

## **Predicting yield in new environments from a variety-testing trial network using random regression models**

Joint work with Daniela Bustos-Korts, Willem Kruijer and Fred van Eeuwijk (&... ??)

Predicting yield in new environments from VCU (Value for Cultivation and Use) crop variety trials remains a challenge in plant breeding. In VCU trials, new varieties are typically tested in field trials for a few consecutive years in multiple locations before being released and replaced by others. Genotype-by-trial (or genotype-by-year-by-location) analysis (GxE) of these trials focusses on the ranking of varieties in terms of mean yield across environments, using variance component linear mixed models (LMM). A major question in the VCU context is whether the current trial system can be improved (e.g. by increasing the number of test years per variety) to get better predictions for each genotype, and whether environmental covariates can be used to predict phenotypes in novel environments.

Using a fully balanced (simulated) dataset of wheat varieties across multiple years and locations, we tested the predictive performance of random regression models (RRM) within and outside observed environments, and compared it to conventional LMMs. We specifically looked at the number of training and test years needed for accurate prediction, as well as whether leveraging phenotypic information from a secondary location (or region) in the RRM would further increase accuracies. Our results show that RRM performed admirably well compared to traditional LMMs and that prediction accuracies in unsampled environments were moderate to high, depending on the location, the number of years used for training and prediction, and the incorporation of phenotypes from a secondary location. RRM are thus a useful tool for prediction in a VCU context, but the design of the VCU network will be an important factor that determines the performance of these models.