

AgMIP Calibration Activity, Phase 4

Description of French data set

Introduction

Phase 4 concerns the same data set from France that we used in phase 3, except that now there are additional observations of yield and yield components, in-season biomass and protein contents in each environment. The objective is to predict phenology (as in phase 3) and also the other variables.

French data set

The data come from variety trials in France carried out by Arvalis – Institut du végétal, Paris. Multi-year trials were run at multiple locations across France for two winter wheat varieties, named A and B here. The trials have three repetitions and follow standard agricultural practices, with N fertilization sufficient to avoid N stress. From the 35 trials conducted for each variety, 22 were used in this study (Table 1). The data were separated into a calibration subset with data from 14 environments (six different sites, five different years, but not every year was represented at each site) and an evaluation subset with data from eight other environments (five different sites, two different years). The calibration and evaluation subsets had neither site nor year in common, so the evaluation is a rigorous test of how well the calibrated model simulates for out-of-sample environments. The data set includes observations of nine different variables. These are days to phenological stages BBCH30 (beginning of stem elongation), BBCH55 (middle of heading) and BBCH90 (maturity, estimated) for all environments, yield and yield components (ears/m², grain number) for all environments with one exception, in-season biomass in some environments and nitrogen content in final biomass and protein content in grain in some environments (Table 2).

The input data include information about the sites (latitude, soil texture, field capacity, wilting point), management (sowing dates, sowing density, irrigation and fertilization dates and amounts), and daily weather data (precipitation, minimum and maximum air temperature, global radiation, and potential evapotranspiration). Initial soil water and N content were not measured in these experiments, but best estimates were provided by the experimental scientists. The phenology measurements of this data set were used in Wallach et al. (2021), where some further details about the environments can be found.

Table 1 Environments (site-year combinations) that provided the data for the French data set. C = calibration environments. E = evaluation environments. (based on Wallach et al. 2021).

Site	Harvest year						
	2010	2011	2012	2013	2014	2015	2016
Site_1			E	E			
Site_2	C	C			C	C	
Site_3			E	E			
Site_4		C					
Site_5			E				
Site_6				E			
Site_7					C		
Site_8						C	
Site_9			E	E			
Site_10	C	C			C	C	
Site_11					C	C	C

Table 2 The variables in the French data set for each wheat variety, and the number of measurements in the calibration and in the evaluation subsets. The development stages are BBCH30 (stem elongation), BBCH55 (heading) and BBCH90 (maturity).

variable	Variable group	number of measurements in calibration data	number of measurements in evaluation data
days from sowing to BBCH30	phenology	14	8
days from sowing to BBCH55	phenology	14	8
days from sowing to BBCH90	phenology	14	8
aboveground biomass at various dates	plant_biomass	44	35
ears/m ²	ears	3	0
grains/m ²	grain_number	13	8
fraction protein in grain	seed_protein	13	8
fraction N in final biomass	plant_N-content	9	8
grain yield	yield	13	8

References

Wallach et al., 2021. How well do crop modeling groups predict wheat phenology, given calibration data from the target population? Eur. J. Agron. 124, 126195. <https://doi.org/10.1016/j.eja.2020.126195>.