes will not change with environment.



Various measures of stability have been developed. For example, Lin and Binns based a measure on the differences between performance of each genotype and the performance of the best in that environment(Lin & Binns 1988). A low value of the superiority index indicates a genotype that does well across many environments. g4 and g18 have similar overall mean values, but the superiority score is much lower for g18 than g4. This suggests g18 is likely to be more consistent across environments than g4, which is likely to be performing very well in certain environments but less well in others.

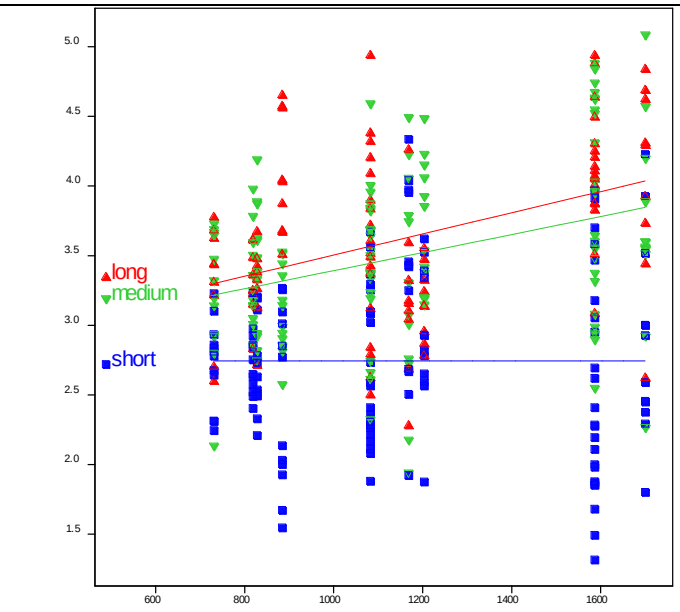
- There are many approaches to statistical analysis of GxE or OxC interaction, many identified by the original authors names and some often generating intense argument over their relative merits. Most can be implemented starting with the O x C table or mean performance data.
- The methods all fall into one of three groups (Paolo Annicchiarico 2002):

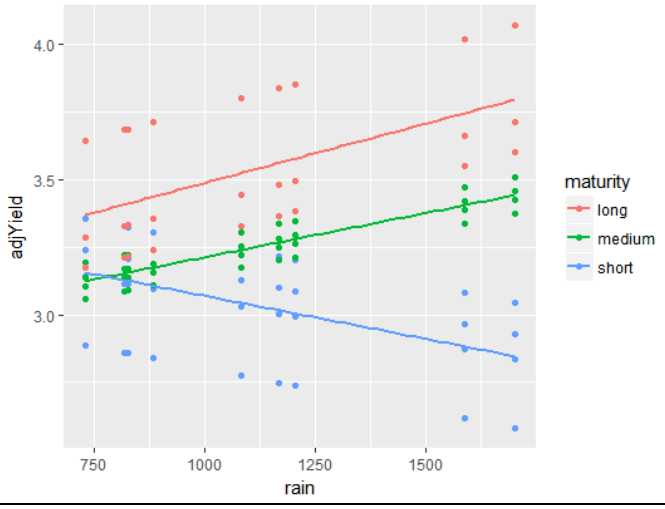
1. *Modelling response to measured environment variables.*

    - If we measure characteristics of each environment (e.g. seasonal rainfall) then we could relate the performance of each genotype to that measure by regression modelling. Differences in the relationship for different genotype are a component of GxE.
    - The method has the strong advantage that it suggests mechanisms and interaction, and would allow some prediction about what would happen in a new environment.
    - The method has the disadvantages of (a) requiring some insight into what might be responsible for GxE (having a hypothesis) so that you know what to measure at each site; (b) being able to summarize the many dimensions of difference between environment in a few simple measures; and (c) assuming that the relationships can be modelled in a simple way by regression.
    - It may also be possible to measure some characteristic of each genotype (eg duration length) and get further insight into interaction, as in this example:

```
Analysis of Variance Table

Response: yield
                Df Sum Sq Mean Sq  F value    Pr(>F)
site            10 734.48  73.448 506.9100 < 2.2e-16 ***
variety         10  24.25   2.425  16.7332 < 2.2e-16 ***
site:block      22   2.87   0.130   0.8998    0.5953
rain:maturity    2   4.35   2.176  15.0208 7.676e-07 ***
site:variety    98  89.58   0.914   6.3088 < 2.2e-16 ***
Residuals      220  31.88   0.145
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

An example with 11 varieties over 11 sites recorded the rainfall at each site. Varieties fell into 3 groups of short, medium and long duration. A usefully large part of the GxE interaction is due to differences in the way varieties of different duration respond to rainfall.

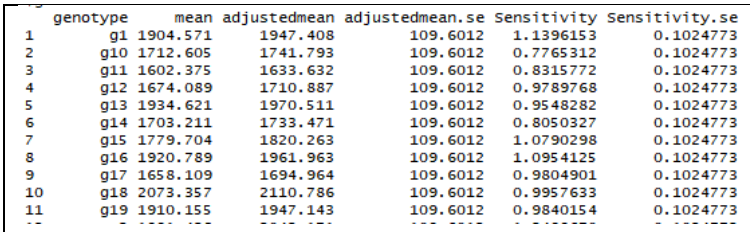
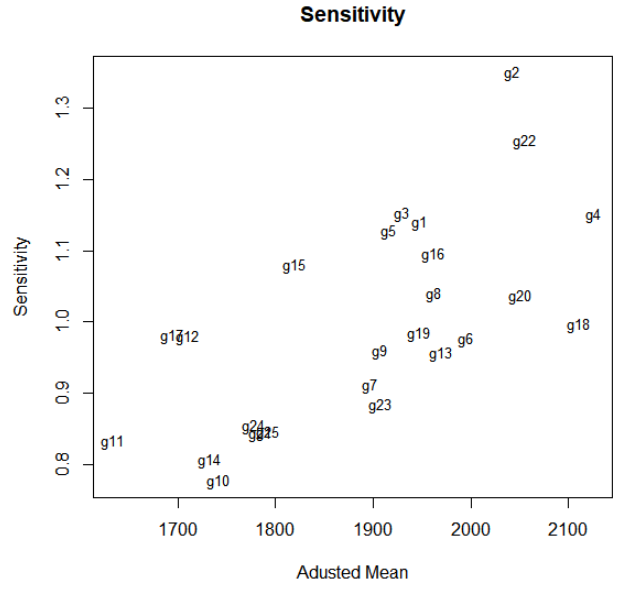




A graph of the results shows that, on average, long and medium duration varieties have a positive response (increasing yield) to rainfall, but short duration varieties do not respond to rainfall. Maybe they reach maternity before water stress becomes an issue. They appear more stable with respect to rainfall, but also produced lower yields.

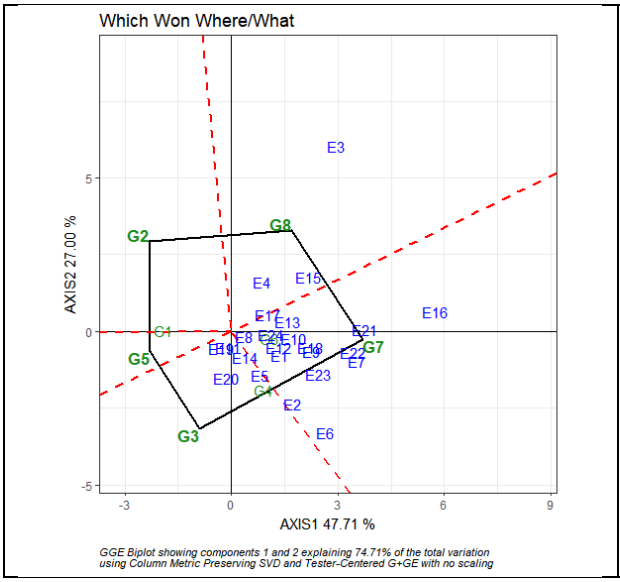
2. *Constructing an environment index and modelling response to it*

- If we did not measure something useful about each environment we can construct an 'environmental index' for each from the genotype performance data then do an analysis similar to 1 above.
- The method is motivated by noticing that a common form of GxE interaction is that some genotypes do well in 'good' environments but others do relatively better in 'poor' environments. We define the environment quality by the average performance of all the genotypes, using that average as an index.
- The method has the advantage of being data driven and not requiring addition environmental data.
- The method has the disadvantages of (a) only revealing and describing very specific types of GxE interaction; (b) giving results for a given genotype that is dependent on which others happen to have been included in the trial.
- There are many variations on the method (Hildebrand 1983; Finlay & Wilkinson 1963).
- The results are often interpreted in terms of stability.

	The yields from the groundnut example with 25 entries and 8 environments was analysed for sensitivity to environment index. Results for just 11 genotypes are shown. The Sensitivity parameter is the regression slope. Genotype g10 is relatively stable. Genotype g16 does better in good environments.
	Plotting sensitivity against mean shows the trade-off. Genotype g11 is not sensitive (or is stable) but has a low mean. Genotype g2 has a high mean but is sensitive – it will perform relatively poorly in poorer environments. Entry g18 has a high mean and average sensitivity. Remember you can always get perfect stability (sensitivity=0) by selecting a genotype that produces 0 everywhere! Nothing can be more stable than that, but it is not very useful. Reliability indices combine estimates of mean and stability.

3. *Pattern analysis, looking for structure in the OxC table*

- Statistical methods can be used to identify structure in the 2 way table of genotype by environment means. A number of methods can be used based on methods of principal components and biplots.
- Two popular sets of methods are AMMI (additive main effects, multiplicative interaction) models and plots (Gauch 1992) and GGE (genotype and genotype by environment) biplots (Yan et al. 2000). Both have their supporters and detractors. Both have their uses in different situations.
- See Easy Biplots [link] for details.
- Advantages of these methods are that they can show up patterns that might otherwise be missed.
- Disadvantages of the methods are (a) the difficulty in working out exactly what the result show and understanding how they were constructed; and (b) the fact that they generally only work on complete GxE tables. The methods fail if some genotypes did not appear in some environments.
- The previous method of constructing an environmental index is closely related to this method. You will need to get familiar with the mathematics of what is going on in these analyses. See the Easy Biplot guide [link] for a start.



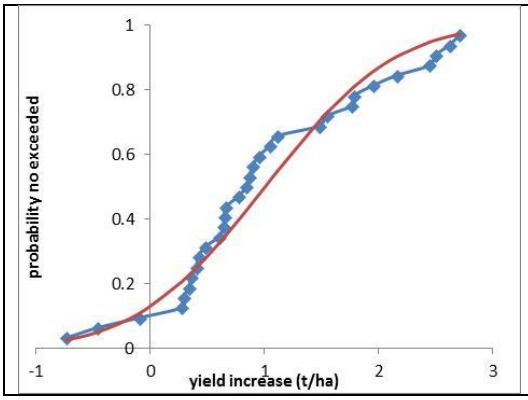
Example of a GGE biplot for data on 8 genotypes (green) in 24 environments (blue). See Easy Biplots [link] for more examples and interpretation.

- While we have presented these as three alternatives, there are connections and links between them. Each approach has many variations and enhancements.
- Using the methods described above for understanding OxC or GxE interaction can be powerful. But it is also easy to get seduced by the method and lose track of the objectives and interpretation. When you get lots of output, 'significant' results and sophisticated looking graphs, step back a little and ask:
  - Do these results make scientific sense? Can you explain what they show in terms of people and plants, microbes and money or whatever?

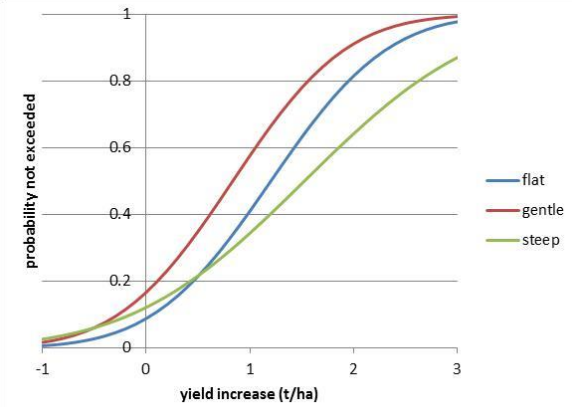
- Are the results of practical use – for example, are interactions large enough to be gainfully exploited or to influence any decision?
- Do the results take you nearer to meeting your objectives?

7. Risk analysis

- Risk is important to farmers and many projects have an aim of reducing risk. Yet few quantitative analyses attempt to measure risk. This may be for two reasons: there are many different types of risk, each would require its own study, and getting sufficient data to estimate risk can be difficult. Here we look at just two well-defined examples of risk and the way they could be estimated from METs.
- Researchers often aim to make recommendations to farmers, such as crop and management options to use, which are based on the empirical results of experiments. Results always show variation between farms and farmers. We typically quote means and use these to make recommendations. But, is the mean relevant to a new farmer taking a decision to adopt? The variation in performance between farms that have compared alternatives gives insight into the risk faced by the new farmer. Hence, we can use data from trials conducted across many farms to look at the variation and risks.
- A large-N design is essential for being able to carry out a good risk analysis, as it requires the estimation of the full distribution of results – not just the average value. This requires more data!
- The details depend on the exact objectives and design. One example is shown in the box below.

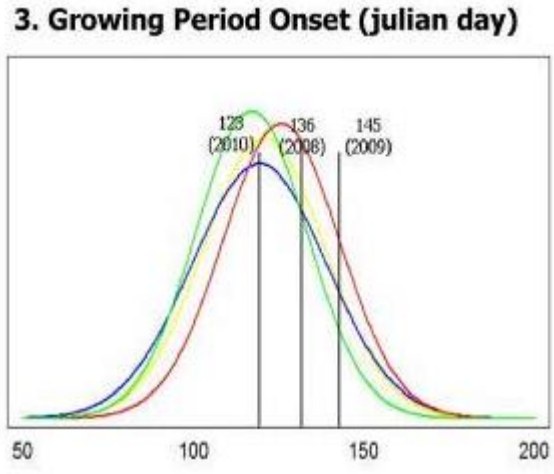


Yield increase with a soil fertility technology showed much variation between 31 farms. This is risk faced by an individual farmer basing decisions on the mean result. The empirical risk curve is in blue. A simple normal model (red) seems a good approximation to risk.



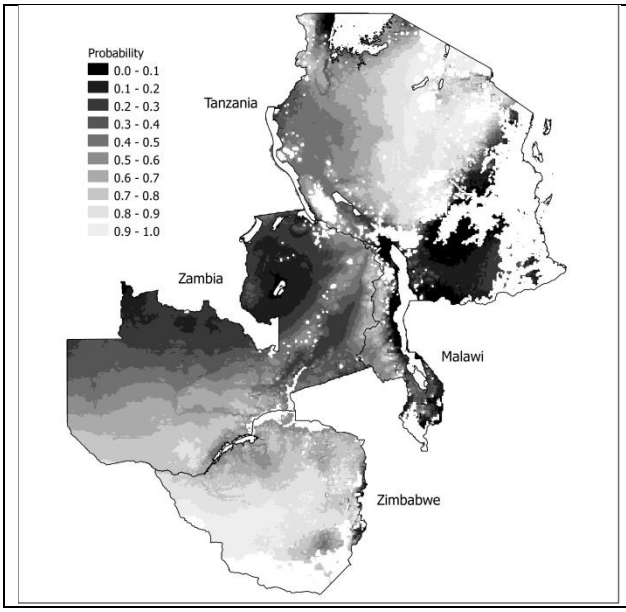
There was a suggestion that some of the risk is associated with landscape position – it should work better on flat land. Breaking down by slope class shows that indeed changing from flat to gentle shifts the risk curve. For steep land the chance of both very high or low yield increase is greater than for flat land – it is uncertain. But the sample size for steep land was very small.

- Risk considerations can be built into GxE analyses of variety trials. See (Paolo Annicchiarico 2002). Reliability indices try to account for trad-eoff between mean and risk in performance.
- Note that farmers may well be interested in finding a range of options that they can adapt to different niches or contexts, rather than simply knowing the best 'on average'. Careful analysis of farmers aims and interests may suggest different approaches to analysing risk data and presenting results
- A second source of risk is season-to-season variation in the weather, with rainfall often the most important. Estimating that directly from experimental data requires long series of experiments, and these can rarely be done. Anyway, we want answers long before they would be finished. Two options are (a) using variation in rainfall between different places as a proxy for variation over time (b) looking at variation in weather records that do exist and trying to infer something about effect on the options.
- One simple way is to conduct an experiment over several seasons and try to 'guess' the weather characteristics that have led to the result (e.g. a long dry spell after planting, good rain during grain filling), then use weather records to estimate the risk of those conditions.



In this example (Traore et al 2012) the date of the start of the growing season (x-axis) is shown for 3 years of an experiment, together with the distribution of this date for four different reference periods (coloured curves). We see that in 2010 the start date was about average. In 2008 and 2009 it was later than average but not unusually so. If the late start in 2009 caused a problem for your crop, you better do something about it, as a start even later than that is fairly likely.

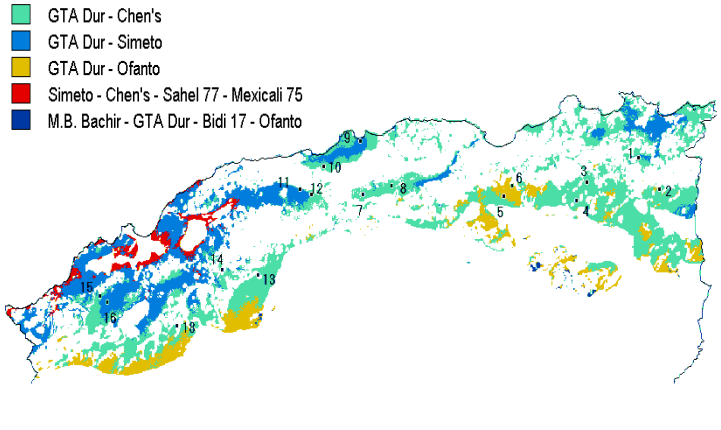
- The difficulty with this method is that performance is not related to a single weather event, but to the cumulative weather experience of the season. Crop-weather models try to simulate that and can sometimes be used to substitute for empirical field data.
- If realistic relationships between performance and environment variables can be found, then it may be possible to map risk results.



An MET was carried out in about 80 environments in southern Africa. Performance of a soil management option was estimated in each and related to rainfall, altitude and soil variables. Using information on the variation in rainfall etc. across the region, it was possible to produce this risk map. It shows the probability that a farmer adopting the new technology gets only a small return. Details in Coe et al (to appear).

8. Extending the scope

- The analysis of an MET should not end with statistical analysis. The statistics is a small part of the overall task of merging what you have learnt with the existing knowledge base and exploring the implications.
- Tools to use might include economic models, farm system models, GIS, crop models and more. The limitation is your imagination!



In this example (P. Annicchiarico et al. 2006) results of trials that showed GxE have been combined with rainfall and temperature data to produce maps of recommendation domains for different varieties (colours on the map)

9. Reporting

- Just do it! But it will take a plan and coordinated effort.
- Remember you will need different types of reports for the very different audiences.
- Anyone who participated in the MET, particularly farmers, must get feedback on the results in a format that is understandable and useful to them.

10. References and further reading

- Confidence and Significance (SSC)
  http://www.reading.ac.uk/ssc/n/resources/Docs/Inferential_Statistics.pdf

All the others are docs for this w/s

Annicchiarico, Paolo, 2002. *Genotype x Environment Interactions - Challenges and Opportunities for Plant Breeding and Cultivar Recommendations*, Rome: FAO. Available at: http://www.fao.org/DOCREP/005/Y4391E/Y4391E00.HTM#Contents.

Annicchiarico, P., Bellah, F. & Chiari, T., 2006. Repeatable genotype x location interaction and its exploitation by conventional and GIS-based cultivar recommendation for durum wheat in Algeria. *European Journal of Agronomy*, 24, pp.70-81.

Finlay, K.W.. & Wilkinson, G.N., 1963. The analysis of adaptation in a plant-breeding programme. *Australian Journal of Agricultural Research*, 14, pp.742-754.

Gauch, H.G., 1992. *Statistical Analysis of Regional Yield Trials - AMMI analysis of factorial designs*, Amsterdam: Elsevier.

Hildebrand, P.E., 1983. Modified Stability Analysis. *Agronomy Journal*, 76, pp.271-274.

Lin, C.S. & Binns, M.R., 1988. A superiority performance measure of cultivar performance for cultivar x location data. *Canadian Journal of Plant Science*, 68, pp.193-198.

Yan, W. et al., 2000. Cultivar Evaluation and Mega-Environment Investigation Based on the GGE Biplot. *Genetics*, 40(3), pp.597-605.