Q & A: THE USE OF A NEW STATISTICAL PROCEDURE BY NTEP

Mr. Kevin Morris, NTEP and Dr. Scott Ebdon, University of Massachusetts
May 1, 2009

1. What is the new statistical procedure that NTEP is now using?

The new procedure is called 'AMMI' for short (stands for Additive Main Effect and Multiplicative Interaction). This procedure was developed and tested by statisticians at the University of Massachusetts and Cornell University over the last ten years. In contrast, the statistical procedure NTEP has been using since its inception, ANOVA (Analysis of Variance), has been used for statistical analysis of agricultural experiments since the 1930's.

2. Why did NTEP feel the need to change the procedure it uses to statistically analyze its data?

Because NTEP trials are evaluated at many locations across the U.S., cultivars and experimental selections (called genotypes) vary in their performance, based primarily on the location (environment) in which they are tested. These differences in performance result from what is termed a 'genotype by environment interaction' (G x E for short). In other words, different climatic conditions, weather patterns, soil types, diseases, etc. among the locations affect the performance of the grasses NTEP tests. The traditional ANOVA procedure is not designed to adequately analyze the way the genotypes interact with the different environments. AMMI was developed to more accurately analyze the 'G x E' interactions in crop yield trials. This is the first instance of AMMI being used for turfgrass cultivar trials.

3. How is AMMI different than what NTEP has used in the past?

AMMI identifies the genotypes (or turfgrass entries in this case) and how they interact with each environment (trial location), captures true structure in those interactions (in a data form), and improves the accuracy of the statistical analysis. The result of this additional analysis is that a more accurate representation of each cultivar's performance is represented in the data. For those who use NTEP data analyzed by AMMI the greater accuracy provides more reliable plant recommendations. In contrast, ANOVA does not capture the G x E interaction structure because ANOVA cannot analyze it. As a result, true cultivar differences that exist as a result of genotypes (cultivars) interacting with the trial locations, is ignored when using ANOVA. Therefore, one that looks at data analyzed by ANOVA, particularly where several locations are grouped together, will get an inaccurate and unreliable picture of turfgrass performance.

4. If AMMI is so much better, why did NTEP wait until now to start using it?

In 1999, NTEP funded five research projects of different statistical methods in an effort to improve our data analysis. AMMI was one of those projects. Results from the computer simulation study of past NTEP trials showed that AMMI increased statistical accuracy anywhere from 50 to 500% over ANOVA (depending on the species). Due to these promising results, a field validation study was established to evaluate which statistical method, AMMI or ANOVA, would more accurately predict the top cultivars, when planted at various locations. The project, coordinated by NTEP, Dr. Scott Ebdon (University of Massachusetts) and Dr. Hugh Gauch (Cornell University), was conducted for four years at six locations. The field validation study confirmed what was seen in the computer simulations. At this point, the NTEP Policy Committee voted recently to implement this new procedure, almost ten years after the initial research project began.

5. For what trials is NTEP using AMMI?

NTEP has decided to start using AMMI only on new trials, where no data has yet been statistically analyzed. Trials started in 2007 (bermudagrass, seashore paspalum, st. augustinegrass and zoysiagrass) are the first candidates for the use of AMMI. In 2010, new trials of bentgrass and fineleaf fescue will also use the AMMI procedure, as will new trials from there on out. One caveat, not every trial is a candidate for AMMI. A test procedure is conducted on a data set to determine if there is a significant statistical gain in accuracy by using AMMI. If not, the traditional ANOVA procedure will be used.

6. What differences in data, if any, we will notice with this new procedure?

First, AMMI is only being used on *turfgrass quality data*. Second, since AMMI basically 'adjusts' means (by reanalyzing them), data that is found in tables where several trial locations are listed, may look different than data on a page with only a single trial location.

7. Sometimes, I may look at turfgrass quality data on a state page (for instance, by going to http://www.ntep.org/states/states.htm and clicking on a state), and the data in the 2008 data table may look different than it is in the national data table. Is this correct?

Yes, as noted above, when AMMI reanalyzes the G x E interactions, then the means for each location may be 'adjusted'. Therefore, the turfgrass quality data in a table on a state web page is analyzed using the traditional ANOVA procedure because that is the appropriate procedure where data from only one trial location is present. However, when data from multiple trial locations are analyzed together (the statistical term is 'pooled'), then AMMI is the more appropriate procedure.

8. If the turfgrass quality data on a state web page is different from the data in an AMMI grouping, which data should I consider the most accurate, and therefore best to use?

Since AMMI means are 'adjusted' and therefore, more accurate, these are the most appropriate to use. Therefore, we recommend that a user consult the turfgrass quality data reported in an AMMI grouping of locations instead of the turfgrass quality data in a state web page. Also, if your state is not represented in anyone of the AMMI groupings of state locations, then choose an AMMI group with a location and management closest to your planting conditions. The other data on a state web page (i.e. genetic color, leaf texture, disease, etc.) is perfectly appropriate to use since the AMMI procedure is not utilized for this data. Please keep in mind, that for now, these recommendations are only valid for trials where AMMI is utilized (2007 trials – bermudagrass, etc. – see Question 5 above).

9. What happened to the turfgrass quality data analyzed by regions (i.e., Northeast, Southeast, Transition Zone, etc.)?

NTEP analysis of turfgrass quality by region was used to group locations that were geographically close to each other. However, that method of grouping locations had no statistical basis. In other words, the locations that are grouped together now (based on the AMMI program) are in those groups because AMMI found them to be statistical similar in their G x E interaction structure. Unlike previous grouping methods, top performers are similar from location to location for the method of grouping used by AMMI. Therefore we are not grouping locations based on a region, but on a sound statistical foundation.

10. Why are the turfgrass quality LSD values for each state location in a table the same using the new procedure (AMMI), when in the past each state had a different LSD value?

Because of the way AMMI reanalyzes and then 'adjusts' means, the procedure pools all the data for locations resulting in a unified (single) LSD value. The ability to achieve greater accuracy and more reliable planting recommendations using AMMI is based on pooling of all the data. This is a normal and valid outcome of utilizing AMMI.

11. In the past, NTEP data included an 'Appendix' table. What happened to the 'Appendix' table?

The Appendix table included an average of the data from all trial locations, plus a measure (percentage) of trial locations where each entry finished in the top 25% of all entries. However, since AMMI reanalyzes and 'adjusts' means so that trial locations that are statistically correlated are grouped together, the turfgrass quality means of the

different ‘environments’ (see question 8) are what should be used when evaluating turfgrass performance. Therefore, because of the AMMI analysis, the information in the Appendix table is no longer statistically valid and not included.

12. Are there any things that we need to know to interpret the data, versus what we have used in the past?

Even though the AMMI procedure itself does not compute LSD (Least Significant Difference) values and C. V. (Coefficient of Variation) values, NTEP is computing those to add to the AMMI analysis. Therefore, use and interpretation of the data is the same for the end-user, albeit more statistically accurate than ever before.