

Answers Triticale assignment

1. **Summary statistics.** See Excel file Triticale Answers.xlsx, sheet entitled “Question 1”.
 - a. The best environment is Coruna. The worst is Toledo.
 - b. The variance is very heterogeneous. In Coruna it is 2.406, while in Toledo it is only 0.022. The mean clearly correlates with the variance ($r = 0.66$, and even 0.75 with the SD).
 - c. Correlations vary greatly between environments, from -0.46 between Coruna and Sevilla_E, to 0.81 between Coruna and Orense. Hence, these values clearly indicate GxE-interaction.
 - d. The above results suggest quite strong GxE-interaction.

2. Finlay-Wilkinson regression. For the data file with the E-added, see file Triticale GxE data_FW.xls. Results of fitting the main model are in Triticale_main_G_E.asr and Triticale_main_G_E.sln. You will see that the ranking of the E-terms in the sln file is the same as in the raw data (question 1), because the data are balanced. From the main-effects model, the full GxE-variance is 0.453 (residual variance of main-effects model).
 - a. *i*) The P-value for the GxE-interaction term (thus for $\beta_i E_j$) is 0.071; so significant at $P < 0.10$ but not at $P < 0.05$ (see file Triticale_FW.asr). *ii*) The mean beta is not zero, because ASReml arbitrarily puts one of the sensitivities to zero; this is confounded with the mean E. When you omit E from the model, the mean β becomes 1, as expected. Note that P becomes much smaller when omitting E, because now the term also covers the main environmental effect. Hence, you cannot test for GxE when you include the main environmental effect in $\beta_i E_j$. *iii*) Residual variance of the FW-model equals 0.423. Hence, this is only a little smaller than the full variance (the 0.453), so FW-regression accounts for only ~7% of the GxE-interaction variance. *iv*) If FW-regression explains only a little of the GxE, most of the GxE is not due to the main environmental effect, E, and we may want to look for a better model. You also see that the difference among the β 's is rather small.
 - b. $\beta_{C1} = 1.17$, indicating that it is above average sensitive (to the (main) environment (effect)). The $G_{C1} = 4.735$ (in a model without mu), which is above average. So general adaptation of C1 is above average.
 - c. C4 and S1 show similar general adaptation (similar G; 4.282 vs. 4.262), but C4 has somewhat greater sensitivity (1.144 vs. 1.053). Hence, C4 is probably better in a good environment.
 - d. See Triticale Answers.xlsx, sheet Question2d. S1 has much greater stability than C4. The standard deviation of the difference between predicted and observed y-values equals 0.33 for S1 but 0.76 for C4. Hence, the FW-regression predicts S1 quite accurate, but C4 not so accurate. (From a statistical perspective, this probably means we should allow for residual variance that is heterogeneous by cross).
 - e. General adaptability of the Complete genotypes is much better than of the Substituted genotypes, as indicated by the average of the genotypic values which is clearly higher. However, the Substituted genotypes are *relatively* better in poor environments, as indicated by their lower sensitivity (beta).

3. **Factorial regression on specific environmental parameters.**
 - a. pH does not explain the GxE ($P = 0.266$). Also altitude does not explain GxE ($P = 0.836$). The Genotype x Rainfall interaction, however, has a significant effect ($P = 0.033$). Residual variance is 0.413, so rainfall explains around 9% of the residual variance. Hence,

in terms of fit, this model is better than FW-regression. However, it also has the advantage that one can predict another environment for which rain-fall is known. The FW-regression cannot predict outside the current data.

4. AMMI model and biplots

- a. *i)* The AMMI model explains a substantial amount of the GxE. The residual variance goes down from 0.453 to 0.082, so $(0.453-0.082)/0.453 = 82\%$ of the GxE-variance is explained by the first two PC of the AMMI model. *ii)* The AMMI explains much more variance than the FW (See answer for FW above).
- b. *i)* The C genotypes clearly group together, as do the S-genotypes. The C-genotypes are poorly adapted to PC1, while the S-genotypes are well adapted to PC1. *ii)* C4 has strong positive interaction with Orense (“well-adapted to Orense”), because the projection of C4 on the vector of Orense is in the same direction as this vector and quite far from zero. S4 has strong negative interaction with Orense. C4 has very little interaction with Salamanca (the projection is close to zero), S4 has moderate positive interaction with Salamanca. *iii)* S2 and S8.
- c. No, it is NOT! The biplot follows from the singular value decomposition of the GE_{ij} interaction matrix. This does NOT contain the main genotypic effects. So S2 and S8 are RELATIVELY good in Sevilla and Lleida, compared to the performance of those two genotypes in other environments. However, their absolute performance cannot be judged from the biplot.