User guidelines for running TSE software (Time-Series cultivar coefficient Estimator – for DSSAT crop models) – Standalone DSSAT plug-in

Emir Memic (University of Hohenheim, Stuttgart, Germany, February 2021)
E-mail: emir.memic@uni-hohenheim.de; emir_memic@windowslive.com

This cultivar coefficients optimisation program was developed and tested with DSSAT4.7 CROPGRO-Soybean model. The program was partially tested with CERES-Maize and Wheat models. The program was designed and written to enable optimisation of cultivar coefficients of all available crop models in DSSAT4.7 shell and should work with future versions.

Contents

1. The TSE concept overview

The cultivar coefficients estimation program was developed to work with DSSAT crop growth models (CROPGRO and CERES). It will work only if fully functional DSSAT files are available such as experiment file (FileX), weather file, soil profile, functional species, ecotype and cultivar files, time-series file containing in-season observations (FileT) and summary file (FileA - optional) are available.

Overall program run can be separated in three steps:

1. Selection of the FileX and corresponding treatments defined in FileX. Based on the selected experiment corresponding in-season observations (as time-series data) are red in as temporal inputs (all available time-series observations can be used), and cultivar coefficients designated for optimisation (phenology- and/or growth-related) are selected with desired coefficient ranges (Figure 1, step 1).

2. For each cultivar coefficient combination defined in first step crop model is executed with simulated outputs being coupled with in-season observations and saved for later analysis (Figure 1, step 2).

3. After all cultivar coefficient combinations were used and simulated outputs coupled with in-season observations statistical analysis is conducted based on the target variables used in the cultivar coefficient estimation process and the combination with the best statistical agreement between simulated and observed is selected as “optimum”.

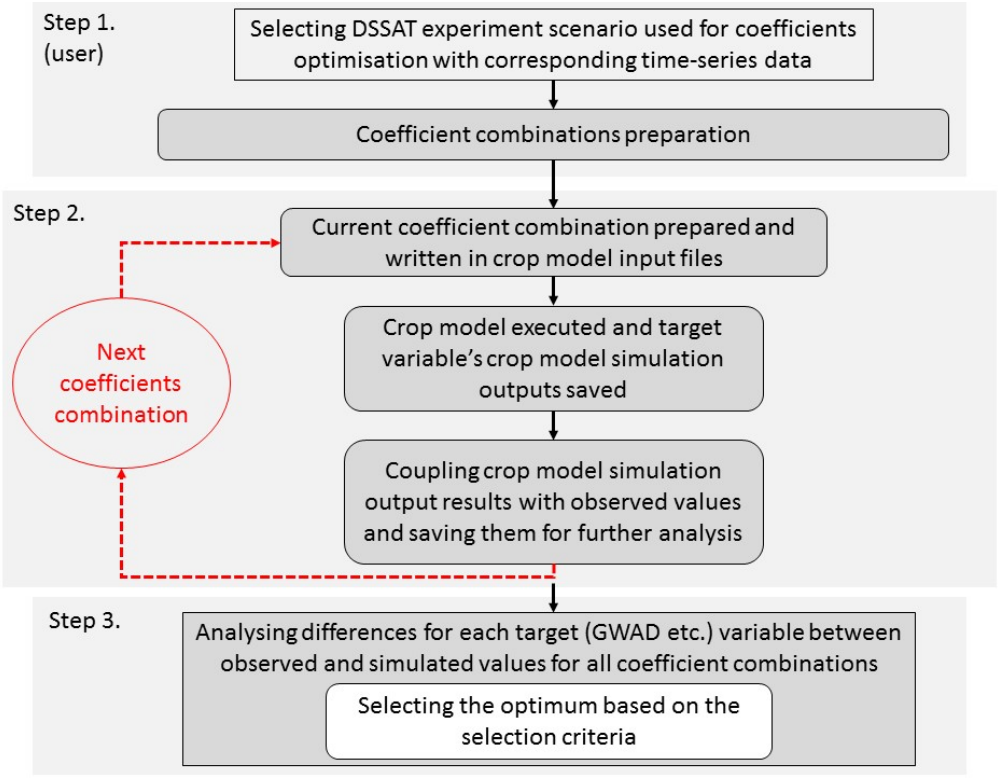


Figure 1: Flow chart showing overall program run in three steps (GWAD – grain weight) (Memic et al. 2021)

The program can be used for optimising cultivar coefficients based on single or multiple experiment data sets. User should first optimise phenology-related cultivar coefficients. After phenological events are correctly optimised, growth-related cultivar coefficients are optimised. Phenology-related cultivar coefficient optimisation is not conducted based on the time-series data (FileT), but by using FileA observations and simply minimising the difference between simulated and observed phenological event as day after planting. Growth-related cultivar coefficients are optimised by using FileT in-season observations by means of normalised RMSE throughout season. The combination providing lowest difference between simulated and observed values based on nRMSE is selected as optimum. Cultivar coefficient ranges initial values are predefined (by existing minimum and maximum values in cultivar files) but can be modified according to user's needs.

2. General TSE program settings overview

The program was developed (in Windows environment, for use on Windows) as standalone DSSAT plug-in and has no specific installation requirements. All required files for running the program are shared with gitHub_TSE_mmyy.zip file.

Before using the program user should look into User Guidelines shared with the program.

The "GitHub_[mmyy].zip" file has to be unzipped as "GitHub_[mmyy]" working directory. From "GitHub_[mmyy]" directory "**TSE**" folder copied to the Tools directory: "C:\DSSAT47\Tools" (Figure 2).

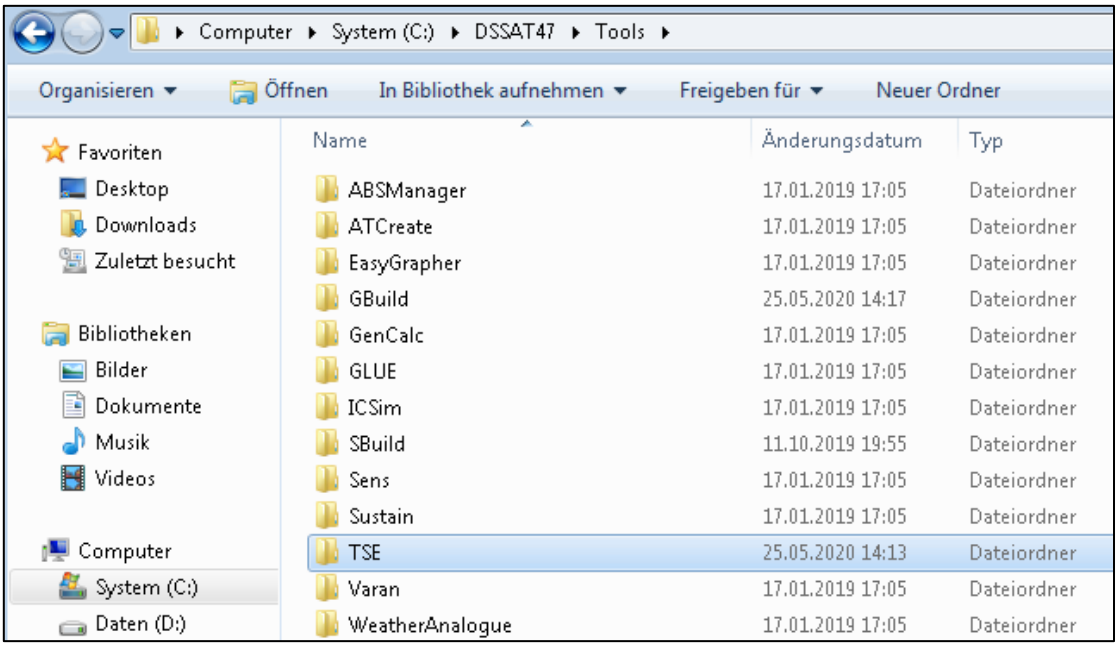


Figure 2 "C:\DSSAT47\Tools\TSE"

In folder "TSE" "C:\DSSAT47\Tools\TSE" (Figure 3) **"TSE_calibrator_DSSAT.exe"** windows runnable has to be **executed as "Administrator"** (Figure 4).

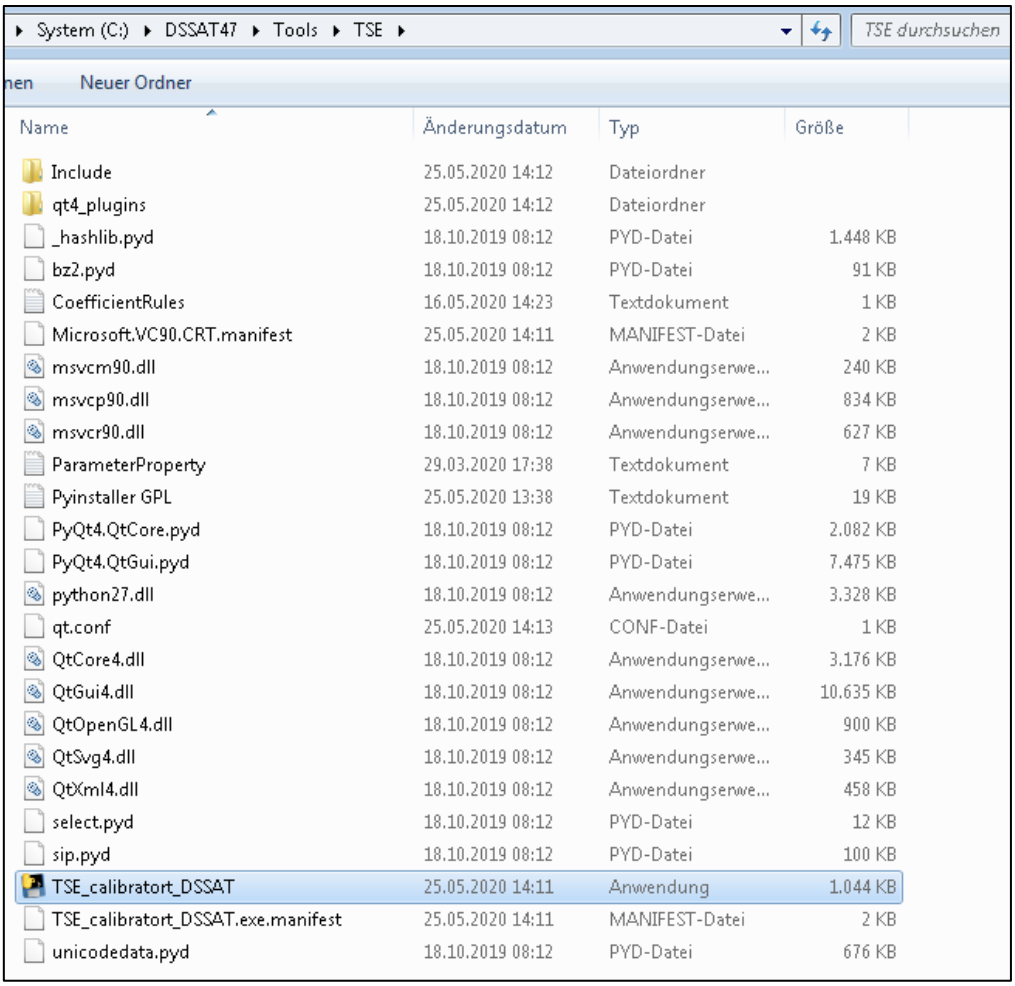


Figure 3 **"TSE_calibrator_DSSAT.exe"** windows runnable

VERY IMPORTANT:

I. PlantGro.Out crop model outputs are coupled to those in File-T (growth-related) time-series in-season observations. **Only variables available in File-T and PlantGro.OUT can be used in the optimisation process.**

II. Evaluate.Out crop model outputs are used (phenology-related) as DAY observations for optimising Phenological events (Still hard coded. Not enough freedom in optimisation process!)

III. If sub-model (eg. WHAPS) is initialised in the FileX, the calibrator might not work! (in File-X in *SIMULATION CONTROLS in GENERAL line, column SMODEL **do NOT initialise sub-models such as WHAPS, IXIM etc.!**)

IV. Only variables such as LAID, CWAD, GWAD etc. initialised in first time occurring "**@TRN….**" line in File-T are actively used by TSE

V. For multi TRT optimisations only target variables simultaneously available in all FileT/s (for corresponding FileX/s Treatment/s) are accessible for optimisation

VI. If PrameterProperty.txt is NOT in the C:\DSSAT47\Tools\TSE folder, **P**/**G** (**P**henlology/**G**rowth-related flags) are not available in the interface!

VII. File-T observations: **All in-season observations available including 0 are used!** Only "-99" values are ignored by the program.

VIII. The program is matching DOY from File-T with those in the PlantGro.OUT. If the user setup in the File-X reporting frequency for example every fifth day and exact observation DOY is not present in the PlantGro.OUT as it is written in the File-T, the program will not be able to match them for comparing simulated with observed.

The optimiser program is creating additional directory "**TSE_workspace**" (**C:\DSSAT47\TSE_workspace**) (Figure 5) and modifying the cultivar file in that directory, which is then executed by main DSSAT model executable.

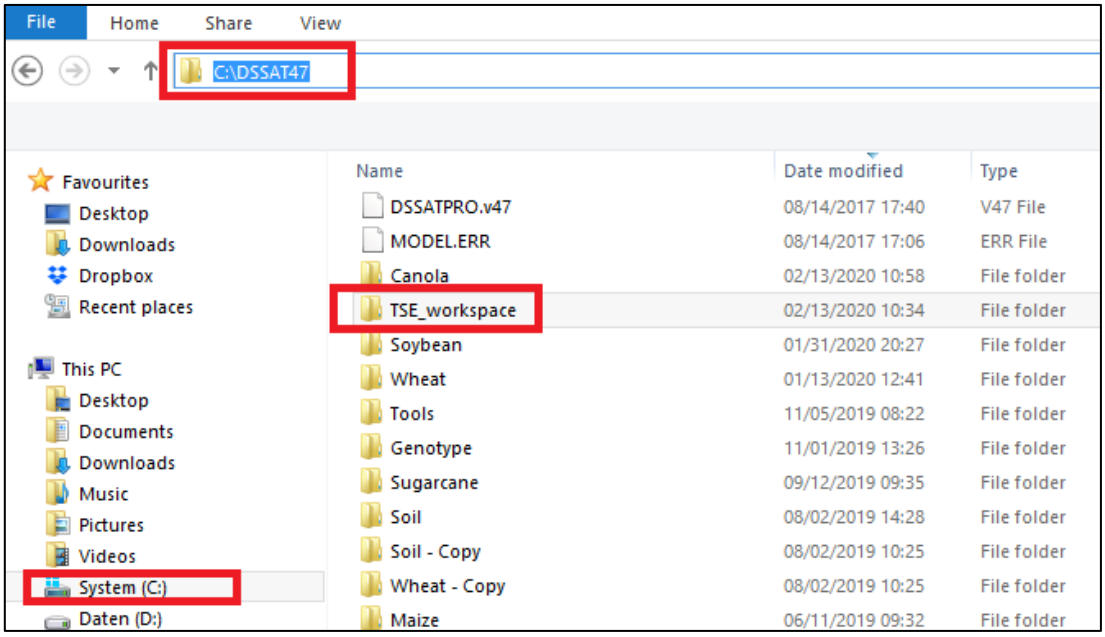


Figure 5 **C:\DSSAT47\TSE_workspace**

The optimiser program is NOT modifying core DSSAT files in their original directories!

Program run is considered: Start- "TSE_calibrator_DSSAT.exe" executed until "Exit" push button is pressed. Any form of optimisation done in-between is temporary saved in the cultivar file in "TSE_workspace" directory.

After TSE program is started (**TSE_calibrator_DSSAT.exe executed**) all modifications on Cultivar file are conducted in **C:\DSSAT47\TSE_workspace**. During one program run (until "**Exit**" push button is pressed) different coefficients (or different target variables) can be optimised one after another or simultaneously and cultivar changes will be saved if accepted as "optimums" in cultivar file in **C:\DSSAT47\TSE_workspace**. If user is satisfied with the cultivar coefficient values based on nRMSE fit cultivar coefficient combination should be copied to **C:\DSSAT47\Tools\Genotype** located cultivar file **MANUALLY**. If TSE program is started again without saving the combination in **C:\DSSAT47\Tools\Genotype** located cultivar file new TSE program start will copy original **C:\DSSAT47\Tools\Genotype** located cultivar file and overwrite user's working cultivar file in **C:\DSSAT47\TSE_workspace**.

After model run finished and before the user click "**Exit**" push button they can open GBuild and check visual and statistical fit (RMSE, d-statistics within GBuild) of the experiment file executed with the "optimum" genetic coefficient combination found in the last model run. With GBuild the user opens **PlantGro.OUT** from "**C:\DSSAT47\TSE_workspace**" directory,

because TSE will create parallel files it requires in this folder, without modifying the original files in DSSAT directory.

The more coefficients are “activated” (used in estimation process) the longer will optimisation last. **For each new coefficient and additional increment step (Inc) number of model runs will increase exponentially**.

For example:

Every time TSE_calibrator_DSSAT.exe is executed, original cultivar (SBGRO047.CUL) file from C:\DSSAT47\Genotype will be copied to C:\DSSAT47\TSE_workspace directory (Figure 6), and overwrite cultivar file in that directory (if exist, if not then just copied). If user wants to keep the genetic coefficient combination, it has to be copied to the original cultivar file in C:\DSSAT47\Genotype directory into SBGRO047.CUL manually.

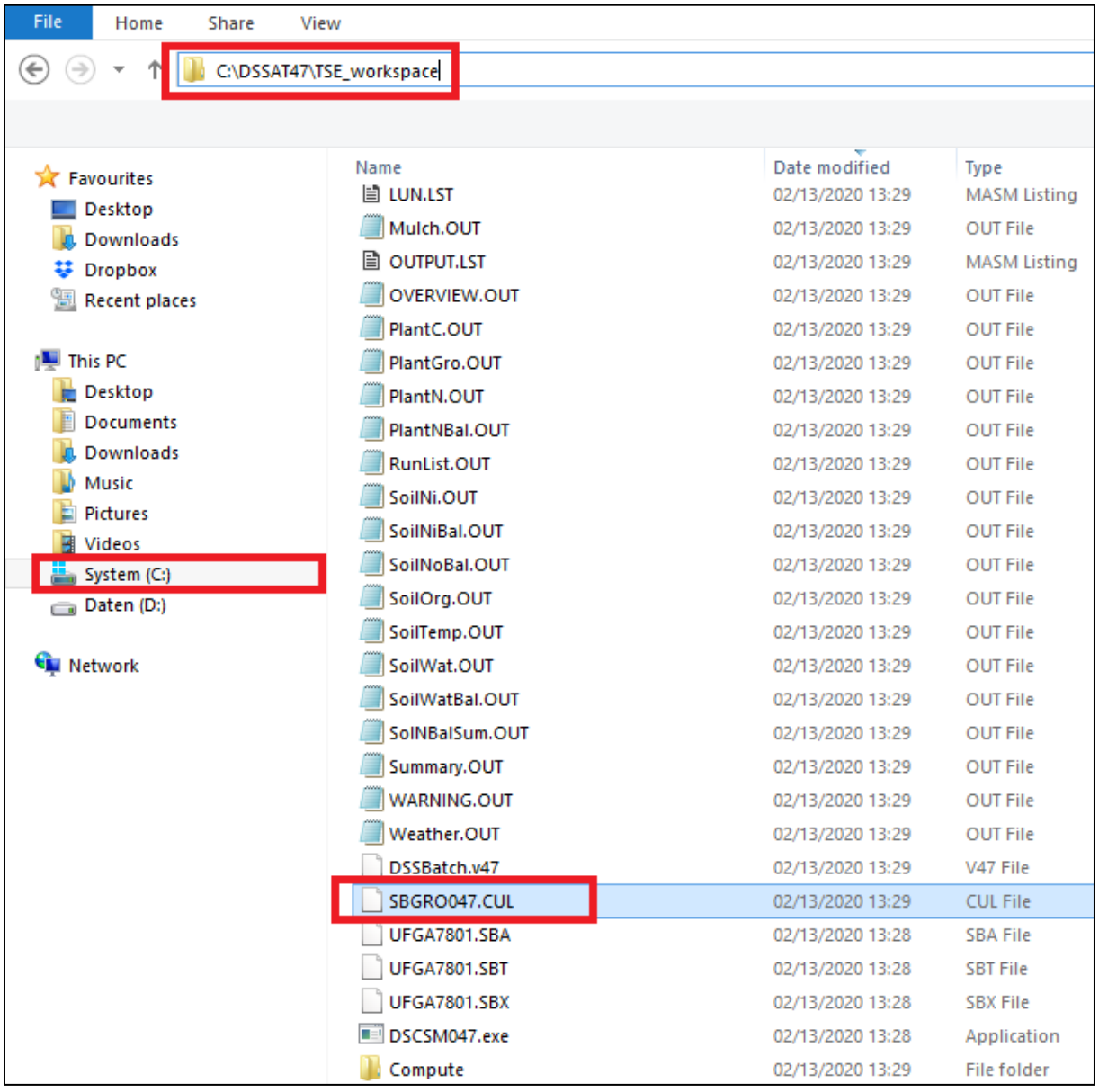


Figure 6 C:\DSSAT47\TSE_workspace\SBGRO047.CUL

The original cultivar is saved as **"!Old_timestamp_cultivarID…."** (Figure 7, text editor line 71) and the new one (Figure 7, text editor line 74) is saved in that working cultivar file (in **TSE_workspace**) and model is executed.

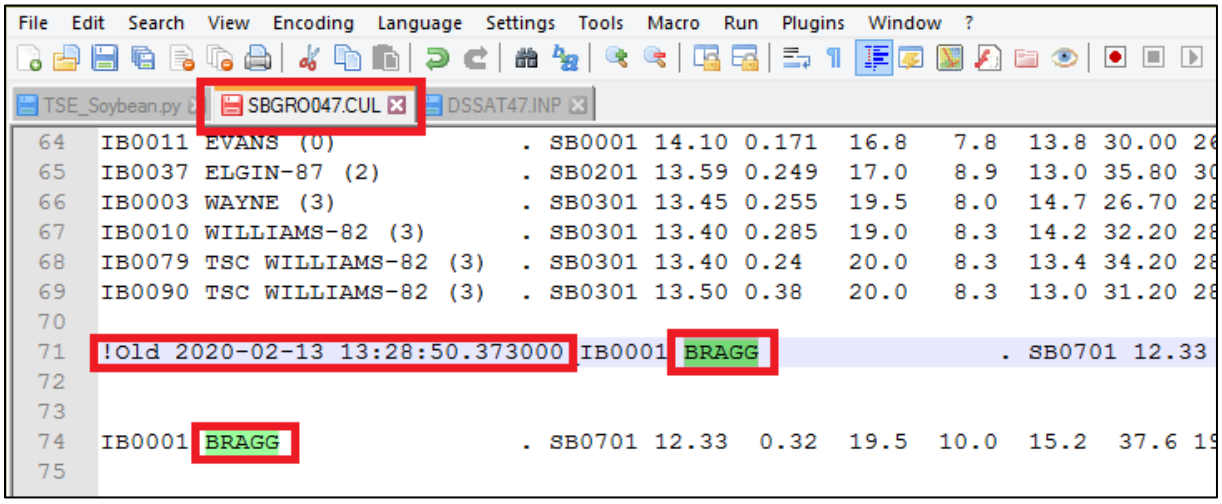


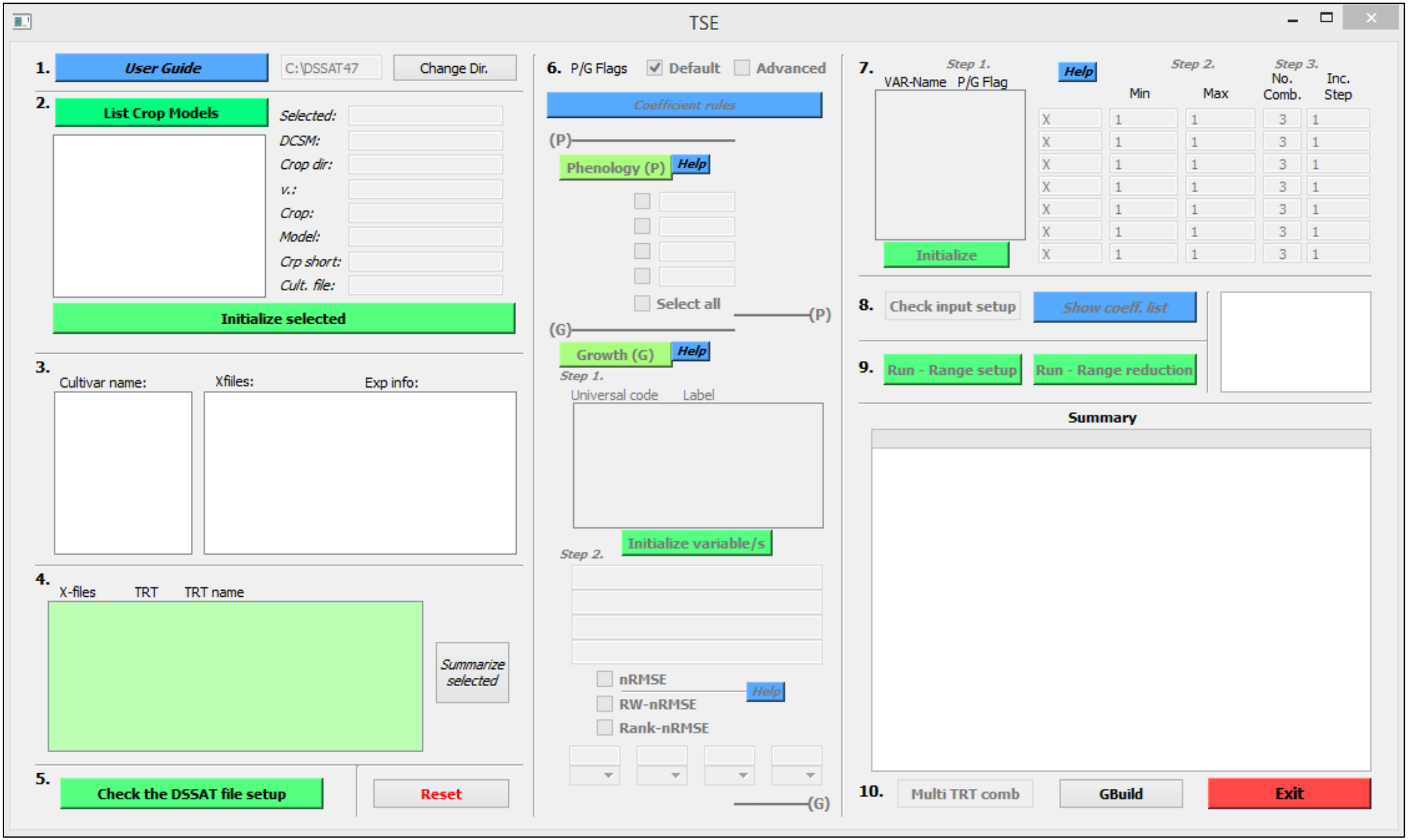
Figure 7 Old and new cultivar

3. Running TSE program

(The steps of preparing the estimator for run are enumerated in the interface)

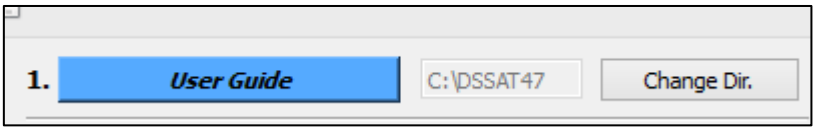
1. **If directory path shown is "C:\DSSAT47", do NOT modify! If the path is not "C:\DSSAT47" (This means that TSE folder was not copied to the "C:\DSSAT47\Tools"), then navigate to TSE folder and select it. It will be explained later in more details.**
2. **Select desired model and Initialize it!**
3. **Select cultivar from model corresponding list and File/s-X from list containing selected cultivar.**
4. **Select corresponding Treatment/s based on the File/s-X containing selected cultivar.**
5. **Execute selected treatment/s with DSSAT model to check if core DSSAT files are runnable, and select Default/Advanced.**
6. **Select optimisation of Phenology/Growth -related coefficients and corresponding methods.**
7. **Selecting desired coefficients and coefficient ranges and increment steps.**
8. **Check if optimisation software setup is correct.**
9. **Run the model!**
10. **Estimate Multi treatment based cultivar coefficient combination or execute GBuild!**

Interface

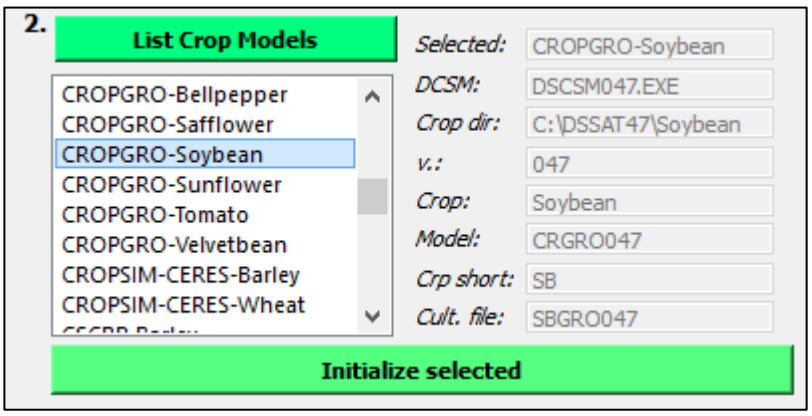

TSE
1.
User Guide
C:\DSSAT47
Change Dir.
2.
List Crop Models
Selected:
DCSM:
Crop dir:
v.:
Crop:
Model:
Crp short:
Cult. file:
Initialize selected
3.
Cultivar name:
Xfiles:
Exp info:
4.
X-files
TRT
TRT name
Summarize selected
5.
Check the DSSAT file setup
Reset
6.
P/G Flags
Default
Advanced
Coefficient rules
(P)
Phenology (P)
Help
Select all
(P)
(G)
Growth (G)
Help
Step 1.
Universal code
Label
Initialize variable/s
Step 2.
nRMSE
RW-nRMSE
Rank-nRMSE
Help
(G)
7.
Step 1.
VAR-Name
P/G Flag
Help
Step 2.
Min
Max
Step 3.
No. Comb.
Inc. Step
Initialize
8.
Check input setup
Show coeff. list
9.
Run - Range setup
Run - Range reduction
Summary
10.
Multi TRT comb
GBuild
Exit

1. Do NOT modify!

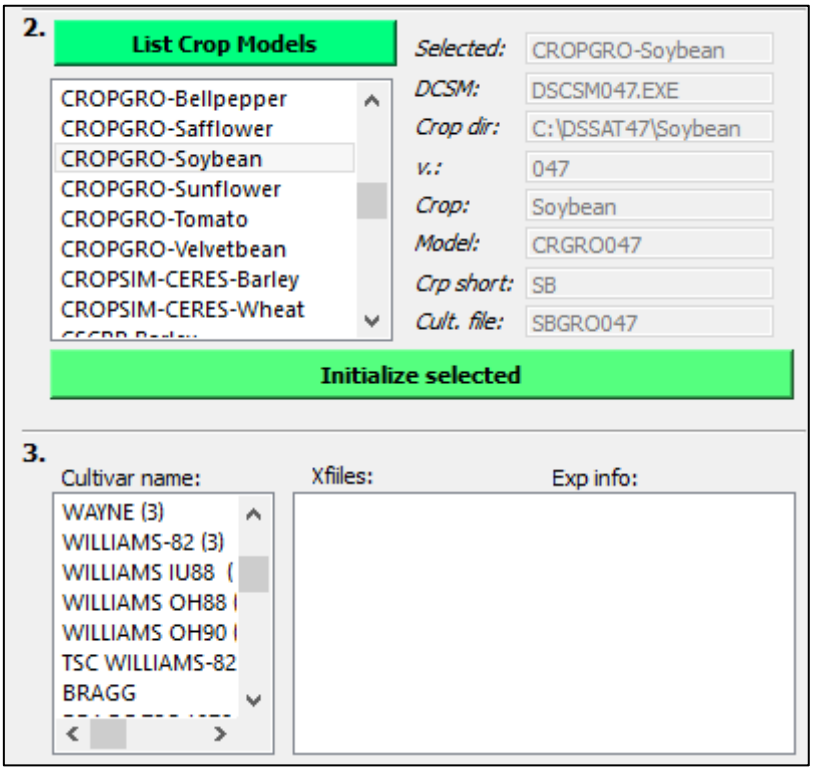
If directory path shown is "C:\DSSAT47", do ***NOT*** modify! If the path is not "C:\DSSAT47" then copy TSE folder (not entire "GitHub_[mm_yy]" extracted folder!), only the "TSE" folder from it into the "C:\DSSAT47\Tools". After re-executing **"TSE_calibrator_DSSAT.exe"** from TSE -> "C:\DSSAT47\Tools\TSE" the correct path should be read in.



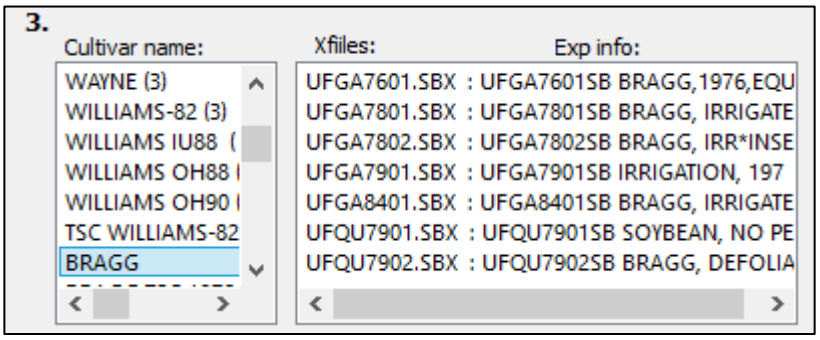
2. Select desired model and Initialize selected!



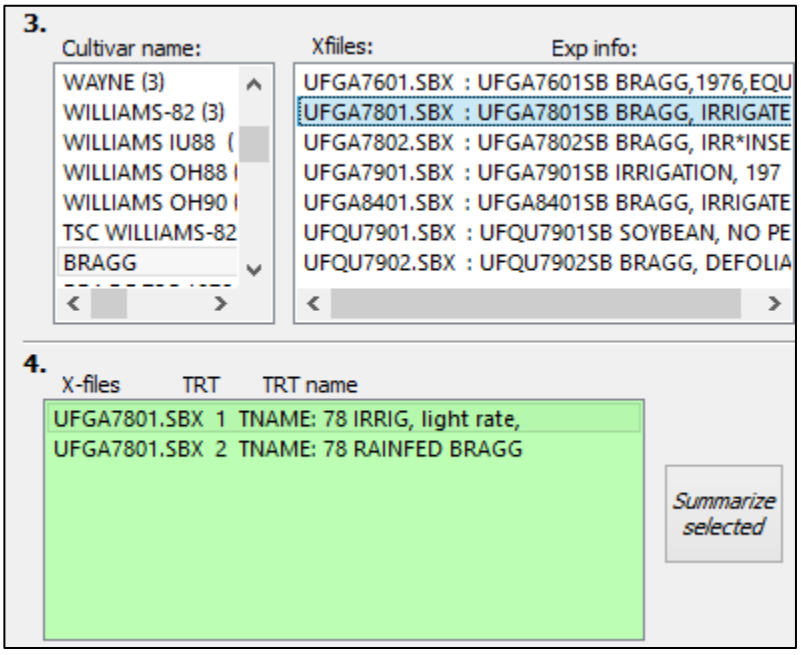
The list of model available in the DSSAT shell are uploaded into the TSE interface after List Crop Model is clicked. After selecting desired model the Initialize selected push button has to be clicked.



3. Select cultivar from model corresponding list and File/s-X from list containing selected cultivar.

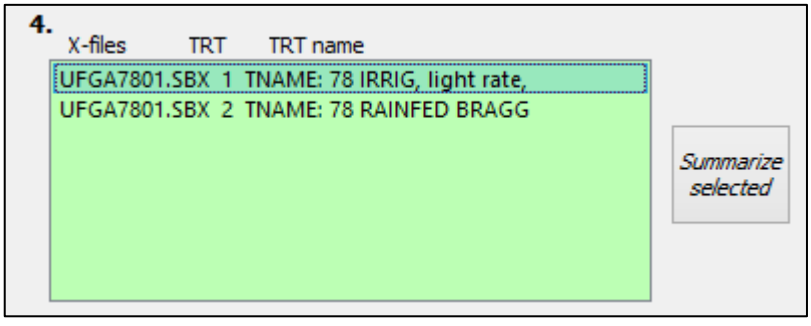


After the model is initialized the list of the model corresponding cultivar names are read in the interface ("Cultivar name :"). After cultivar name (e.g. "BRAGG") is selected the File/s-X containing that cultivar name are read in the interface.



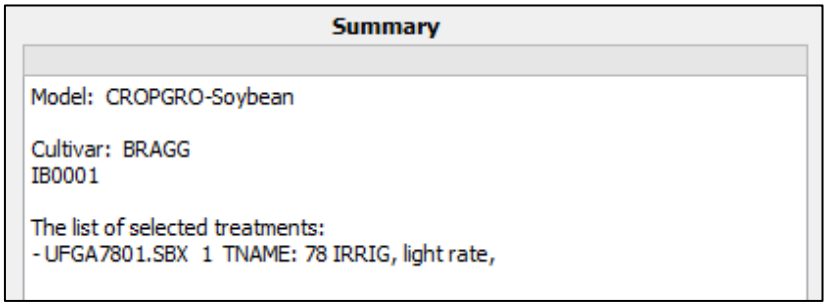
In the list File-X (or multiple File-Xs) can be selected as show (e.g. UFGA7801.SBX).

4. Select corresponding Treatment/s based on the File/s-X, containing selected cultivar.



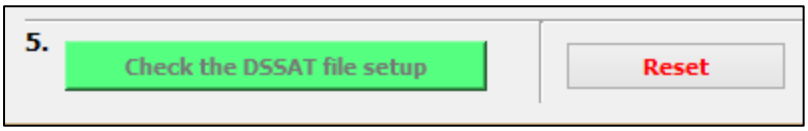
After selecting File-X in the interface the list of the treatments in the File-X containing the cultivar is shown and offered for selection (green list widget box).

The green box (green list widget box in the interface) is where the experiment file and treatments for optimising cultivar coefficients are **selected**. Multiple treatments can be selected (multiple treatment/experiment selection is done by: ***Ctrl+ mouse left button click*****).**



The "Summarize selected" push button will display selected options in the Summary window in the interface.

5. Execute (Check the DSSAT file setup) selected treatment/s with DSSAT model to check if core DSSAT files are runnable.

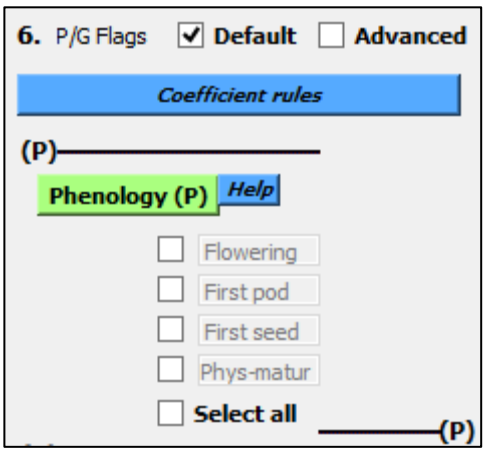


The "**Reset**" button will unlock the access to the selection of the crop model, cultivar and File-Xs, if selection was wrong.

6. Select optimisation of Phenology/Growth -related coefficients and corresponding methods.

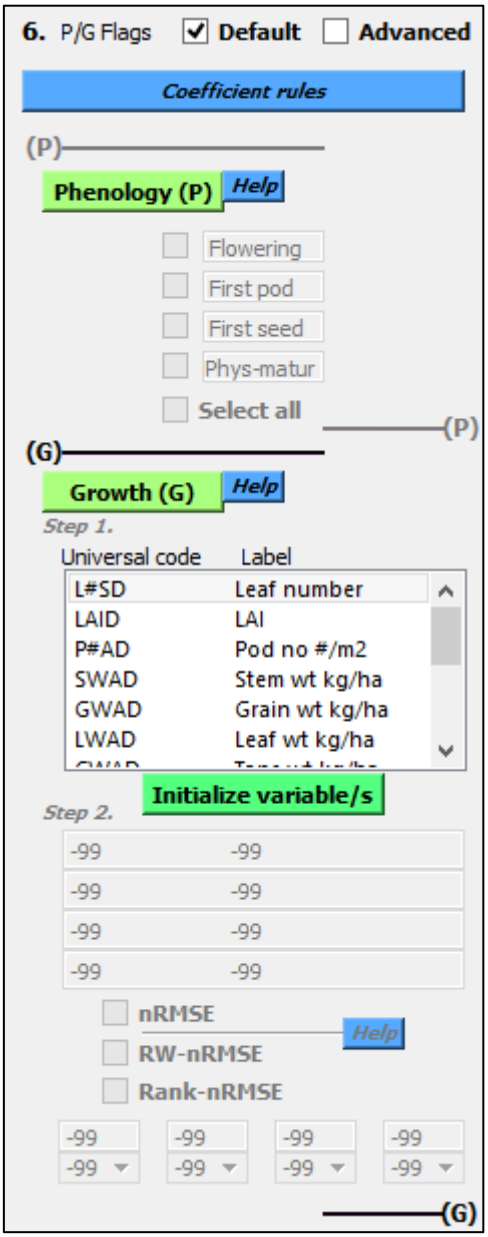
Cultivar specific coefficient flags (Phenology and Growth – P/G) are listed in the "C:\DSSAT47\Tools\TSE" in a file "ParameterProperty.txt". ***Default*** (check box) will upload only coefficients with predefined P/G flags (according to selection of the Phenology or Growth). ***Advanced*** (check box) will upload all available cultivar coefficients from cultivar file for potential optimisation.

Phenology:

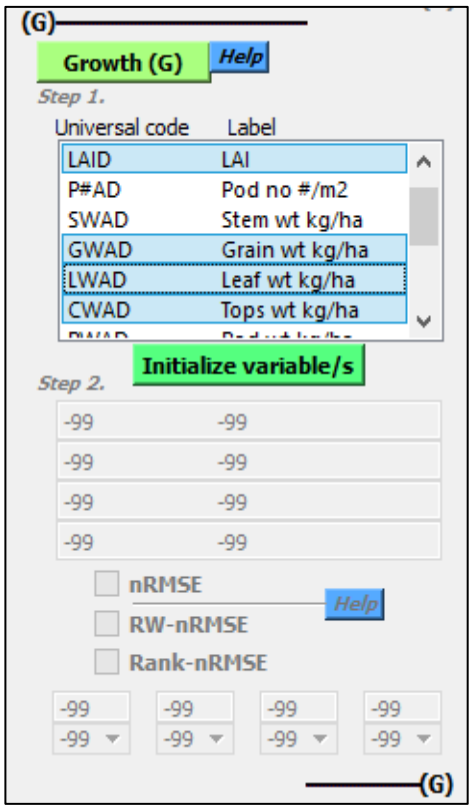


Currently Phenology optimisation is hard coded and will work only with these four phenological events. In future it will be made more flexible. This part of the program was so far tested with CROPGRO-Soybean and some CERES models.

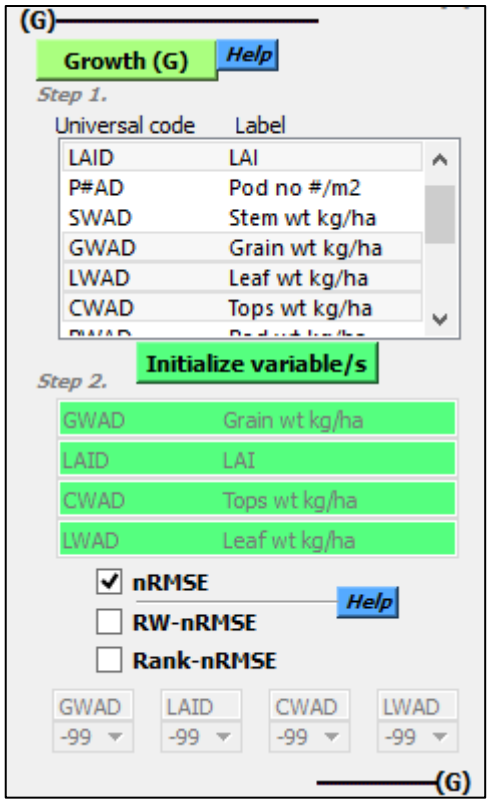
Growth:



In this step error minimisation method (nRMSE or RW-nRMSE or Rank -nRMSE, explained in the APPENDIX) is selected. Maximum four different target variables (e.g. LWAD, CWAD …) can be selected. If four are not available, one or two or three can be selected. In first round of optimisation four can be selected and in second next four, or with some other combinations.

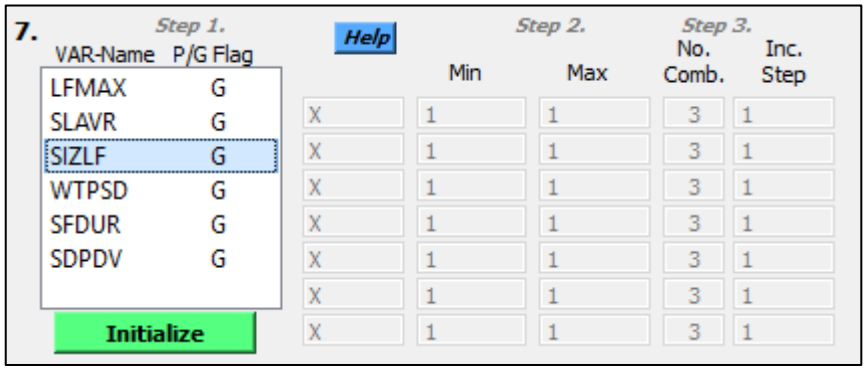


"*Ctrl* + *mouse left click*"- for selecting multiple target variables simultaneously!

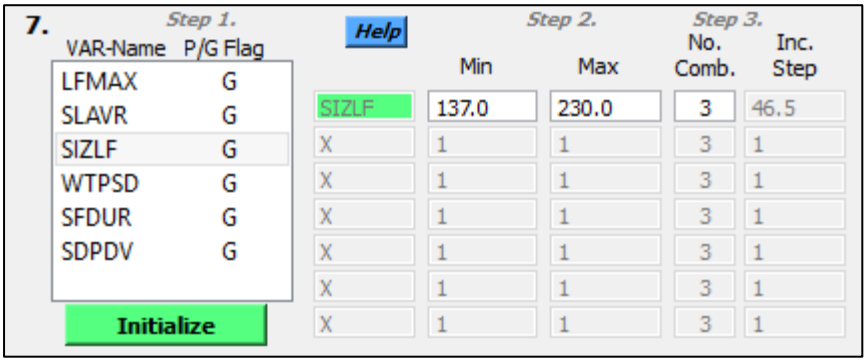


After selecting target variables Initialize target variable/s has to be clicked.

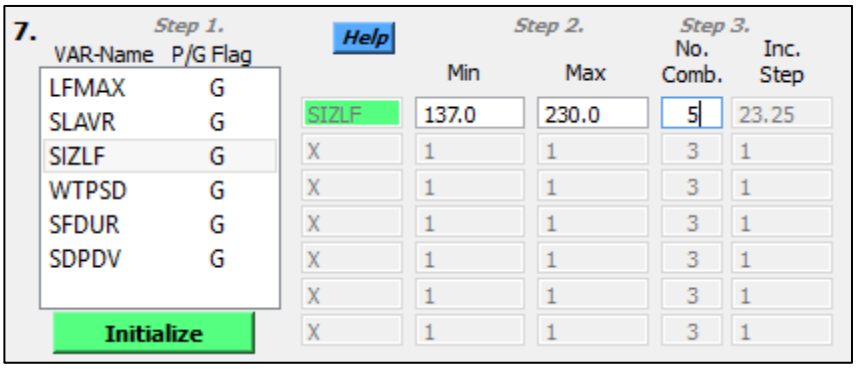
*7. **Selecting** desired coefficients and coefficient ranges and increment steps.*



First coefficient/s are selected from the list:



After selecting coefficient or multiple coefficients (***Ctrl+ mouse left button click*** – for multiple selection) "Initialize" push button is clicked!

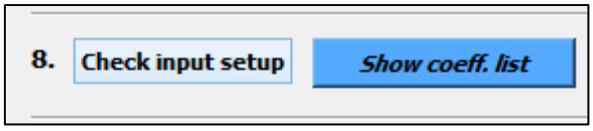


After selecting desired coefficients in section 7. from list widget window (step 1.) Initialize push button will initialize coefficient Labels in edit boxes and populate Min/Max coefficient ranges with min/max value available in corresponding cultivar file for selected cultivar coefficient (step 2.). Automatically this will create simple range of Min/Max value with three coefficient combinations that are going to be passed into the cultivar file, where after each model is executed. If the user wants more combinations in between given Min/Max range number of desired coefficient combinations between Min/Max can be given in "No. Comb." edit boxes. After giving the desired number of combinations between Min/Max increment steps are automatically calculated.

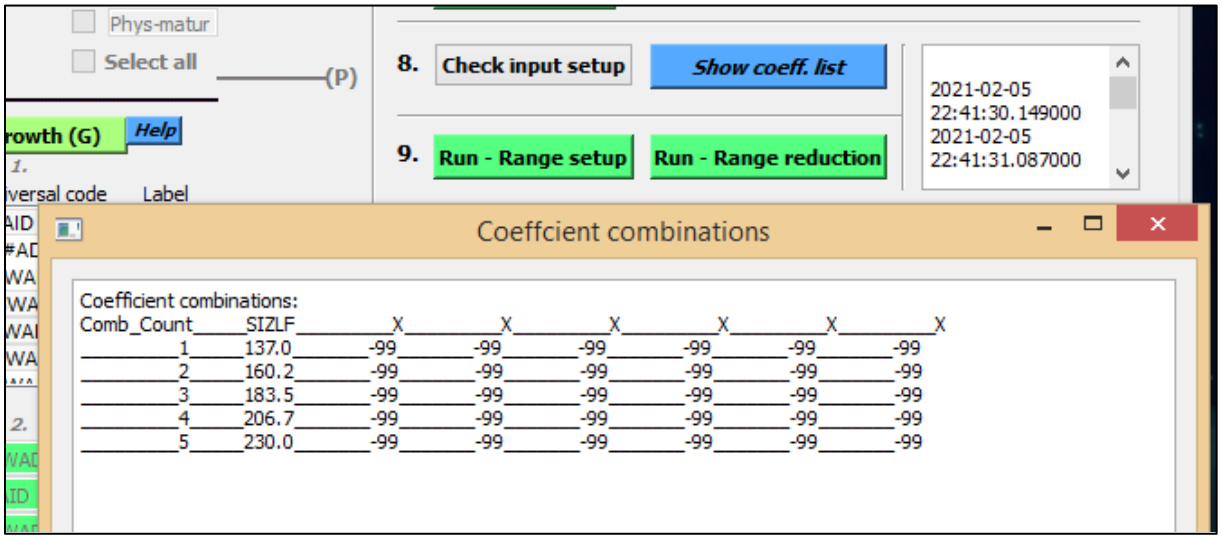
If user would like to modify Min/Max values it can be done directly (manually) in edit boxes below Min/Max labels in step 2. If Min/Max are modified based on the number of combinations values are automatically calculated and populated in corresponding edit boxes.

If the user wants to fix the value of some coefficient, in other words not to vary it during the program run, then they set Min and Max to equal value (<u>same value in the Min edit box as in the Max value edit box</u>).

8. Check if optimisation software setup is correct.



Show coeff. list push button will display coefficient to be used in the estimation process.



9. Run – Range setup & Run – Range reduction!

Range reduction is described on an example in Appendix A.5.

10. Additional options

Multi TRT comb

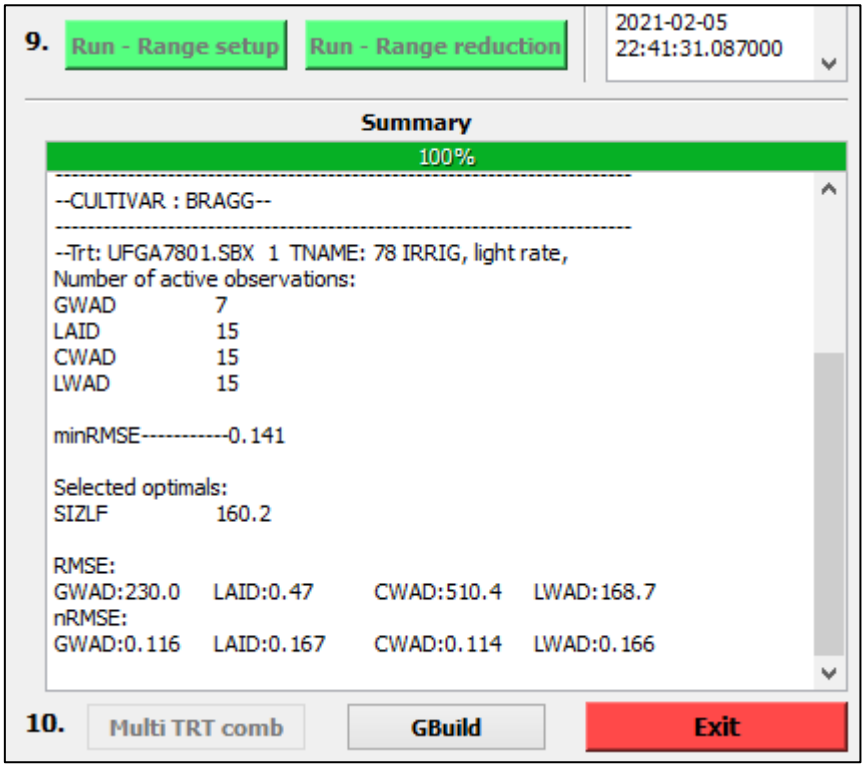
If this push button clicked multi treatment based cultivar coefficient combination will be estimated, as described in appendix A.2.

Gbuild

With **GBuild** PlantGro.Out can be opened **in TSE_workspace** directory to see how good new combination is. If more than one experiments were used (if Multi-TRT) combination was found, all used treatments can be seen in Gbuid simultaneously…

If the **PlantGro.Out** in "**C:\DSSAT47\TSE_workspace**" is open with GBuild visual and statistical fit across all used experiments of coefficient optimisation can be seen in GBuild.

Exit the program and all running threads



APPENDIX

A.1 The nRMSE error minimisation method (nRMSE)

For quantifying the variation between simulated (Si) and observed (Oi) values the statistical method of nRMSE (Eq. 2) is used. The nRMSE, is RMSE (Eq. 1) normalized by mean ($\overline{O}$) for each observed crop variable.

$$RMSE = \left[\frac{1}{n}\sum_{i=1}^{n}(S_i - O_i)^2\right]^{0.5} \qquad (1)$$

$$nRMSE = \frac{RMSE}{\bar{O}} \qquad (2)$$

The *nRMSE* is a simplified selection criterion that is applicable across multiple target variables with different scales. The coefficients are estimated across multiple target variables, with specific objective of locating coefficients combination with lowest *nRMSE* over all targeted variables (Eq. 3).

$$AVG_{nRMSE(n)} = (GWAD_{nRMSE(n)} + LAID_{nRMSE(n)} + CWAD_{nRMSE(n)} + LWAD_{nRMSE(n)})/4 \quad (3)$$

The selection of the coefficient combination with the lowest nRMSEs average across all target variables proves to be a good solution (Table 1, AVG-nRMSE=0.12).

Table *1* The *nRMSE* - simplified example of varying one of the cultivar parameters affecting growth (G) related target variables (GWAD, LAID, CWAD, and LWAD) of Gainesville 1978 treatment with calculated nRMSEs for multiple target variables (each) and optimum selection based on the average nRMSE (AVG-nRMSE) over multiple target variables (Memic et al.2021, under review).

LFMAX	**nRMSE**				**AVG**	
(G -Growth**)**	**GWAD**	**LAID**	**CWAD**	**LWAD**	**nRMSE**	
0.8	0.208	0.22	0.185	0.203	0.204	
0.912	0.131	0.153	0.119	0.146	0.137	
1.024	**0.082**	0.145	0.109	0.144	**0.12**	**0.12**
1.136	0.078	0.171	0.137	0.173	0.14	
1.248	0.109	0.209	0.178	0.211	0.177	

LFMAX – CROPGRO coeff., *GWAD* - grain weight, *LAID* - leaf area index, *CWAD* - tops weight, *LWAD* - leaf weight

A.2 The nRMSE multiple-treatment based goodness of fit criteria (nRMSE)

Cultivar coefficients can be optimised based on the in-season observation of one experiment (single treatment/experiment) or multiple experiment data sets (multiple season and/or locations). Using only one experimental data set for optimisation will lead to better statistical agreement between simulated and observed, but potentially result in over-fitting of simulation outputs with observations. Cultivar coefficients should perform well across multiple seasons and locations and as such should be derived based on multiple season and locations. In Table 2 are shown single treatment "optimums" and multiple treatment "optimum".

For demonstrating single treatment and multiple treatment based cultivar coefficient optimisation one cultivar coefficient (LFMAX) and three treatments have been selected

(Gainesville 1978, Gainesville 1979 and Gainesville 1984). Cultivar coefficient value minimum (0.85), maximum (1.25) and increment step (0.1) were written in the corresponding cultivar file and crop model was executed. For each cultivar coefficient simulated and observed GWAD, LAID, CWAD and LWAD average nRMSE was calculated for localising "optimums". First single treatment "optimums" are localised based on the lowest average nRMSE (AVG-nRMSE, Table 2, section a, grey fields).

In second step multiple treatment cultivar coefficient "optimum" is localised based on the single treatment based AVG-nRMSEs with formula: $[(TRT_1+TRT_2+TRT_3)/3]$. Based on the averaging of the single treatment based AVG-nRMSEs multiple treatment based optimum is selected (Table 2, section b, grey fields) with lowest treatment based average nRMSE (Table 2, section b, blue field).

Table 2 'Bragg' cultivar (*DSSAT Default*) is shown for soybean experiments conducted in Gainesville 1978, Gainesville 1979 and Gainesville 1984, with LFMAX cultivar coefficient variations from 0.85 to 1.25 with increment step 0.1 with 5 coefficient combinations for each treatment.

a) Single treatment "optimums"				b) **Multi treatment based "optimum"**					
Year	TRT	LFMAX	AVG nRMSE	Year	TRT	LFMAX	AVG nRMSE	Multiple treatment average	Optimum
1978	1	0.85	0.17	1978	1	0.85	0.17		
1978	1	0.95	0.128	1979	2	0.85	0.119		
1978	1	1.05	0.122	1984	3	0.85	0.408	0.232	
1978	1	1.15	0.142	1978	1	**0.95**	0.128		
1978	1	1.25	0.175	1979	2	**0.95**	0.14		
1979	2	0.85	0.119	1984	3	**0.95**	0.374	**0.214**	**0.214**
1979	2	0.95	0.14	1978	1	1.05	0.122		
1979	2	1.05	0.177	1979	2	1.05	0.177		
1979	2	1.15	0.21	1984	3	1.05	0.355	0.218	
1979	2	1.25	0.239	1978	1	1.15	0.142		
1984	3	0.85	0.408	1979	2	1.15	0.21		
1984	3	0.95	0.374	1984	3	1.15	0.346	0.233	
1984	3	1.05	0.355	1978	1	1.25	0.175		
1984	3	1.15	0.346	1979	2	1.25	0.239		
1984	3	1.25	0.345	1984	3	1.25	0.345	0.253	

LFMAX-defined in Table 3,*TRT*-treatment, *AVG-nRMSE*-average of normalised RMSE over four target variables (grain weight, leaf area index, Tops weight, Leaf weight)

A.3 The Relative Weight –nRMSE (RW-nRMSE)

A *Relative Weight -nRMSE* (*RW-nRMSE*) complementary method was developed with relative weights assigned to the selected target variables (1 to 5). A relative weight can be assigned to each target variable separately, where 5 is the most accurate and 1 the least accurate, in relative terms. Each accuracy weight has weight coefficient assigned to it in the TSE code. With relative weights each target variable's accuracy can be selected, giving users more flexibility in terms of target variable prioritisation in the cultivar coefficients estimation process. The relative weights multiplier (rw) in Eq. 4 is used to artificially inflate nRMSE of specific target variables. By doing so the size of specific target variable is artificially increased in the overall AVG-nRMSE that results at the end in cultivar coefficient combination prioritising those variables over others.

$$AVG_{nRMSE(n)} = (rw_1 * GWAD_{nRMSE(n)} + rw_2 * LAID_{nRMSE(n)} + rw_3 * CWAD_{nRMSE(n)})/3 \quad (4)$$

It will be additionally explained!

Table 3

A.4 The Rank Weighting –nRMSE (Rank-nRMSE)

A *Relative Weight -nRMSE* (*RW-nRMSE*) complementary method was developed with priority ranks (1 to 5). A rank can be assigned to each target variable separately, where 1 is the most accurate and 5 the least accurate, in relative terms. Each accuracy rank has weight coefficient assigned to it in the TSE code (Table 4), rank 1 with weight coefficient of 0.01 and rank 4 with values 0.04 used in selection criteria. With ranks each target variable's accuracy can be selected, giving users more flexibility in terms of target variable prioritisation in the cultivar coefficients estimation process (Table 4), as shown in the following example with gwad having rank 1 and the other target variables rank 4.

Table 4 RW-nRMSE method target variable ranking with weight coefficients

RW-nRMSE		
	Rank	**Coefficient weight**
LAID	4	0.04
CWAD	4	0.04
LWAD	4	0.04
GWAD	**1**	**0.01**

Initially the weight coefficients are defined (by ranking) for creating a mathematical threshold used in the program (Figure A.1a). The initial selection criteria was set to 0.0 (Figure A.1a). With the initial selection criteria the *while* loop (flow controlling statement executed repeatedly until specified conditions are meet in the programming language) is initialised (Figure A.1b). Within the *while* loop, the *if* condition is set for finding an overall statistically acceptable solution (Figure A.1c). The *while* loop will be stopped only if the calculated nRMSEs for multiple target variables satisfies the RW-nRMSE selection criteria.

If accurate grain yield is the goal (other target variables less accurate) of the optimisation then the relative weight coefficients based on RW-nRMSE criteria (Table 4) gwad rank 1 and the other three variables rank 4 offers more accurate results of gwad target variable. The rank 1 for gwad and rank 4 for laid, cwad and lwad initializes RW-nRMSE mathematical threshold in to code for while loop first run of 0.01 for nRMSE-gwad and 0.04 for nRMSE-laid-cwad-lwad. In practice this means when the first loop run is searching through calculated nRMSEs for all target variables only coefficient combinations that have calculated nRMSE-gwad lower than 0.01 and nRMSE-laid-cwad-lwad lower than 0.04 is selected as optimum. If in first while loop run none of the coefficient combinations have calculated nRMSE as defined with relative weights the while loop is restarted for the second time with new mathematical threshold for nRMSE-gwad lower than 0.02 and nRMSE-laid-cwad-lwad lower than 0.8. If the newly formed mathematical threshold doesn't provide coefficient combination with calculated nRMSEs the while loop will be started again with nRMSE-gwad mathematical threshold of 0.03 etc. until coefficient combination providing calculated nRMSEs for target variables is located.

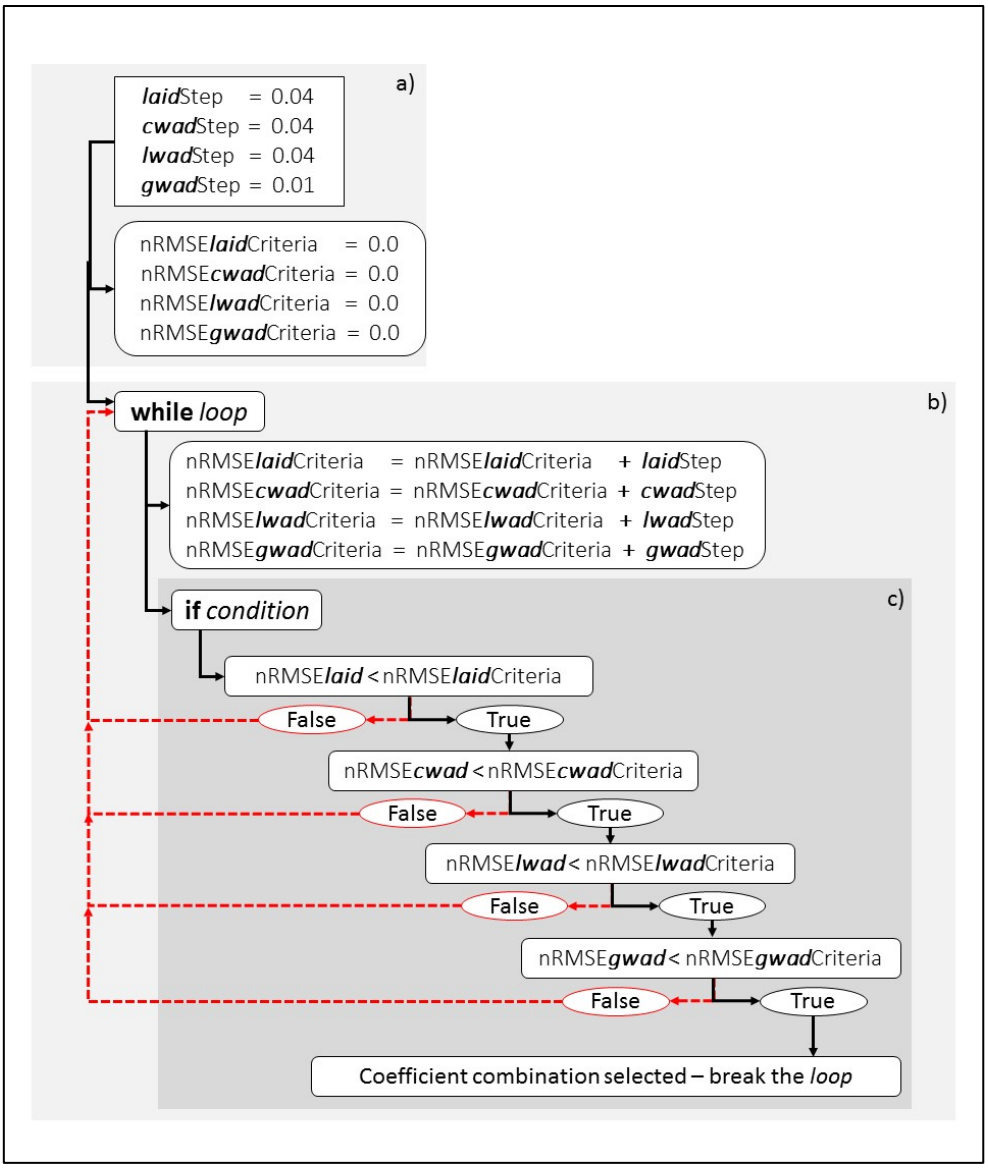


Figure A.1 The RW-nRMSE cultivar coefficient combination selection method

With RW-nRMSE-gwad rank 1 and RW-nRMSE-laid-cwad-lwad rank 4 while loop with if conditions is repeated 8 times (Table 7) with desired LFMAX coefficient of 1.136 providing the most accurate simulations for grain weight (gwad). This coefficient values is different from the one selected with simple nRMSE criteria results shown in Table 1.

Table *5* The *RW-nRMSE* - simplified example of varying one of the cultivar parameters with calculated nRMSEs for multiple target variables demonstrating selected combinations with target variable with relative weights ranking in TSE.

	nRMSE				
LFMAX (G)	**gwad**	**laid**	**cwad**	**lwad**	**AVG-nRMSE**
0.8	0.208	0.22	0.185	0.203	0.204
0.912	0.131	0.153	0.119	0.146	0.137
1.024	0.082	0.145	0.109	0.144	**0.12**
1.136	**0.078**	0.171	0.137	0.173	0.14
1.248	0.109	0.209	0.178	0.211	0.177

G – *growth*, gwad - *grain weight*, laid - *leaf area index*, cwad - *tops weight*, lwad - *leaf weight*

A.5 Range reduction – generating coefficient combinations

(This method has to be used carefully! It is designed to provide the best statistical fit between simulated and observed. The best statistical fit providing coefficients might not be Physiologically meaningful!)

The cultivar coefficient estimation process consists of varying the values for each cultivar coefficient and comparing a statistical fit of simulated outputs with field observations in order to determine the coefficient combination providing the best agreement between simulated and observed values. Various cultivar coefficients have potentially wide ranges (minimum and maximum values difference) with many in-between values that depend on the increment step size (Inc). The so called *Exhaustive gridding - coefficient variation* (Table 6) can be used to systematically investigate coefficient ranges in search for coefficient values that provide the best statistical fit. For example P5 coefficient value for minimum 100, maximum 900 and increment step 3.3 can be passed into the cultivar file and for each coefficient value model executed. In this example as shown in Table 6 it can be seen that for a coefficient range from 100 to 900 with increment steps of 3.3 a total 243 coefficient variations are executed for the P5 coefficient with Exhaustive gridding method. In order to overcome time losses in the process of cultivar coefficients estimation based on the statistical fit (lowest nRMSE), a range reduction method (Röll et al. 2020) was implemented (Table 6). With range reduction method four global pahses are conducted in the process of estimating cultivar coefficient with smallest average nRMSE. Greater increment steps are used in the first phase for each given coefficient range with P5 coefficient having minimum 100, maximum 900 and increment step 200. Based on the lowest nRMSE, the value for coefficient is selected, i.e. 300. In the second Phase, the new coefficient ranges with a narrower increment step are executed with P5 having minimum 180, maximum 420 and increment step 60. Based on the lowest nRMSE, the new coefficient "optimum" is selected, i.e. P5=300.In the third phase new range for each coefficient is defined with P5 having minimum 270, maximum 330 and increment step 15. In the final pahse P5 minimum is 263.2 and maximum 276.7 with increment step of 3.3 is passed into the cultivar file and based on lowest nRMSE P5=263.2 is selected. Based on the range reduction approach, 48% fewer combinations are executed when compared to exhaustive gridding coefficient variations. The range reduction method as described in Table 6 is expected to retain a systematic optimum localisation approach (achievable with exhaustive gridding variation), and is expected to provides more realistic values for the coefficients when compared to the random generation of cultivar coefficients for allowed ranges. The range reduction is flexible programmed and will work with different scale coefficients. The method will be described in more detail in additional publication with corresponding programming code written in python.

Example with CERES-Wheat P5 coefficient

Exhaustive gridding		**Range reduction method**							
P5		Phase 1		Phase 2		Phase 3		Phase 4	
Min	100	100		180		270		263.2	
Max	900	900		420		330		267.7	
Inc. step	**+3.3**	**+200**		**+60**		**+15**		**+3.3**	
								263.2	→263.2
								266.5	
				180		270	→270	269.9	
	100	100		240		285		273.3	
	103.3	300	→300	300	→300	300		276.7	
	…	500		360		315			
	→263.2	700		420		330			
	…	900							
	900								
No. comb.	243	5		5		5		5	
Total	**243**	**20**							

References

Memic E., Graeff S., Boote K., Hensel O., Hoogenboom G., (2021). Cultivar coefficient estimator for the cropping system model based on time-series data – A case study for soybean. (***Under review***)