AgMIP calibration: where are we and what are the results with the STICS model?

Samuel Buis, Sabine Seidel, Taru Palosuo, Henrike Mielenz, Peter Thorburn, Marie Launay, Benjamin Dumont, Eric Justes, Daniel Wallach

XIIIth Seminar of developers and users of the soil-plant model STICS
Aerocampus Aquitaine Bordeaux Nov. 13-16, 2023

The AgMIP Calibration project

- Co-Leaders
  - Daniel Wallach (retired from INRAE – University of Bonn)
  - Taru Palosuo (Natural Resources Institute Finland - LUKE)
  - Sabine Seidel (University of Bonn)
  - Peter Thorburn (CSIRO)
  - Henrike Mielenz (Julius Kühn-Institut)
  - Samuel Buis (INRAE)

- Objectives

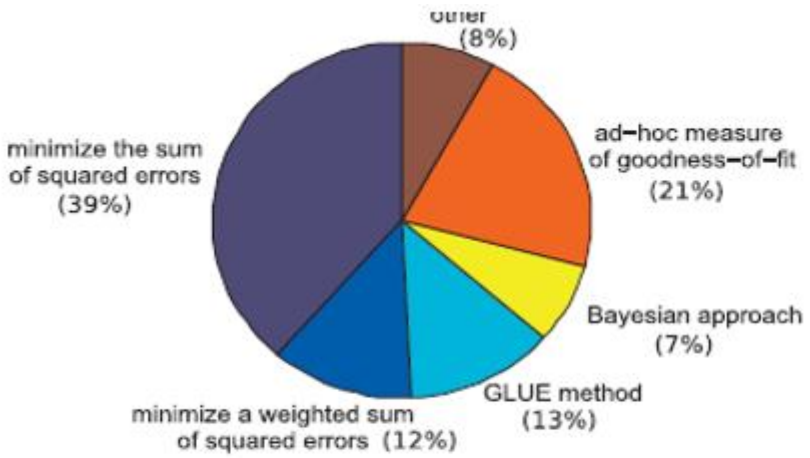
1) Build a knowledge base concerning **calibration** practices for crop models.
2) Develop and test guidelines / methods for improved calibration practices.
3) Develop tools for application of improved calibration practices

The AgMIP Calibration project

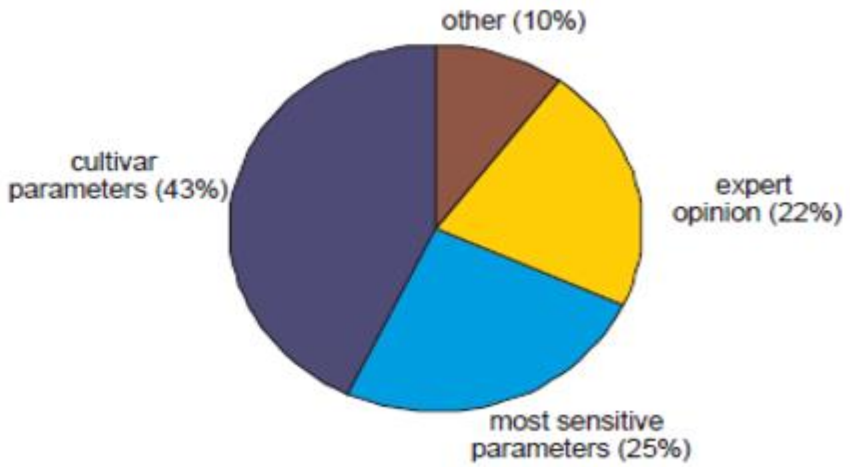
2016–2018	2018–2021	2021–2023	2023–2024	2024–
Survey	Calibration of phenology		Calibration of phenology, growth and yield	
Phase I	Phase II	Phase III	Phase IV	
	Method free	Common protocol	Common protocol	
			Synthetic study	Application to real data
Seidel et al, 2018 EJA	Wallach et al, 2021a, 2021b, 2021c	Wallach et al, 2023a ASD, 2023b (submitted to AFM)	Wallach et al, (submitted to Nature comm.)	

Phase I (2017-2018): the survey

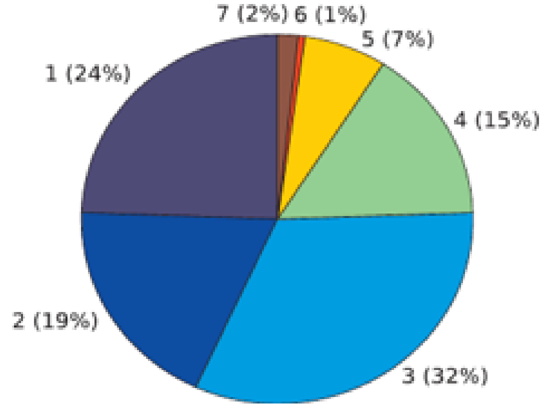
Estimation method used



Parameter selection method used



N° of steps in estimation process



Software : 30% used existing software, 26% wrote their own program, **44% modified parameters by hand**.

⇒ **Diversity of approaches and choices for model calibration**

⇒ **It would be very useful to provide guidelines, with suggestions for good practices**

⇒ **It would be very useful to provide software for crop model calibration**

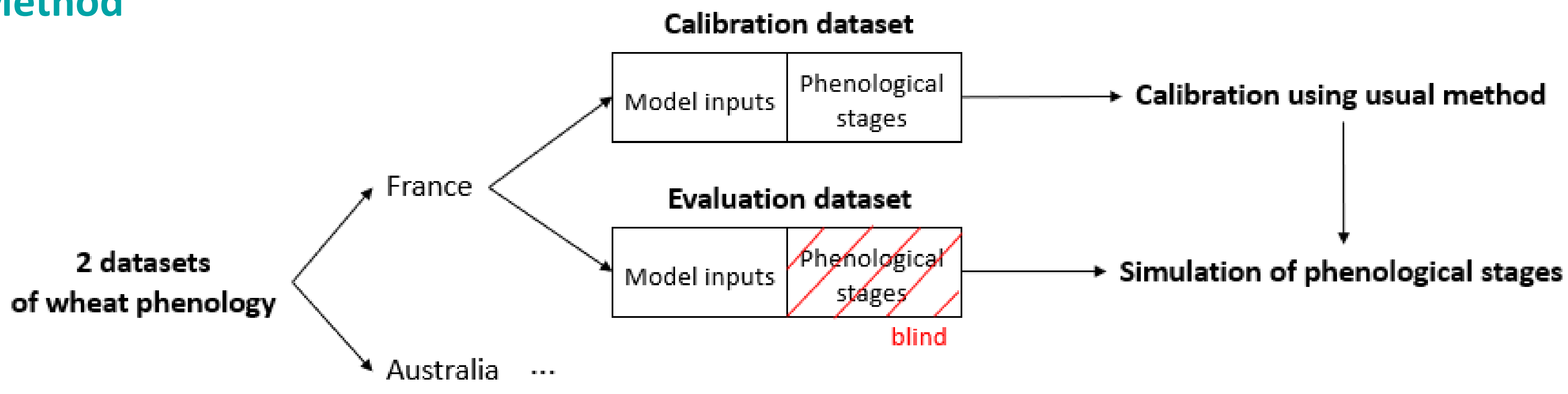
Seidel et al., 2018, EJA

Phase II (2018-2021): Calibrate your model in your "usual" way using phenology data

- **Objectives**

  How well crop modeling groups can predict wheat phenology for current conditions and management?

- **Method**



- **Participants**

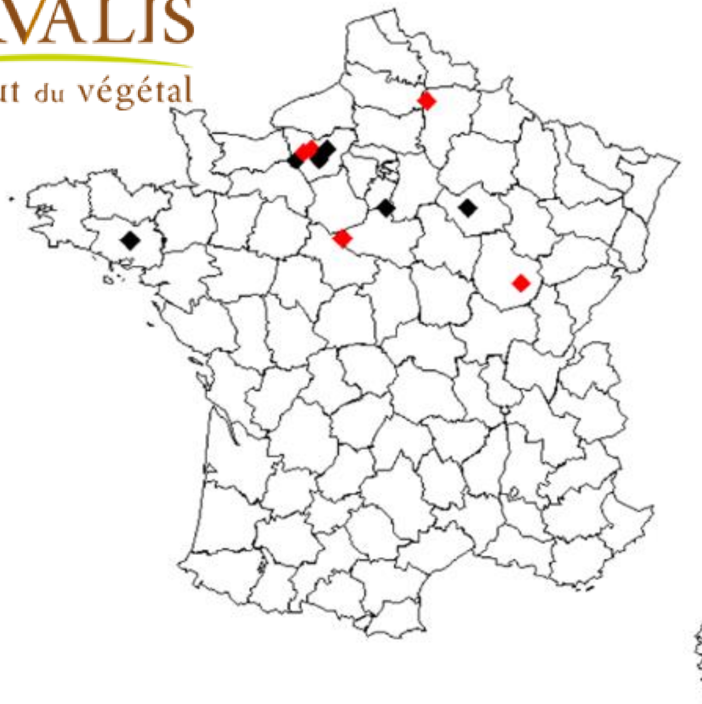
  27-28 modelling group
  20-23 models

AgroC, APSIM (4/3 groups), AquaCrop, CERES-Wheat (4/3 groups), *CoupModel*, *CROPSIM-Wheat*, Cropsyst, DAISY, Nwheat, GECROS, HERMES, LINTUL, MONICA, OpenCrop, *PANORAMIX*, Salus, SPASS, SSM-Wheat, **STICS**, SUCROS, WOFOST (2groups), Wheat-Grow

=> Results published in Wallach et al., 2021, EJA ; Wallach et al., 2021, AFM ; Wallach et al., 2021, EMS

Phase II (2018-2021): Calibrate your model in your "usual" way using phenology data

Data sets

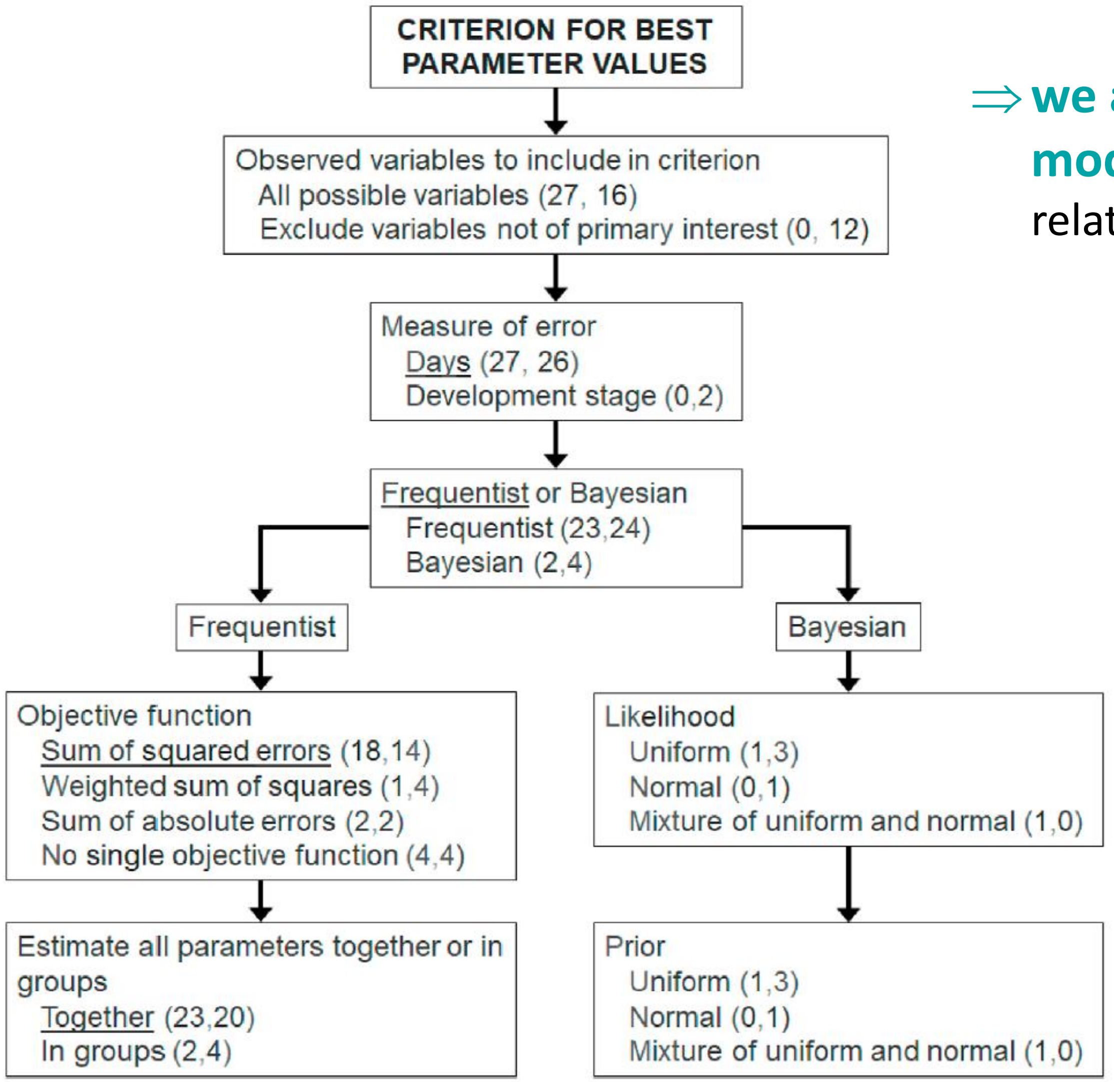


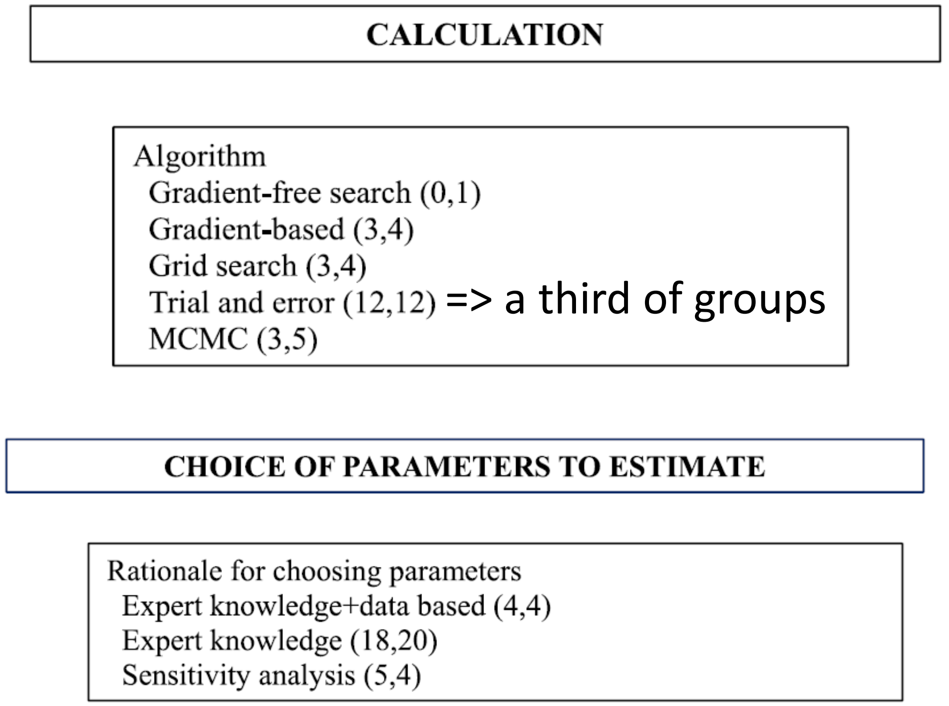
# of cultivars	1	2
# of environments in calibration data set	24 (4 sites, 2 years, 3 sowing dates)	14 (6 sites, 5 years)
# of environments in evaluation data set	18 (6 sites, 1 year, 3 sowing dates)	8 (5 sites, 2 years)
Observations	Dates of most of zadok growth stages	dates of BBCH30, BBCH55
Required stages	dates of BBCH30, BBCH65, BBCH90	dates of BBCH30, BBCH55

Phase II (2018-2021): Calibrate your model in your "usual" way using phenology data

⇒ **substantial variability in calibration approach between modeling groups** (even for same model structure)

⇒ **we are far from having a consensus on how to calibrate crop models** (even for a given model structure and dataset, and even for the relatively simple case which focuses just on phenology)





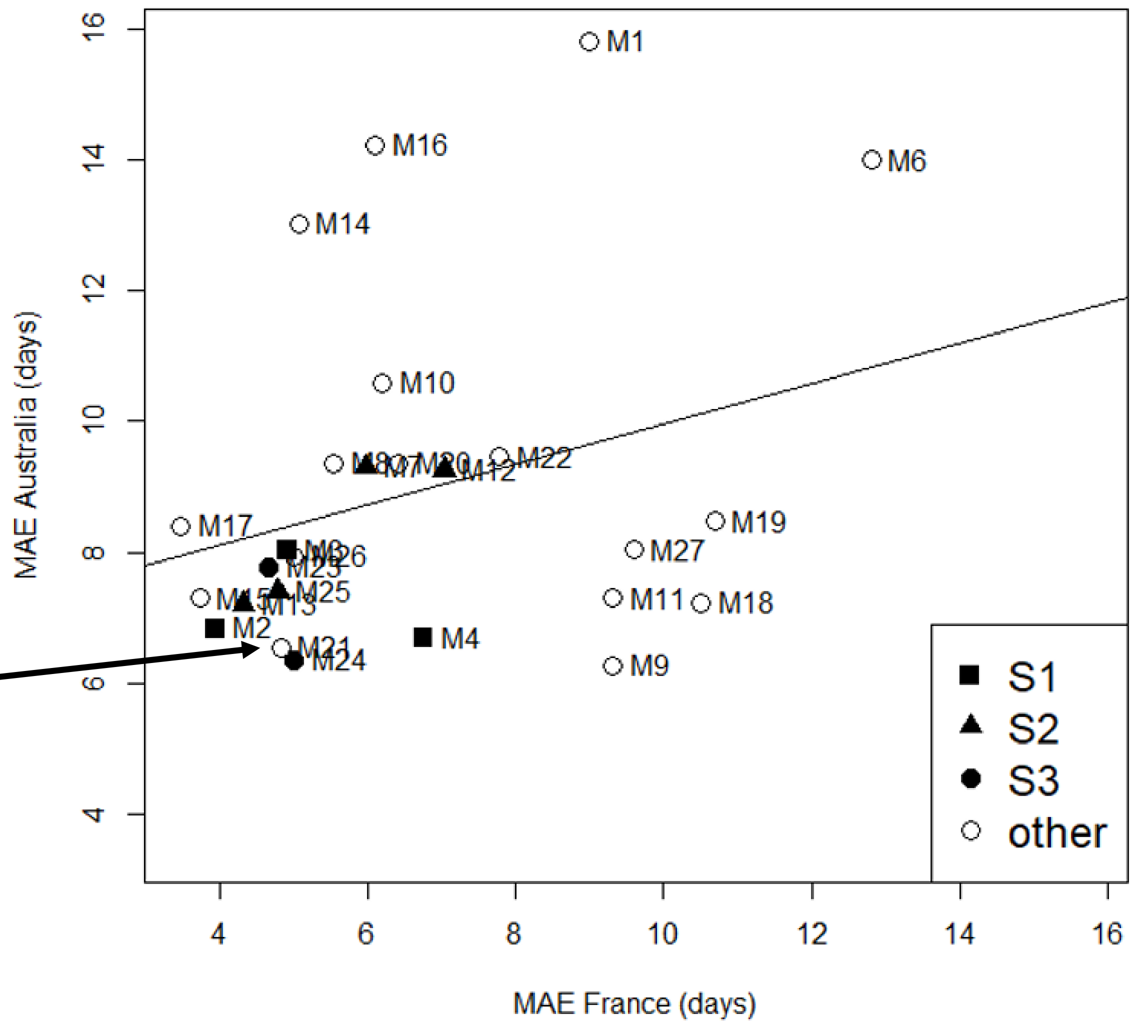
Phase II (2018-2021): Calibrate your model in your "usual" way using phenology data

⇒ **Large differences between modeling groups**

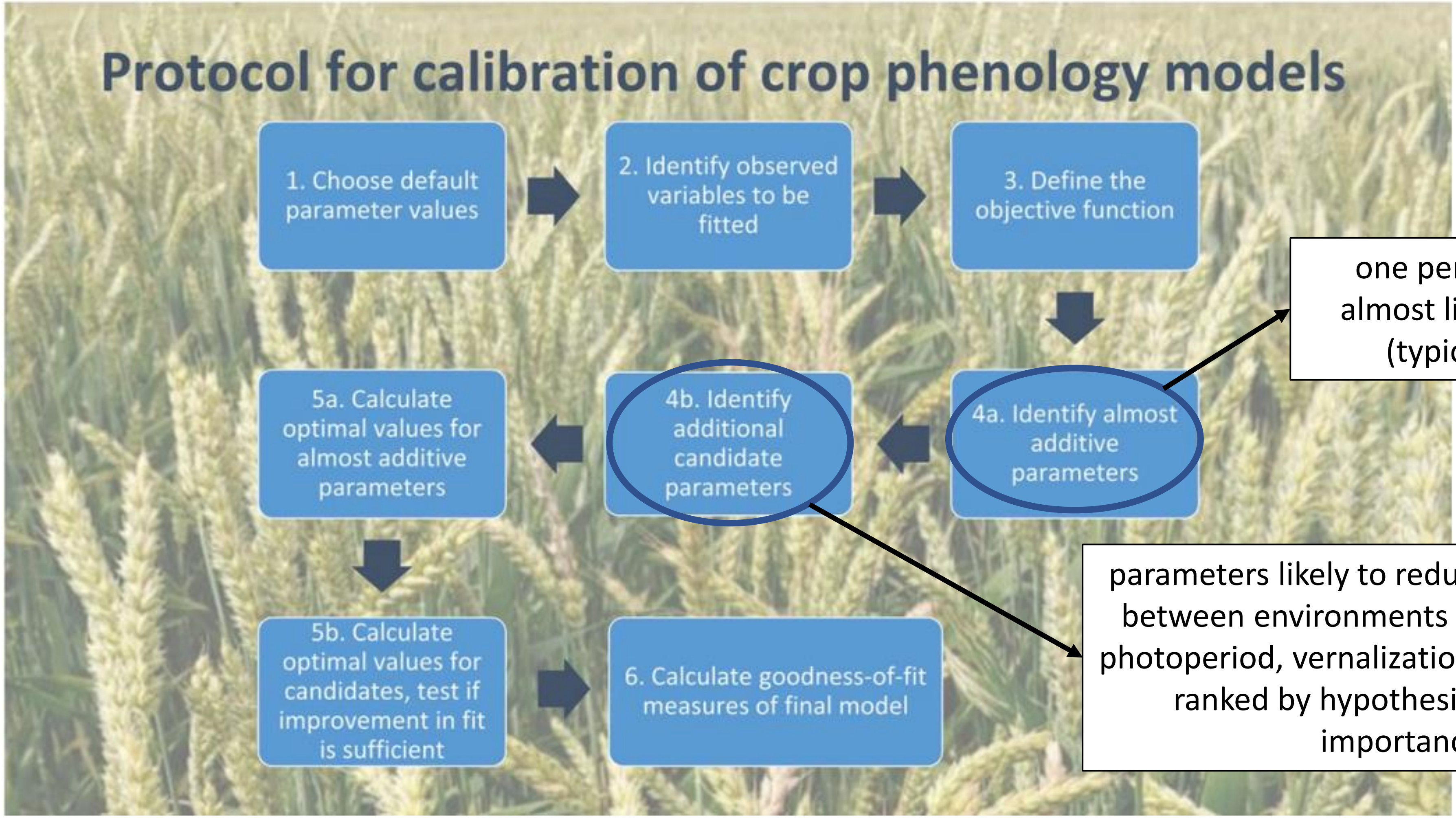
⇒ **No model was ranked first on both datasets**

⇒ **There are modeling groups which performed better than others over a wide range of environments**

⇒ **STICS is on the Pareto front: it offers one of the best compromises**



Phase III (2021-2022): let's do the same but with a common methodology



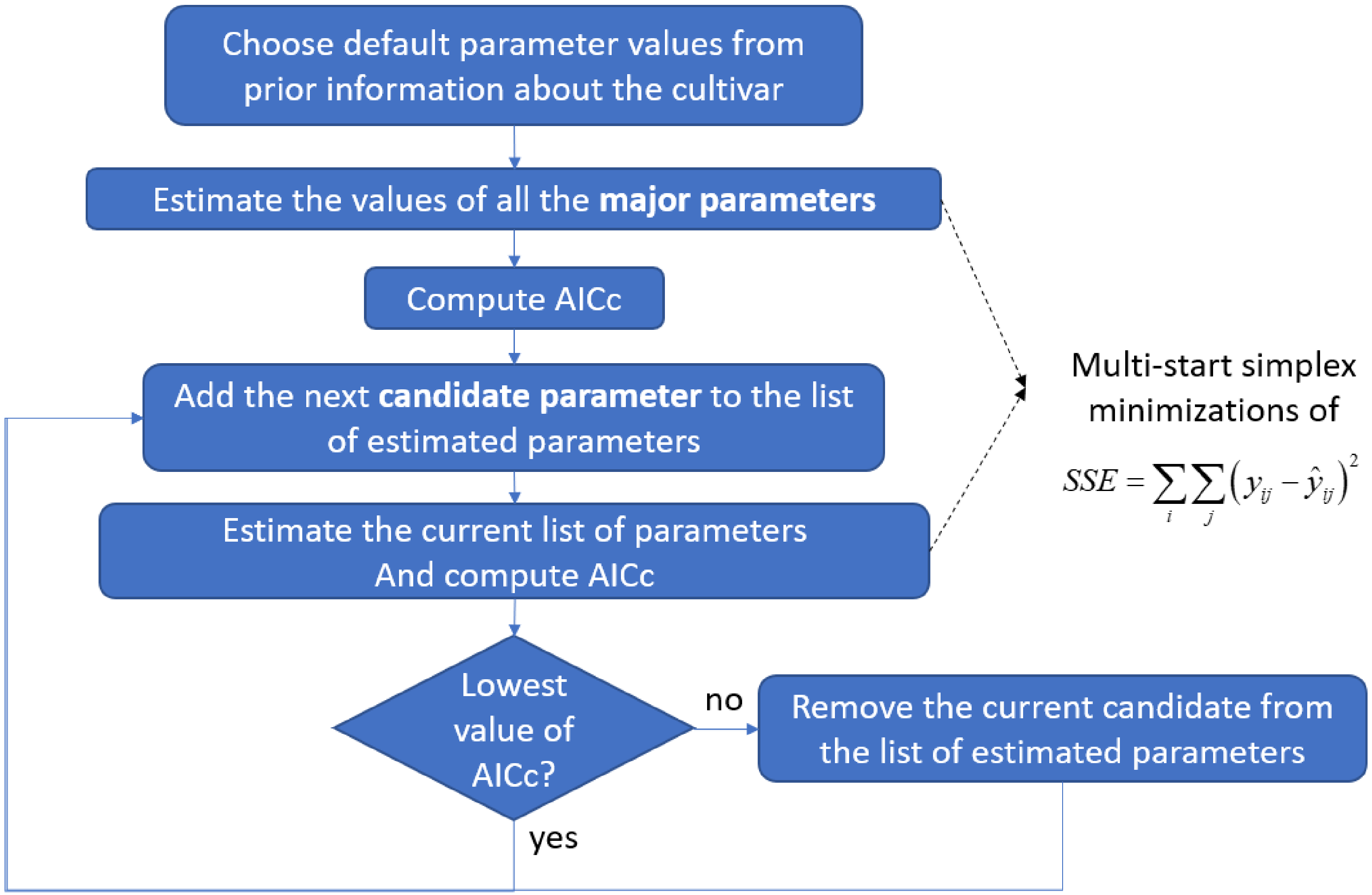
one per observed stage, almost linearly dependent (typically sum of T°)

parameters likely to reduce the variability between environments (e.g. linked with photoperiod, vernalization, water stress …) ranked by hypothesized order of importance

=> Results published in Wallach et al., 2023, ASD; Wallach et al., 2023 (under review in AFM)

Phase III (2021-2022): let's do the same but with a common methodology

Step 5



$$SSE = \sum_i \sum_j \left(y_{ij} - \hat{y}_{ij} \right)^2$$

Example for the STICS model, French dataset

Estimated parameters	Sum of squared errors	BIC
stlevamf, stamflax	405	81.47
stlevamf, stamflax, jvc	349	80.64
stlevamf, stamflax, jvc, sensrsec	322	81.71
stlevamf, stamflax, jvc, belong	349	83.97
stlevamf, stamflax, jvc, jvcmini	319	81.45
stlevamf, stamflax, jvc, stressdev	349	83.97

Total cost : (n° of candidates + 1) = 6 estimations

(Forward regression : 28 estimations,
All combinations : 127)

Phase III (2021-2022): let’s do the same but with a common methodology

- **Implementation in R provided**
- **Based on CroptimizR and CroPlotR packages**
- **More than 10 modelling groups used it in Phase III (>40%)**

CroptimizR 0.5.1 Get started Reference Articles Changelog

Estimating phenology following AgMIP-calibration phase III protocol

Samuel Buis

2023-01-06

Source: `vignettes/AgMIP_Calibration_Phenology_protocol.Rmd`

Introduction

This document presents how the AgMIP phase III protocol, designed to calibrate phenology of crop models, can be easily implemented using CroptimizR and CroPlotR packages.

This protocol is described in detail in Wallach et al (2022).

Its objective is to improve crop model parameters values using calibration for prediction of crop phenology.

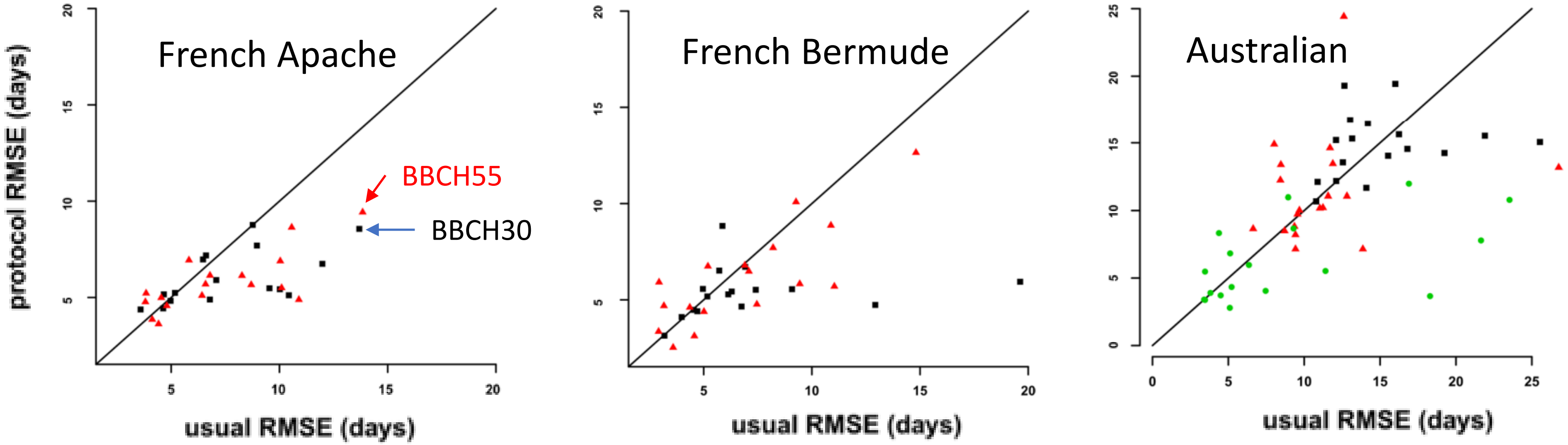
Here is a brief description of its different steps:

- Define the default values for all parameters that impact the simulation of phenology in the considered crop model.
- Identify measured variables (phenological stages) to be fit within the dataset available.
- Identify the nearly additive, obligatory parameters to be calibrated. The obligatory parameters are parameters that are nearly additive, i.e. such that changing the parameter has a similar effect for all environments (typically parameters that control degree days to the measured stages). Ideally, the number of almost additive parameters will be identical to the number of measured

Contents

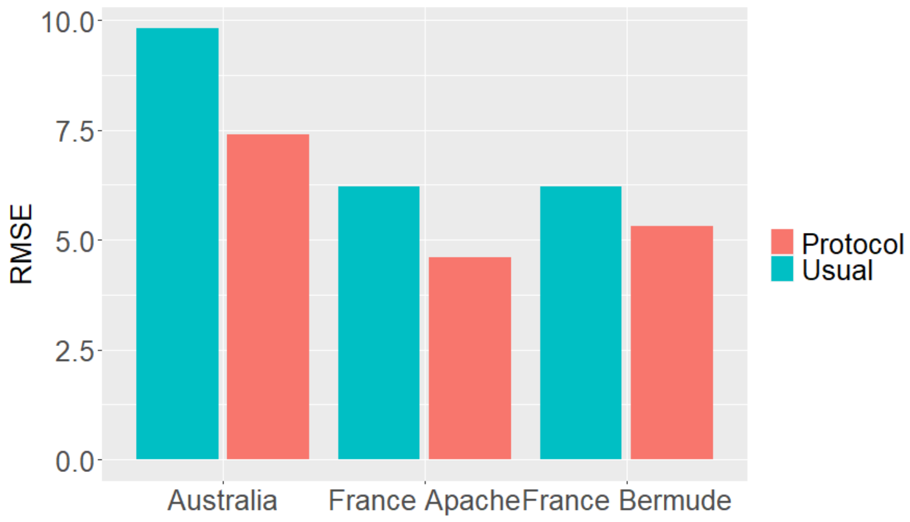
Phase III (2021-2022): let's do the same but with a common methodology

Results



⇒ **On average over modeling teams, the protocol led to a better fit to the evaluation data**

⇒ **The error of e-mean and e-median nearly identical with usual and protocol calibration, but the protocol reduced their uncertainty**



Phase III (2021-2022): let's do the same but with a common methodology

Variability explained by model parameterization

- based on variability in simulated values between different modeling groups using the same model structure :

$$y_{sp} = \mu + \alpha_s + \beta_{sp}$$

$$\sigma^2_{total} = \sigma^2_{structure} + \sigma^2_{parameters}$$

- based on a comparison between two different calibration procedures: usual and new protocol

$$\mathrm{var}_i = \left(y^{usual} - \bar{y}\right)^2 + \left(y^{protocol} - \bar{y}\right)^2$$

$$\hat{\sigma}^2_{parameters} = 1/n \sum_{i=1}^{R} \mathrm{var}_i$$

⇒ **Both methods lead to the same result: almost 70% of results variability is explained by model parameterization, much more than what was reported in literature up to now**

Phase IV (2022-…): now, let's consider biomass and yield data in addition

Datasets: same cultivars and environments, but more observed variables

Example of the French dataset

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/m²	ears	3	0
grains/m²	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

Phase IV (2022-…): now, let's consider biomass and yield data in addition

Model expertise steps

1. Choose default parameter values
2. List observed variables and corresponding simulated variables
3. Define variable groups and their order for calibration
4. Specify major parameters for each variable group
5. Specify candidate parameters for each variable group

Protocol

same as for Phase III but iterated on the different groups of variables

Calculation steps

6. For each variable group select candidate parameters to estimate and estimate values
7. Estimate all parameters together, using all data
8. Evaluate goodness-of-fit

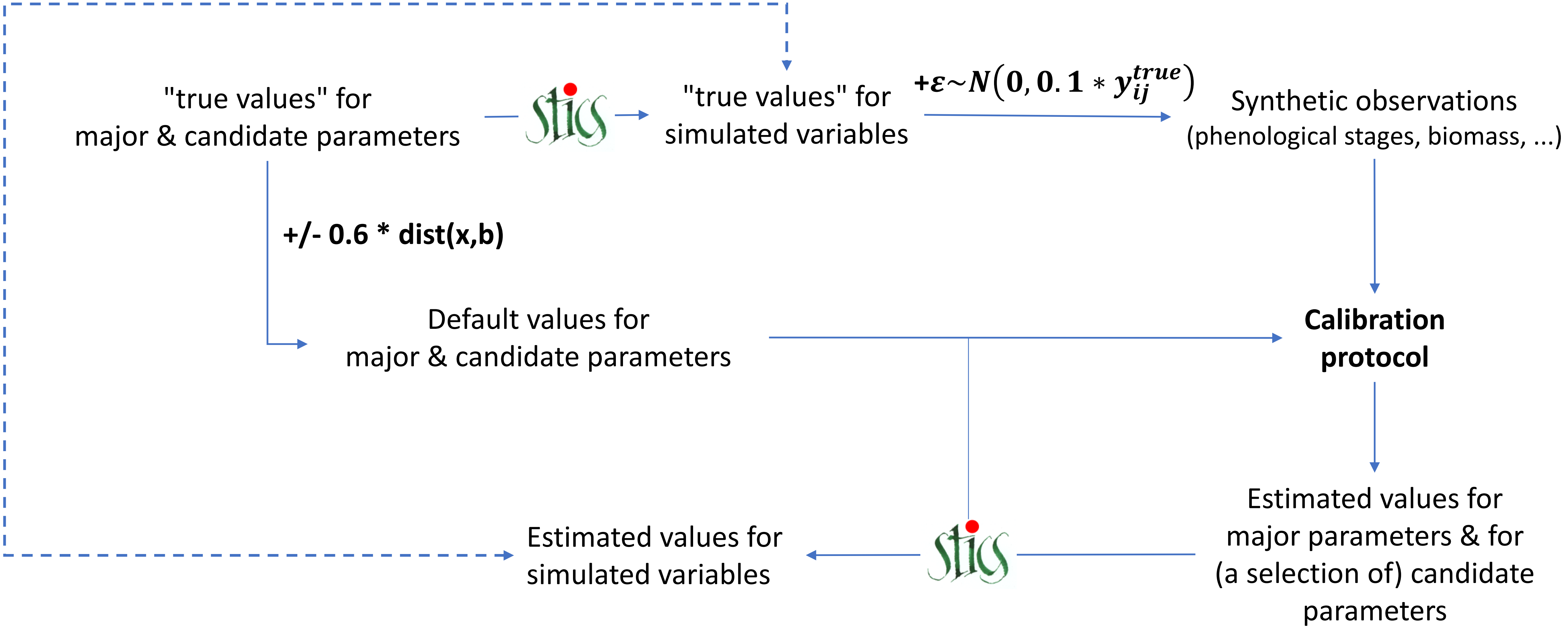
Multi-start simplex minimizations of Weighted Least Squares

$$WLS = \sum_i \frac{\sum_j \left(y_{ij} - \widehat{y_{ij}}\right)^2}{s_i}$$

where s_i is the standard deviation of model error for group i, estimated following step 6

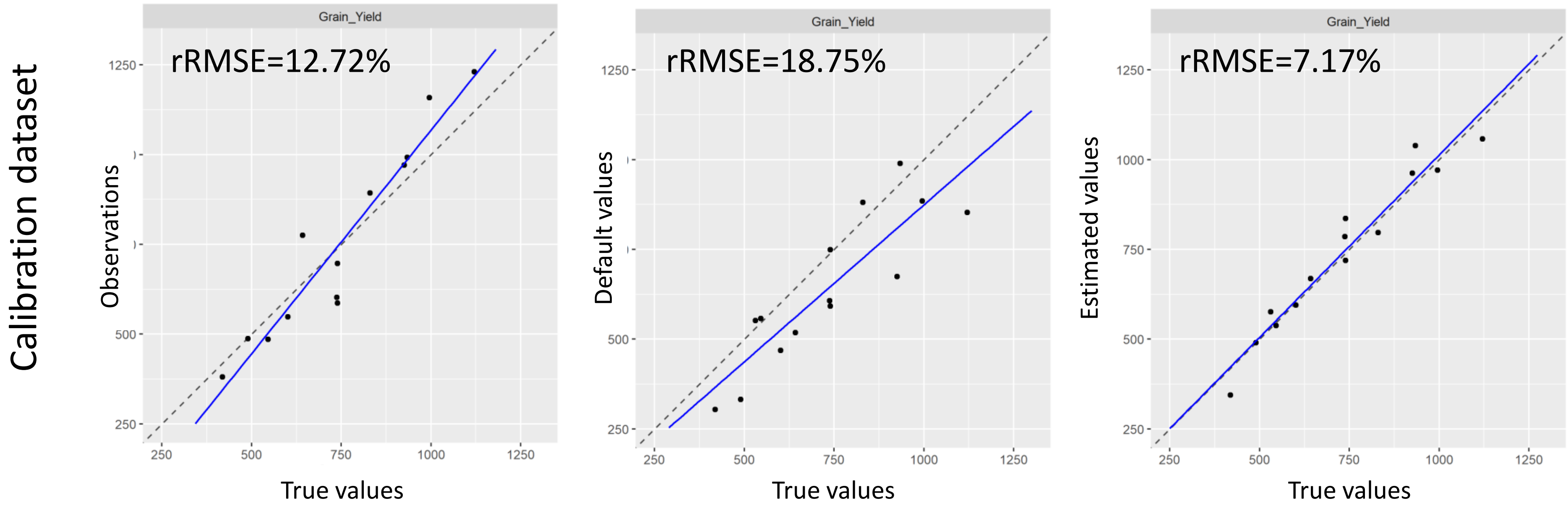
Phase IV (2022-…): now, let's consider biomass and yield data in addition

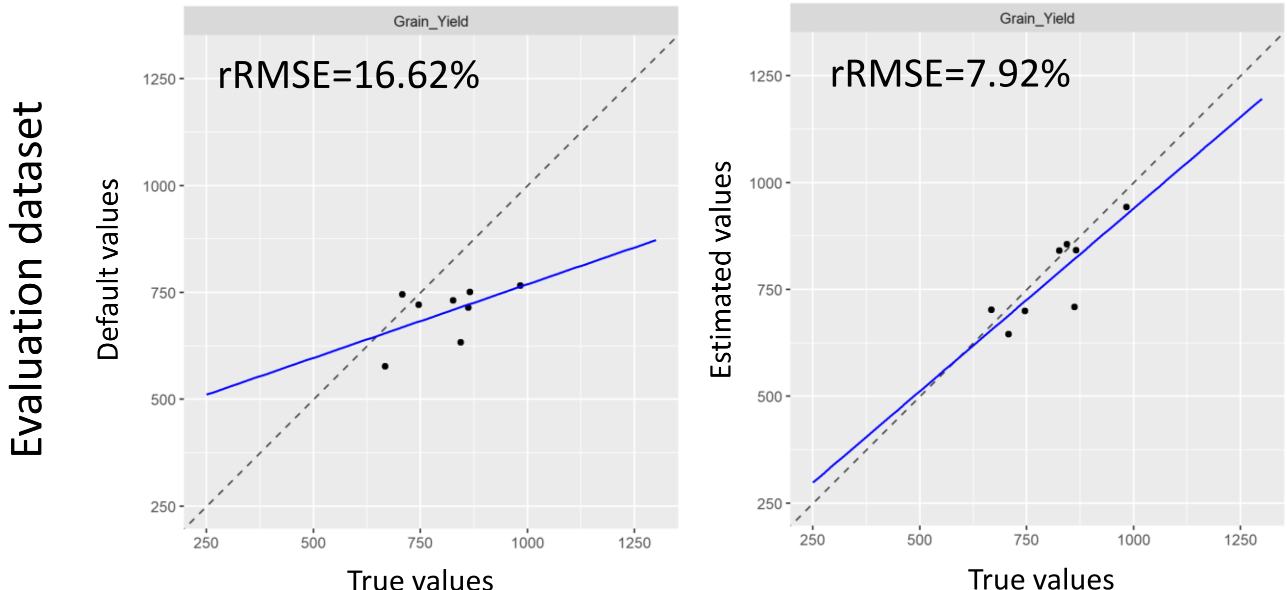
Synthetic experiments



Phase IV (2022-…): now, let’s consider biomass and yield data in addition

Results





⇒ **Calibration substantially improved the fit to the calibration and evaluation data**

Conclusions

- **Many crop model intercomparison studies have shown large variability in crop model results**

- **We have shown that the way crop models are calibrated may largely explain this variability**

- **Up to now, there is no consensus on how to calibrate crop models**

- **AgMIP calibration project proposed protocols and software to calibrate crop models**

- **The first evaluations of these protocols have shown their usefulness in the context of multi-model studies**

Perspectives

- **Apply Phase IV protocol to real data on a (large) ensemble of crop models: implementation based on CroptimizR and CroPlotR provided**

- **Apply Phase IV protocol to other datasets, compare it to other methods, test adaptations ...**
  => to publish in a special issue dedicated to crop model calibration end 2024 in EJA