A calibration protocol for soil-crop models aimed at reducing prediction error and inter-model variability

3.1 Description of the calibration protocol

The proposed protocol for soil-crop model calibration, described in detail below, is composed of eight steps. The first five steps (the model expertise part of the protocol) require detailed knowledge of the model and the data. No calculations are performed here. The result of these steps is a series of tables that contain all the model-specific information needed for the calculations. The last three steps (the calculation steps) describe the calculations to be done. The protocol includes instructions for each step, and the documentation to be produced in each step. The documentation is an integral part of the protocol, insuring transparency and reproducibility of the calibration procedure.

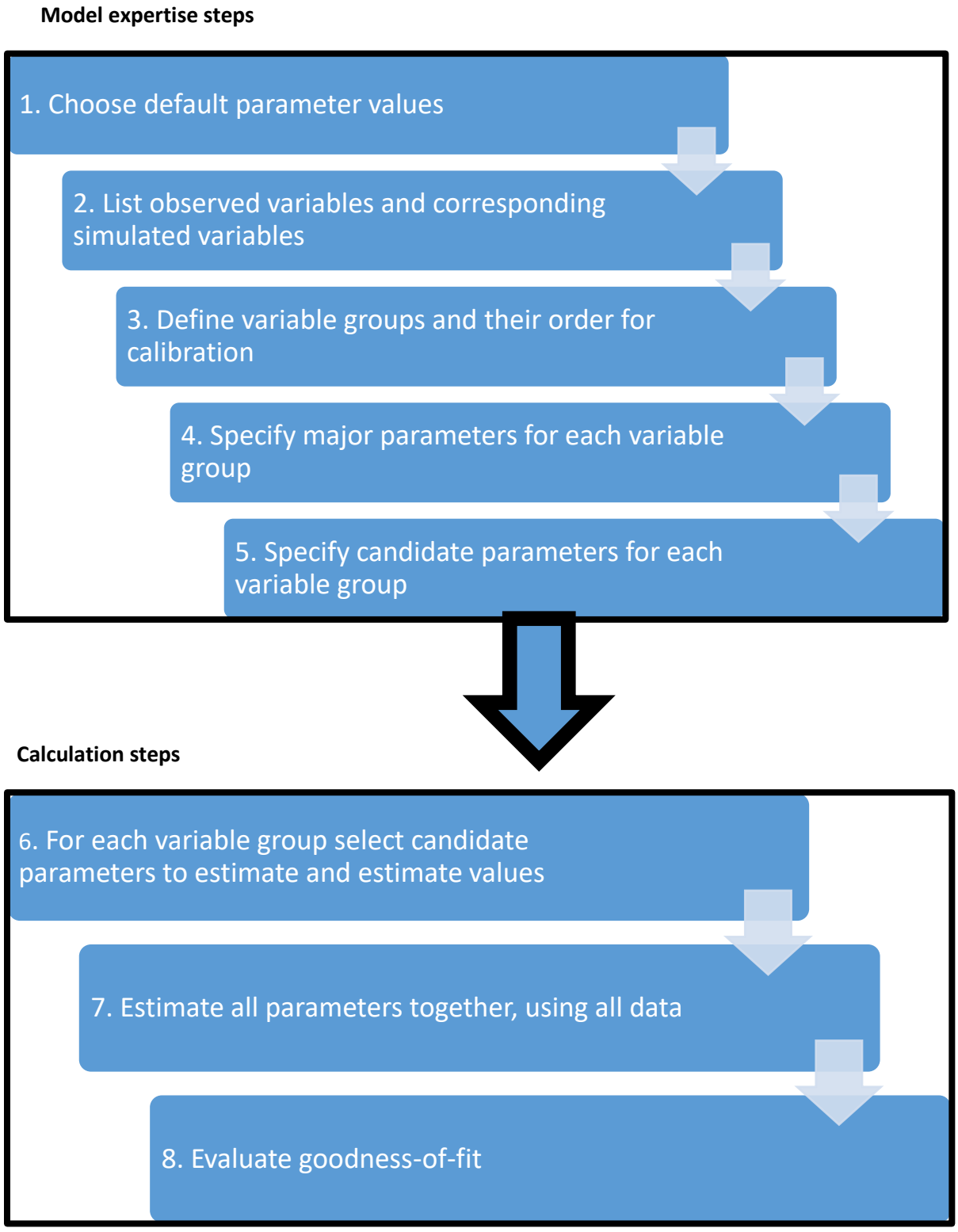

Figure 1: Schema of calibration protocol. The first five steps involve codifying model expertise. Given that information, the calculation steps 6-8 require no farther model-specific inputs.

Step 1. Explain choice of default parameter values and describe the calibration environments.

Only a small fraction of crop model parameters will usually be estimated from the data. The majority of the parameters will remain at their default values, so it is important to choose the default values with care. In particular, one should obtain as much information as possible about the cultivar characteristics (maturity class, photoperiod sensitivity, etc.) and choose default parameter values accordingly. The documentation required here (Table 1) contains the cultivar characteristics and the rationale for the choice of default parameter values.

Step 2. List observed variables and corresponding simulated variables, if any.

The purpose of this step is to identify the correspondence between observed and simulated variables. The documentation required for this step is a table with one row for each measured variable, showing also the corresponding simulated variable (for an example, see Sciebo file French_example_STICS_Buis, tab “variables”).

Step 3. Define groups of variables and order them.

The grouping of observed variables is fixed by the protocol All days to development stages are grouped together in a phenology group. All measurements of a given variable at different times (e.g. biomass) will also be in the same group. Other variables (including all final values such as final yield, grain number, grain protein content etc.) will each constitute a separate group. The order of the groups is the order in which they will be used for fitting the model. The order of the groups is important. It should be such that the simulated values of the variables in a group have little or no effect on the simulated values of variables in previous groups. Phenology will usually be the first group, since the days to different development stages are usually not affected or only slightly affected by other variables. The required documentation here, which is combined with the documentation for step 2, shows the group and order for each observed variable.

Step 4. Identify the major parameter or parameters for each group of variables

The purpose of this step is to identify the major parameters that affect each variable group. There is a strict limit on the number of major parameters for each group, to avoid over-parameterization. If there is only one variable in the group, there can only be one major parameter. For variables with at least two measurements in some environments (e.g. biomass with in-season measurements) there can be at most two major parameters (for example, one that determines rate of increase during vegetative growth and a second that determines rate of increase during reproductive growth).. For phenology, there can be as many major parameters as observed development stages with simulated equivalents. However, each major parameter must affect the time to a different stage.

The major parameters for a group of variables should have an effect on the simulated values in all environments. If a parameter is nearly additive, i.e. has nearly the same effect in all environments, then the estimation of the parameter will make the model bias nearly zero for the associated variable, which is desirable. Thus, the first choice of major parameter for a variable is a parameter that is nearly additive. Thermal degree days to a development stage is usually a nearly additive parameter for days to that stage, since increasing the required number of degree days will, in general, increase the days to the stage by a similar amount for all environments. Parameters that describe the effect of stresses, which only affect the simulated values if the stresses are present, will not be major parameters. The required documentation here is a table which shows the major parameters for each variable group (for an example, see Sciebo file French_example_STICS_Buis, tab “major variables”).

Step 5. Identify candidate parameters for each group of variables.

The candidate parameters are those parameters that are likely to explain a substantial part of the variability between environments and/or management strategies that remains after the major parameters are estimated. Each of these parameters will be tested (in the next step), and will only be included in the final list of parameters to estimate if estimation leads to a sufficient improvement in fit to the data.

The candidate parameters should be ordered from supposedly most to supposedly least important. The number of candidate parameters is not limited, but it is recommended to keep the number fairly small. The required documentation here is a table with the candidate parameters for each variable group (or an example, see Sciebo file French_example_STICS_Buis, tab "candidate parameters").

Step 6. Selection of parameters to estimate for each variable group and first estimation of their values

In this step, each group of variables is treated separately, in the order chosen in step 3. A list of parameters to estimate for each group is initialized with the major parameters. The major parameters for the group are estimated using ordinary least squares (OLS), and the corrected Akaike Information Criterion (AICc Brewer et al., 2016; Chakrabarti and Ghosh, 2011) is calculated as

$$AICc = n\ln(SS/n) + 2p + \frac{2p(p+1)}{n-p-1} \qquad (1)$$

where SS is the sum of squared errors for all variables in the group, n is the number of data points and p the number of estimated parameters. This assumes that all model errors for the group are independent and identically normally distributed.

Once the major parameters have been estimated, each candidate parameter in turn is added tentatively to the list of parameters to be estimated. If estimating all the parameters on the list reduces AICc below the previous smallest value, the candidate is kept on the list of parameters to be estimated. Otherwise, the candidate is removed from the list of parameters to be estimated, and returns to its default value (see flow diagram in Figure 2). AICc is a standard model selection criterion, designed to choose the best predicting model even if none of the proposed models is the true model (Aho et al., 2014).

Biomass should be replaced by the natural logarithm of biomass before the calculation. The reason is that biomass values may go over a wide range of values during the growth period, with an associated increase in the standard deviation of model error. The log transformation will make the standard deviations approximately constant for all dates. Similarly, a log transformation should be used for any other variables expected to vary over a wide range over time.

In the example here, the Nelder-Mead simplex algorithm was used to find optimal parameter values, both in step 6 and step 7 (Nelder and Mead, 1965). However, other optimization algorithms could also be used. When testing candidate parameters, one of the starting points for optimization should be the previous best parameter values, since those will often be close to the new best values.

A first required documentation table here shows the result of adding each new candidate parameter, for each variable group (see example in Table). The optimum parameter values and the fit to the measured data after this step are combined with the documentation of step 7 (Table 3,

Table4).

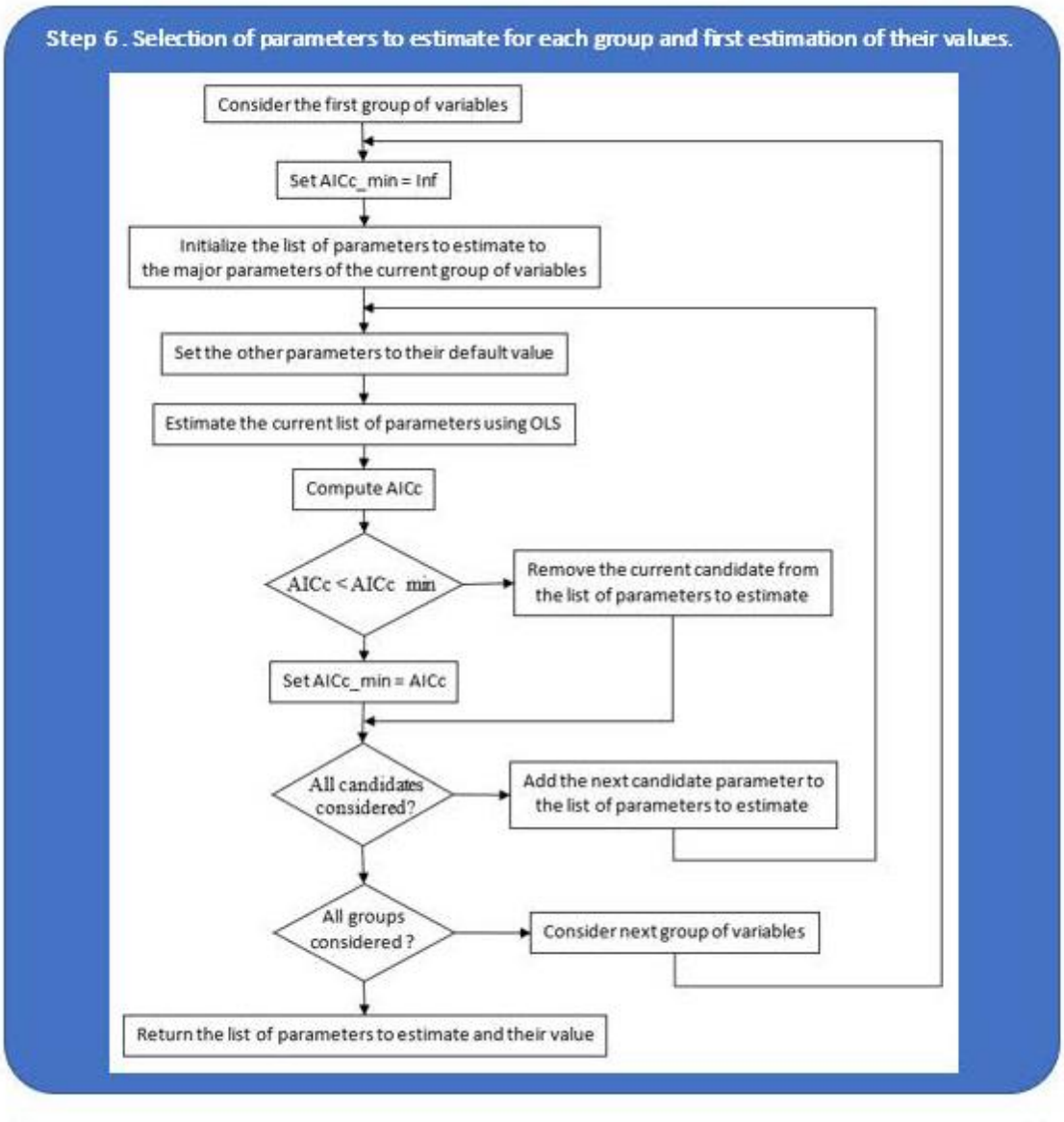

Figure 2:Flow diagram for the selection of parameters to estimate, and first estimation of their values, for one variable group, for step 6 of the protocol

Step 7. Re- estimation of all selected parameters using all variables simultaneously

In this step, all the selected parameters from step 6 are estimated together, using all the data, using weighted least squares (WLS). The objective function, to be minimized, is a sum of terms, one for each variable group. The term for each group is the sum of squared errors for that group, divided by *errVar*:

$$errVar = SS / (n - p) \qquad (2)$$

where SS is the sum of squared errors for all variables in the group from step 6, n is the number of data points for all variables in the group and p the number of selected parameters to estimate after step 6 for this group. The required documentation table here shows the estimated parameter values after steps 6 and 7 (see example in Table 3).

Step 8. Evaluation of goodness-of-fit

In this step metrics of goodness of fit are calculated for the simulations using the default parameter values, using the parameter values after step 6 and using the parameter values after step 7. The required documentation table here shows the metrics for goodness-of-fit at each stage. An example is shown in Table 4. Additional metrics could also be calculated. Graphs of simulated versus observed values for each variable should also be produced (for an example, see Sciebo files scatterPlots_XX.pdf in the Example of result files-folder).

Table 1 Description of cultivar that measurements refer to, and explanation of choice of default parameter values. Sections in red need to be modified by each modeling group.

Cultivar	Characteristics
Cultivar of measurements: A	A soft winter wheat. Stem elongation – semi-early. Heading – early. Vernalization requires 40 days where full vernalization occurs if daily average temperature is between 3°C and 10°C. There is no vernalization below -4°C or above 17°C. Otherwise there is a proportional reduction in vernalization effectiveness.
Cultivar used to provide default parameter values:	Explain here the origin of the default parameter values, and explain that choice.

Table 2: Selection of candidate parameters to estimate. Example of documentation for step 6 of the protocol. Each row shows the list of parameters to be estimated at that stage of the calculations. Candidate parameters that lead to a reduction in AICc compared to the previous smallest value are kept in the list of parameters to estimate. Otherwise, the candidate is removed from the list. This example is for the STICS model applied to the artificial calibration data used here, and for the variable group "biomass". There will be an analogous table for each group of variables. Sections in red need to be modified by each modeling group.

Table 2 Model: STICS contact: Samuel Buis Date(dd/mm/yy): 07/06/23			
Group	Parameters to be estimated	AICc	Candidate to be fit to data?
biomass	efcroiveg,efcroirepro	-127.89	yes (automatically)
biomass	efcroiveg,efcroirepro dlaimaxbrut	-166.13	yes
biomass	efcroiveg,efcroirepro dlaimaxbrut, durvieF	-164.97	no
biomass	efcroiveg,efcroirepro dlaimaxbrut, vlaimax	-178.70	yes
biomass	efcroiveg,efcroirepro dlaimaxbrut, vlaimax,psisto	-176.39	no
biomass	efcroiveg,efcroirepro dlaimaxbrut, vlaimax,psiturg	-177.02	no

Table 3: Parameter values. There is a row for each parameter that is estimated. This table is part of the documentation for steps 6 and 7 of the protocol. This example is for the STICS model applied to the artificial calibration data used here. The column of true values has been added here to facilitate evaluation of the protocol. In practical situations, the true values are unknown. Sections in red need to be modified by each modeling group.

Group	Estimated parameter	True value	Default value	Value after step 6	Value after step 7
phenology	stlevamf	212	324.8	202.14	204.54
phenology	stamflax	367	446.8	356.63	358.95
phenology	stdrpmat	700	820	686.23	686.31
phenology	belong	0.012	0.0228	0.0061	0.0066
phenology	stressdev	0	0.6	0.087	0.093
biomass	efcroiveg	4.25	5.3	4.43	4.55
biomass	efcroirepro	4.25	3.5	3.91	3.43
biomass	dlaimaxbrut	0.00047	0.00318	0.00031	0.00032
biomass	vlaimax	2.2	2.38	2.08	2.09
N in biomass	Vmax2	0.05	0.08	0.014	0.022
grain_number	cgrain	0.036	0.0324	0.037	0.035
grain yield	vitircarbT	0.0007	0.00031	0.00067	0.00068
grain_protein	vitirazo	0.0145	0.0064	0.015	0.014

Table 4: Relative root mean squared error (RRMSE) for the calibration data. The table shows RRMSE for the default parameter values and after parameter estimation in steps 6 and 7, for each variable. This table is part of the documentation for steps 6 and 7 of the protocol. This example is for the STICS model applied to the artificial calibration data used here. Sections in red need to be modified by each modeling group.

	BBCH30	BBCH55	BBCH90	ln(biomass)	N in biomass	grain number	grain yield	grain protein
Default parameter values	0.156	0.182	0.1588	0.051	0.22	0.31	0.25	0.286
After step 6	0.013	0.013	0.0066	0.021	0.13	0.12	0.17	0.071
After step 7	0.012	0.013	0.0065	0.018	0.12	0.11	0.16	0.076