Instructions for AgMIP Calibration activity

Phase 2, step 1 (“usual approach” to phenology calibration with the ARVALIS data set)

11 April 2018

# Introduction to calibration: Phase 2 step 1

The aim of this exercise is to compare calibration approaches and simulation results for the prediction of wheat phenology, where all models are provided the same data for calibration (the “training data”) and each modeling group uses their “usual” method of calibration.

The results will be compared with respect to goodness-of-fit to the training data, with respect to goodness of prediction for a separate set of evaluation data, and with respect to the difference between the two. Analysis will look at the variability between modeling groups, and in particular will see if that is related to the calibration approach (for example, to the number of estimated parameters).

This is a model + calibration-approach comparison. It is quite different than the model comparison exercises done previously in other AgMIP groups, in that emphasis is on the calibration approach more than on the models, a substantial and realistic amount of training data is provided, and evaluation is done using data which have neither site nor year in common with the training data, giving true independence between model errors for the training data and the evaluation data.

The full data set (training data and evaluation data) is from varietal tests in France. The prediction exercise is of practical interest; the question is to what extent models can predict phenology in new locations and years, given a reasonable amount of training data.

In this phase, modeling groups use their “usual” method of calibration. In subsequent phases, all groups will be asked to apply the same calibration approach. We will look at both a frequentist and a Bayesian approach. The approaches used in this first step will help define the approaches to test subsequently.

# General dataset information

The data were provided by ARVALIS - Institut du vegetal, a French technical institute. These are phenology data for two varieties of winter wheat in France, for multiple sites and years.

# What you are requested to do

You are requested to estimate the parameters of your phenology model using the training data provided. Use your usual approach to calibration of phenology. You are requested to furnish 2 files with results.

1. Using the estimated parameters, calculate the dates to stage BBCH30 (beginning of stem elongation) and to BBCH55 (50% heading) for each variety and for all the environments (the environments of the training data and of the evaluation data).
2. Fill out the short questionnaire as to how you did the calibration

# The files attached here

There are five files (not counting this file). Three are with information provided to you. The file “cal2\_phenology mgt soil data.txt” has one row per environment (an environment is defined by a combination of site and planting date) for each variety. There are 70 rows in all, 35 for each variety. Each row has soil data and management data for that environment. Some rows have phenology data for calibration. In other rows, to be used for evaluation that information is not given. The second file “cal2\_explanation of data.docx” has a detailed explanation of the meaning of the variables in the first file, plus some additional information such as average sowing depth and initial conditions. The third file “cal2\_weather.xlsx” has all the weather information for all environments.

The remaining two files are templates to be filled out with your results. The first “cal2\_results numerical modelName\_contactPerson” will contain your simulated phenology dates for the 70 environment-variety combinations. The second “cal2\_results explanation modelName\_contactPerson” will contain the description of what you did.

# The phenology data

The data provided to you are observed dates of two phenological stages, namely BBCH 30 (beginning of stem elongation) and BBCH55 (50% heading). These data are given for 14 environments for each variety (the training data). Those are the data that you will use for calibration. There are input data for 21 additional environments, but no phenology data are given for those. Your results for those 21 environments will be used for evaluation. See the diagram below for an overall picture of the separation between training data and evaluation data.

# Template for results

Please follow the attached template (file “cal2\_results numerical modelName\_contactPerson.txt”) in furnishing your results. Please furnish simulated results for all environments, both varieties. When you send your results, in the name of the results file replace modelName by the name of your model, and contactPerson by the name of the contactPerson. There is a column called “simulated\_date\_emergence\_dd/mm/yyyy ”. If you didn’t simulate time to emergence, please put “NA”s in that column. (We don’t have observations of emergence date, but it will be of interest to compare simulated values between models).

# Explanation of calibration procedure

Please answer the questions in the file “cal2\_results explanation modelName\_contactPerson.docx”. When you send your results, in the name of the explanation file replace modelName by the name of your model, and contactPerson by the name of the contactPerson.

# Training and evaluation data

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| site\harvest year | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 |
| FORESTE |  |  | x | x | xx | x |  |
| MERY | **x** | **x** |  | x | **x** | **x** |  |
| ROUVRES |  |  | x | x | x | x |  |
| Cesseville |  | **x** |  |  |  |  |  |
| Iville |  |  | x |  |  |  |  |
| Villettes |  |  |  | x |  |  |  |
| EPREVILLE |  |  |  |  | **x** |  |  |
| CRESTOT |  |  |  |  |  | **x** |  |
| OUZOUER |  | x | x | x | x | x |  |
| BIGNAN | **x** | **x** | x | x | **x** | **x** |  |
| BOIGNEVILLE |  |  | x | x | **x** | **x** | **x** |

Note that FORESTE has 2 sowing dates in 2014

**x** Training data. 14 values. 6 sites, 5 years

x Evaluation data. 8 values. 5 sites, 2 years . These evaluation data have neither site nor harvest year in common with training data.

x Secondary evaluation data. 13 values. These data have site or harvest year in common with the training data, and therefore may underestimate prediction error.