APPENDIX A: Fixed vs Mixed effect model

To illustrate the difference between the fixed effect model (as proposed by Griffing 1956) and the mixed effect model we suggest in the article, consider an experiment that consists in all pairwise combinations between M varieties, each combination being observed R times (i.e. a repeated half-diallel design), and R replications of one additional mixture corresponding to the mixture of varieties $M+1$ and $M+2$ (Fig. A.1). The statistical analysis of this experimental design can then be performed using a fixed or a random effect model.

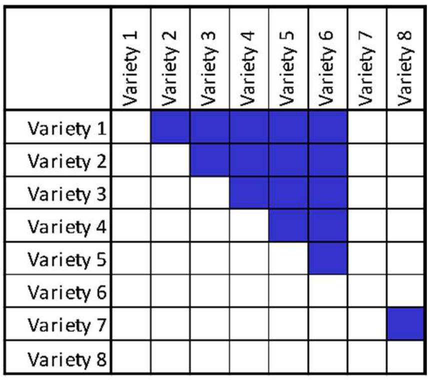

Fig. A.1: Representation of the diallel-like design for the example (with $M=6$ varieties), each colored square representing an observed mixture.

Fixed effect model

The statistical model is

$$Y_{ijr} = \mu + \alpha_i + \alpha_j + \gamma_{ij} + E_{ijr}$$

where the α s and γ s coefficients are fixed as in the Griffing model (1956), and where the errors satisfy

$$Eijr \sim N(0, \sigma_E^2), i.i.d..$$

We also note $\mu_{ij} = \mu + \alpha_i + \alpha_j + \gamma_{ij}$ the average response obtained for the mixture of varieties i and j. In this context, if one defines the GMA for variety i as

$$GMA_i = \sum_{j=1}^{M+2} \mu_{ij} - \sum_{k,j=1}^{M+2} \mu_{kj}$$

then none of the GMA is estimable. Alternatively, if one defines

$$GMA_i = \sum_{j \in C_i} \mu_{ij} - \sum_{k,j \in C} \mu_{kj}$$

where C is the set of all observed combinations and C_i is the set of all combinations involving variety i, then

- for a given set of varieties the definitions of the GMAs depend on the actual experimental design,
- the definition of "GMA" becomes specific to each variety i,
- GMA_{M+1} and GMA_{M+2} are defined to be 0.

This can be extended to the definition of the SMA: if one defines the SMA as in the complete design, then no SMA is estimable as soon as the mixing design is imbalanced, and if one adopts the design-dependent definition of the SMA, then the remarks above hold and one can observe that

- SMA_{ij} can be estimated for $i, j \leq M$,
- $SMA_{i,M+1}$ cannot be estimated whatever $i \leq M+1$,
- $SMA_{M+1,M+2}$ is defined to be 0.

Random effect model

The statistical model is

$$Y_{ijr} + \mu + G_i + G_j + S_{ij} + E_{ijr}$$

where the G s and S s are random effects. In this last model, we assume that all the G_i and S_{ij} effects are independent and that

$$G_i \sim N(0, \sigma_G^2), i.i.d$$

$$S_{ij} \sim N(0, \sigma_S^2), i.i.d$$

$$E_{ijr} \sim N(0, \sigma_E^2), i.i.d \ .$$

One can observe that there is no identifiability problem: the intercept and all variances are identifiable. Also note that while the variance estimates will depend on the actual (i.e. realized) mixing design, the definition of the variances does not depend on the specific mixtures that are observed.

Additionally, the problem of estimating some contrasts (corresponding to the GMA and SMA effects, Wu and Matheson 2000) is now replaced by the problem of predicting the random effects G and S. Considering the experimental design described above, note that the predictions can be obtained from the BLUP formula whatever the actual design. Applied to the design described above, one can see that

- all G effects will be predicted - possibly quite poorly - with non-zero values,
- all S effects corresponding to observed variety combinations will be predicted with non-zero values,
- all S effects corresponding to non-observed variety combinations will be predicted at zero, i.e. nothing will be learned from the experiment for these effects.

REFERENCES

Griffing, B., 1956 Concept of General and Specific Combining Ability in relation to diallel crossing systems. Aust. J. Biol. Sci. 9: 463–93. https://doi.org/10.1071/BI9560463

Wu, H. X., A. C. Matheson, 2000 Analysis of half-diallel mating design with missing crosses: theory and SAS program for testing and estimating GCA and SCA fixed effects. Silvae genetica, 49 (3), 130-136.