Pheno_opt3

Pepijn van Oort, 23-4-2015

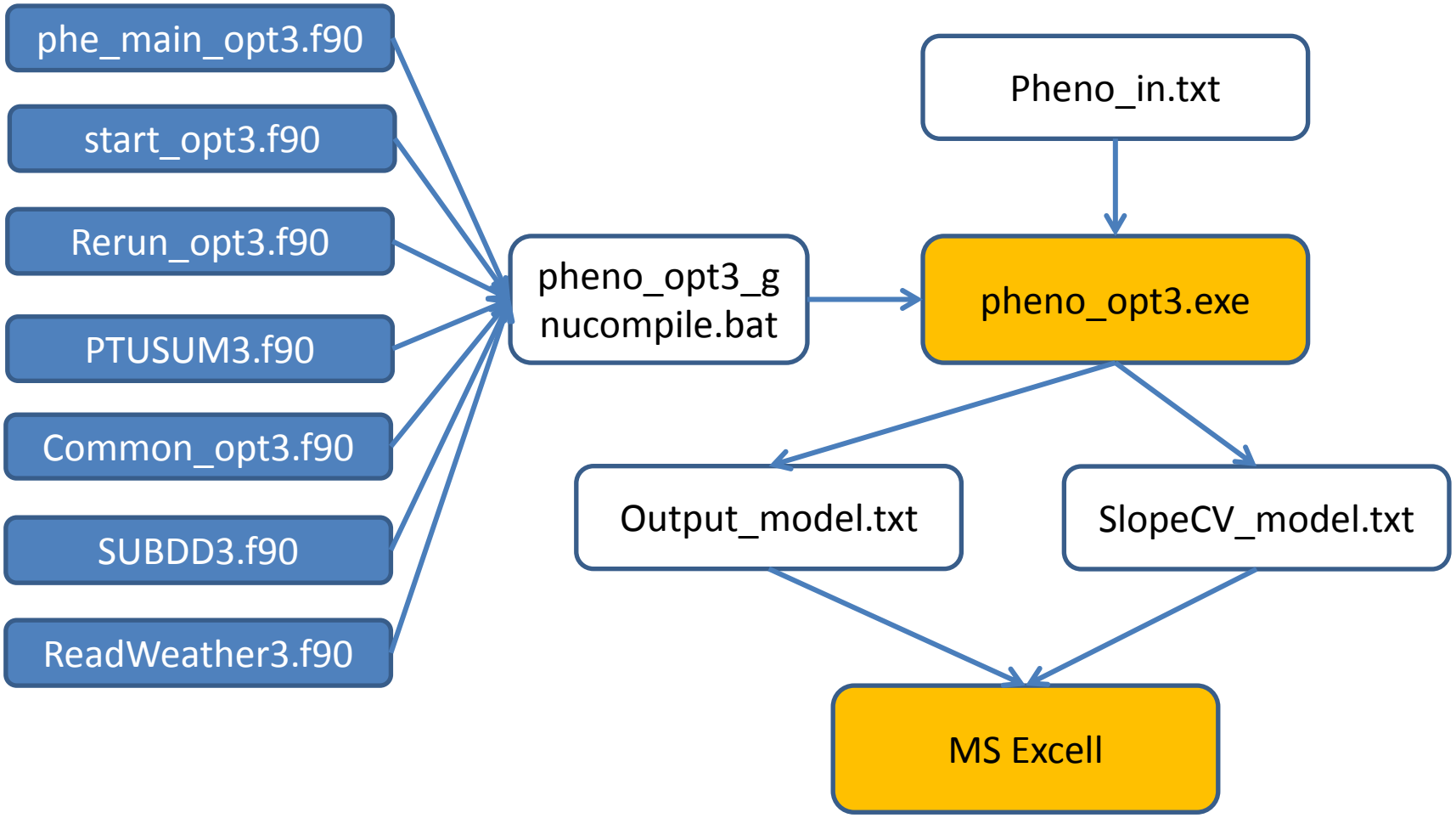
A program to estimate phenological parameters

- Input:
  - pheno_in.txt file with observed phenology, can be multi-site
  - Weather data: Tmin, Tmax (+ latitude in weather files for daylength calculation
- Output:
  - output_model.txt = observed and simulated phenological dates for each parameter set
  - SlopeCV_model.txt = summary statistics for each parameter set

Specs

- All possible parameter sets (can be very large) are evaluated. For each parameter user defines range and step size
- SinusExponential diurnal temperature model (corrects for daylength)
- Daylength and photoperiodic daylength (incl twilight, default is sun angle -4^{o} below horizon)
- 1h timestep calculation
- Different models can be chosen

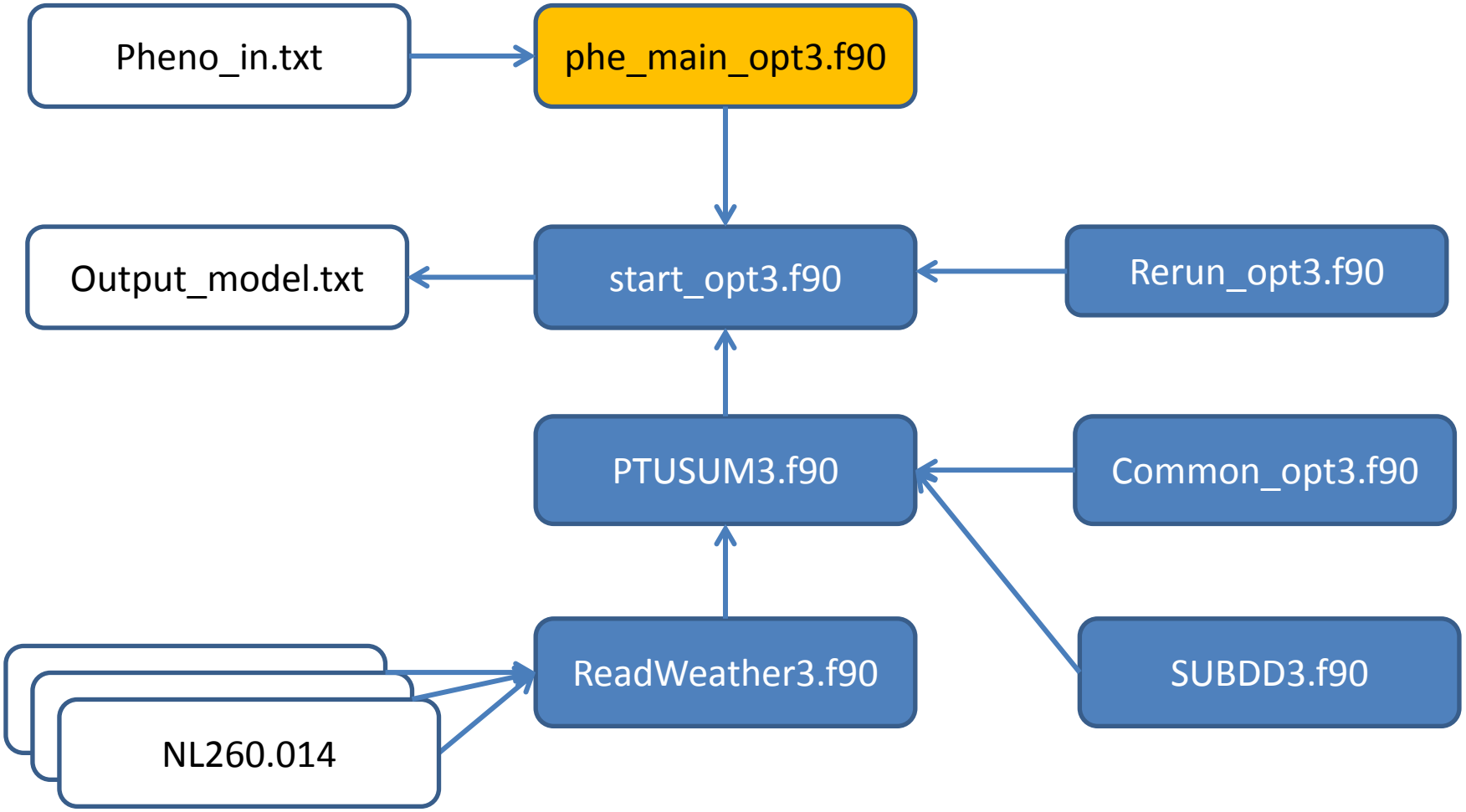
Overview (1)


phe_main_opt3.f90
start_opt3.f90
Rerun_opt3.f90
PTUSUM3.f90
Common_opt3.f90
SUBDD3.f90
ReadWeather3.f90
pheno_opt3_g
nucompile.bat
Pheno_in.txt
pheno_opt3.exe
Output_model.txt
SlopeCV_model.txt
MS Excell

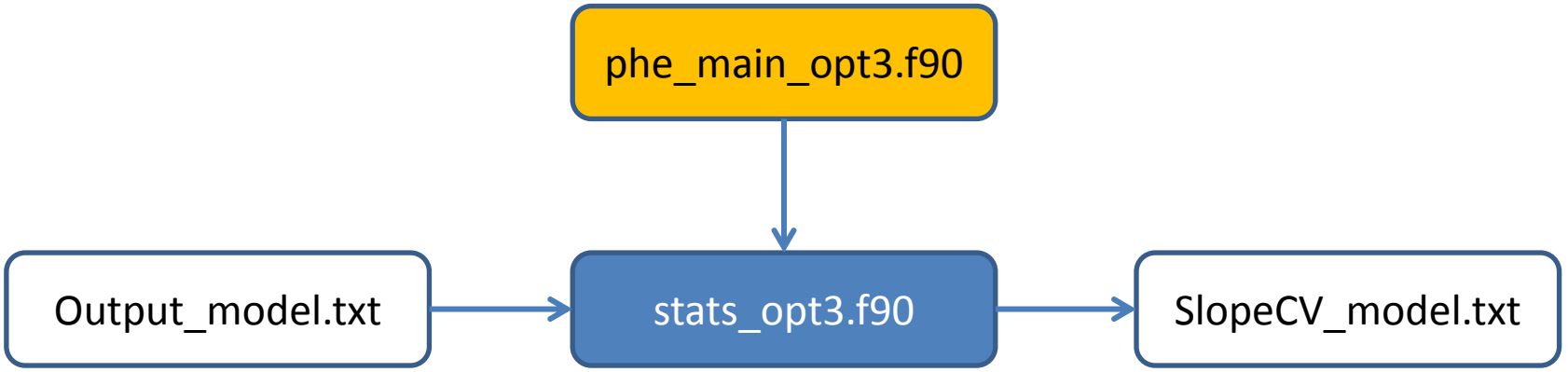
Overview (2)

- The batch file compiles the model
  - Set the correct paths in this file
  - Double click with left mouse button to compile
- The executable is the model
  - Double click with left mouse button to compile
  - Interpret the output in MS Excel. Template Excel files are available

Step 1: Calculations



Step 2: Summary statistics



Phe_main_opt3 is the main program. Here you choose your model
In **rerun_opt3** you set the parameter ranges and stepsize for the selected model
Input is in the **pheno_in.txt** file and in **weather files**

Phe_main_opt3

- first calls **start_opt3** which does all the calculations and writes detailed results to **Output_model.txt**
- next calls **stats_opt3** which reads **Output_model.txt**, calculates summary statistics and writes these to **SlopeCV_model.txt**

Choice from different models (1)

- Temperature: TBD, TOD, TMD
  - For not temperature sensitive set to 0 (never occurs in reality, but interesting for comparison)
- Photoperiod: MOPP, PPSE
  - For not photoperiod sensitive set to 0
- Transplanting shock: SHCKD
  - For non transplanted set to zero
- Combinations: see next slide

Choice from different models (2)

- Check SUBDD3.f90 for if a short or long day sensitive function is activated.
- Alternative photoperiod response models can be activated in SUBDD3.f90.
  - Default is the bilinear function, with linear delay in development above MOPP (short day plant) or delay below MOPP (long day plant)
  - Also available: inverse function
  - Also available: exponential function

Choice from different models (3)

- Bilinear0 = no development below TBD, then always increasing
- Bilinear1 = bilinear0 + plateau at TOD, above which development rate stays constant or decreases
- Bilinear2 = bilinear1 + different TOD for day and for night
- Bilinear3 = bilinear1 + add fixed temperature
- Beta = beta shape temperature & daylength function

Bilinear models

Model[1]	phe_main_opt3	Rerun_opt3	
		Calibrated parameters	Set to fixed value (=not used)
T (TBD)[2]	Bilinear0	TBD	TOD, TMD, MOPP, PPSE
T (TBD, TOD, TMD) [3]	Bilinear1	TBD, TOD, $(TMD)^3$	MOPP, PPSE, $(TMD)^3$
DL	Bilinear1	MOPP, PPSE	TBD, TOD, TMD. Set TBD =99. And TOD must not be set to 99.
T (TBD) & DL	Bilinear0	TBD, MOPP, PPSE	TOD, TMD
T (TBD, TOD, TMD) & DL	Bilinear1	TBD, TOD, MOPP, PPSE, $(TMD)^3$	$(TMD)^3$

[1] T = temperature dependent model, DL = daylength model;
[2] this is the simplest model, in which development rate increases linearly above TBD and never reaches a plateau
[3] TMD can optionally be fixed or calibrated. To keep optimal development rate (= not increasing/decreasing above TOD) set TMD = 1000

Things you need to do only once

- Prepare the phenology input file pheno_in.txt
  - See detailed instructions further on
- Prepare the weather files and place them in a directory ...\weather\
  - See detailed instructions further one
- Set the path of the weather directory in ReadWeather3.f90 (this needs to be done only once)
- Set the paths of the fortran files and the executable in pheno_opt3_gnucompile.bat (this needs to be done only once)

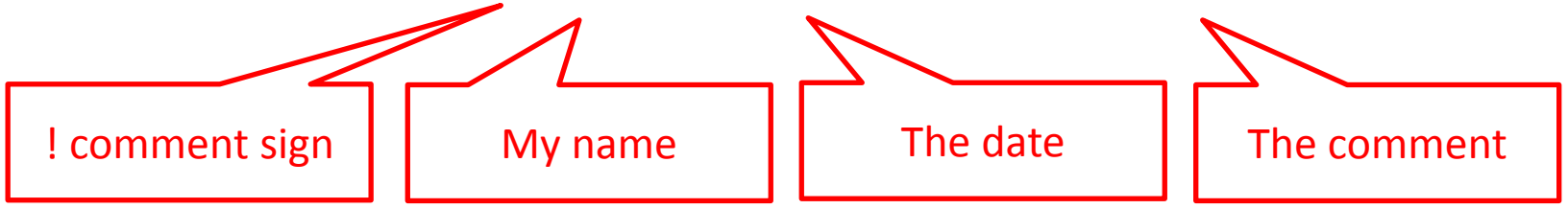
For each new calibration

- Select model in phe_main_opt3.f90
- Set parameter ranges in rerun_opt3.f90
- Recompile model with the batch file
- Interpret results, adjust parameter range if necessary

Versioning

We always make errors (even me!). So

- Don’t just start changing the files
- Make backups often
- Have a document in which you keep track of your major changes and their back-ups
- Inside computer code and text files, add comments on what you changed, why
  - I often put ! Pvo20130912: changed from ... To ...



Prepare phenology data file (1) pheno_in.txt

D:\africarice\pheno_opt_rice4\pheno_opt_rice4mvs2005\Pheno_in.txt - Notepad++

File Edit Search View Encoding Language Settings Macro Run Plugins Window ?

Rerun_opt.f90 | Pheno_in.txt

	*SITE	EMD	EMDYR	IDOYTR	IYRTR	IDOYPI	IYRPI	IDOYFL	IYRFL	IDOYM	IYRM
2	BEN1	230	2011	243	2011	277	2011	304	2011	329	2011
3	*BEN1	10	2012	28	2012	62	2012	81	2012	107	2012
4	*BEN1	233	2012	247	2012	289	2012	300	2012	333	2012
5	SEN6	61	2006	76	2006	130	2006	154	2006	184	2006
6	SEN6	79	2006	96	2006	138	2006	164	2006	198	2006
7	SEN6	108	2006	125	2006	173	2006	199	2006	222	2006
8	SEN6	138	2006	159	2006	202	2006	225	2006	252	2006
9	SEN6	169	2006	188	2006	229	2006	262	2006	285	2006
10	SEN6	199	2006	222	2006	263	2006	287	2006	317	2006
11	SEN6	230	2006	249	2006	289	2006	318	2006	354	2006
12	SEN6	291	2006	310	2006	36	2007	64	2007	96	2007
13	SEN6	322	2006	340	2006	60	2007	90	2007	124	2007
14	SEN6	352	2006	13	2007	84	2007	111	2007	141	2007
15	SEN6	18	2007	37	2007	102	2007	127	2007	156	2007
16	SEN6	49	2007	74	2007	129	2007	154	2007	184	2007
17	SEN6	77	2007	103	2007	152	2007	177	2007	207	2007
18	SEN6	108	2007	124	2007	181	2007	207	2007	239	2007
19	SEN5	19	2006	69	2006	126	2006	162	2006	183	2006
20	SEN5	77	2006	97	2006	159	2006	185	2006	204	2006
21	SEN5	108	2006	125	2006	182	2006	208	2006	236	2006
22	SEN5	138	2006	159	2006	205	2006	232	2006	256	2006

Use * for header line and comment line (data not used)

SEN5 is name of weather file.
Weather data for year 1981 are saved as SEN5.981.
Year 2001 > SEN5.001
You can add data from multiple stations

Prepare phenology data file (2)

- Do not use data from experiments in which stress occurred
- You can prepare in MS Excel, then save as **SPACE** delimited text file. For example as .PRN file.
- Open file pheno_in.txt
- Header of pheno_int.txt :

*SITE = name of weather file. Do not use spaces or other symbols in file name.

EMD = Emergence day. Julian day of year. Example: 31 jan = 31; 1 feb = 32; 31 dec is 365 or 366

EMDYEAR = Emergence year (Julian calendar). Number of 4 digits, example 1990

IDOYTR = transplanting day. In case of direct seeding calculate as EMD+1 and set parameter SHCKD = 0

IYRTR = transplanting year

IDOYPI = panicle initiation day. If missing you can assume flowering day - 30

IYRPI = panicle initiation year

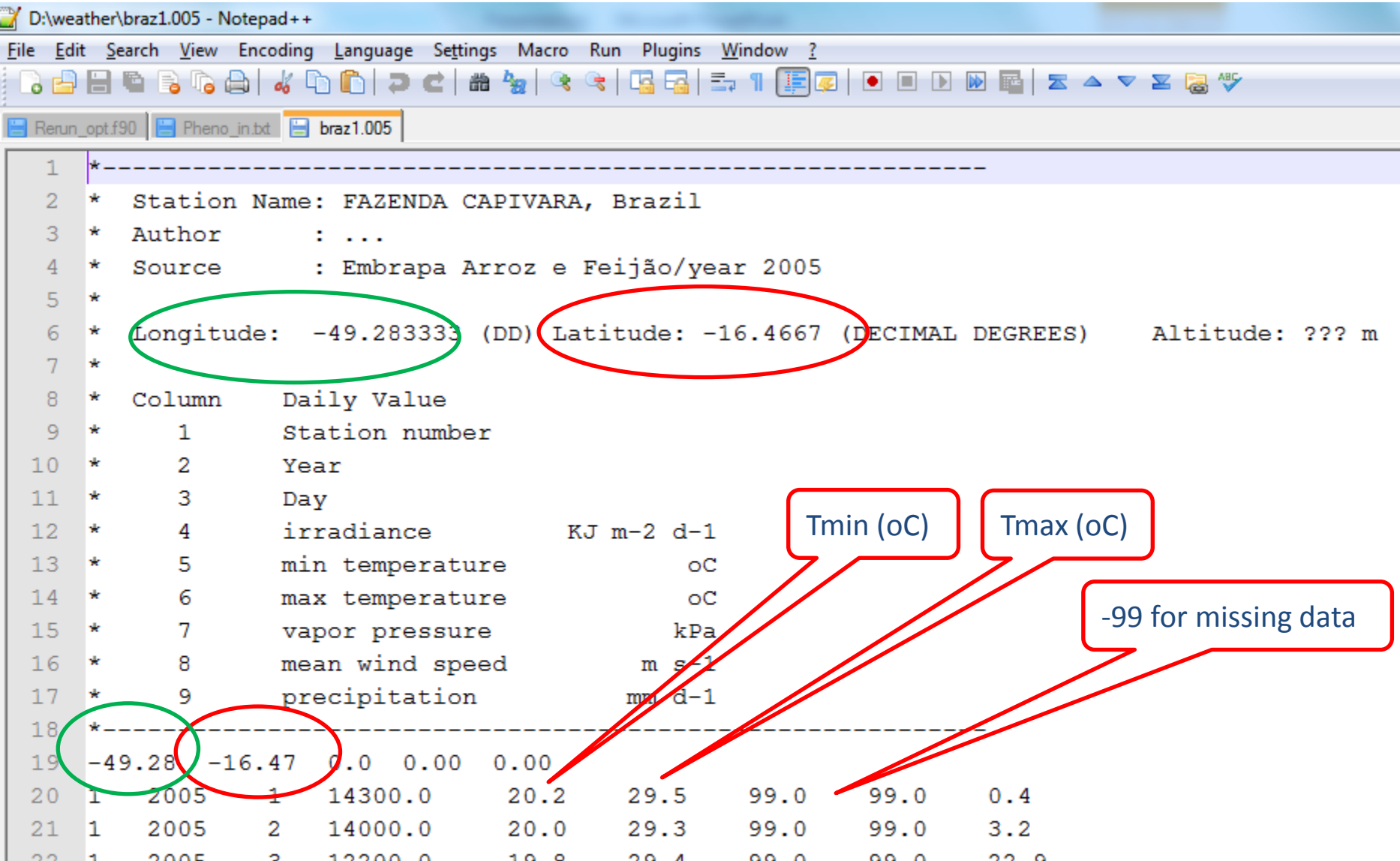
IDOYFL = flowering day

IYRFL = flowering year

IDOYM = maturity day. If missing you can assume flowering day + 35

IYRM = maturity year

Prepare weather data files (1)

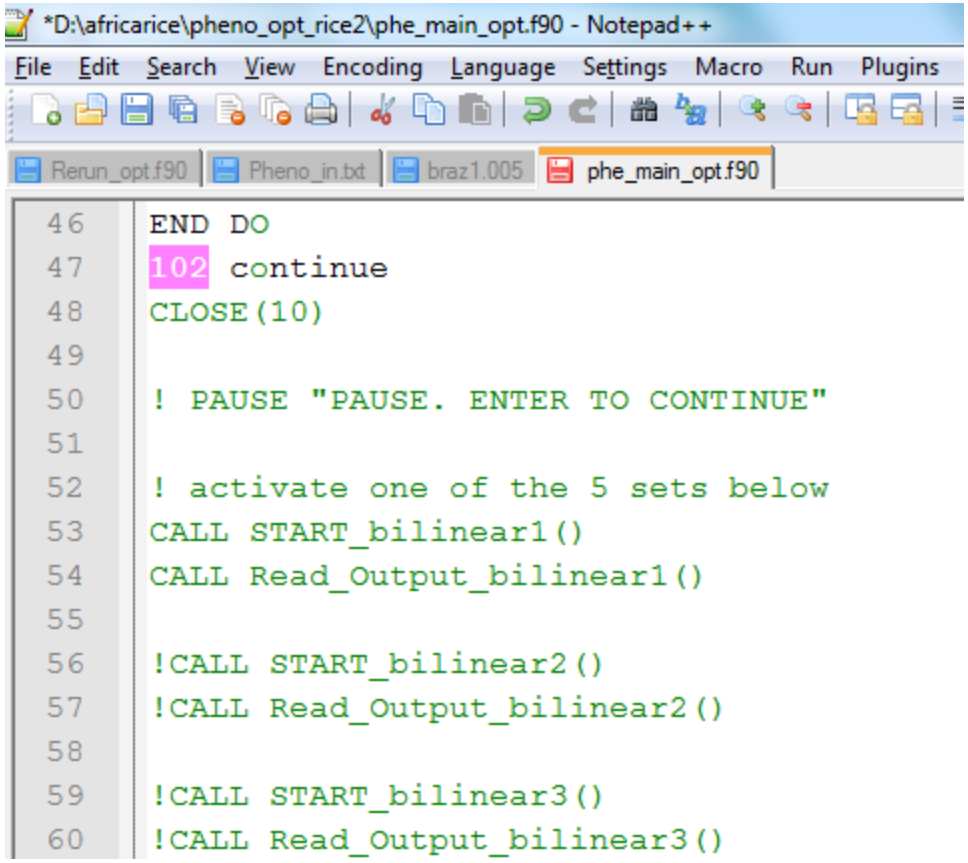


Prepare weather data files (2)

- You can prepare in MS Excel, then save as comma separated (csv) file or space delimited (prn) file.
- Pheno_opt uses Tmin, Tmax and latitude. If for other data no values just fill in -99
- Check quality of weather data
- Every year separate weather file.
- Copy directory ...\training\weather\ to D:\weather\ on your computer
- Open file ...\model**ReadWeather3.f90.** Fill in path in which weather data are stored. Example:

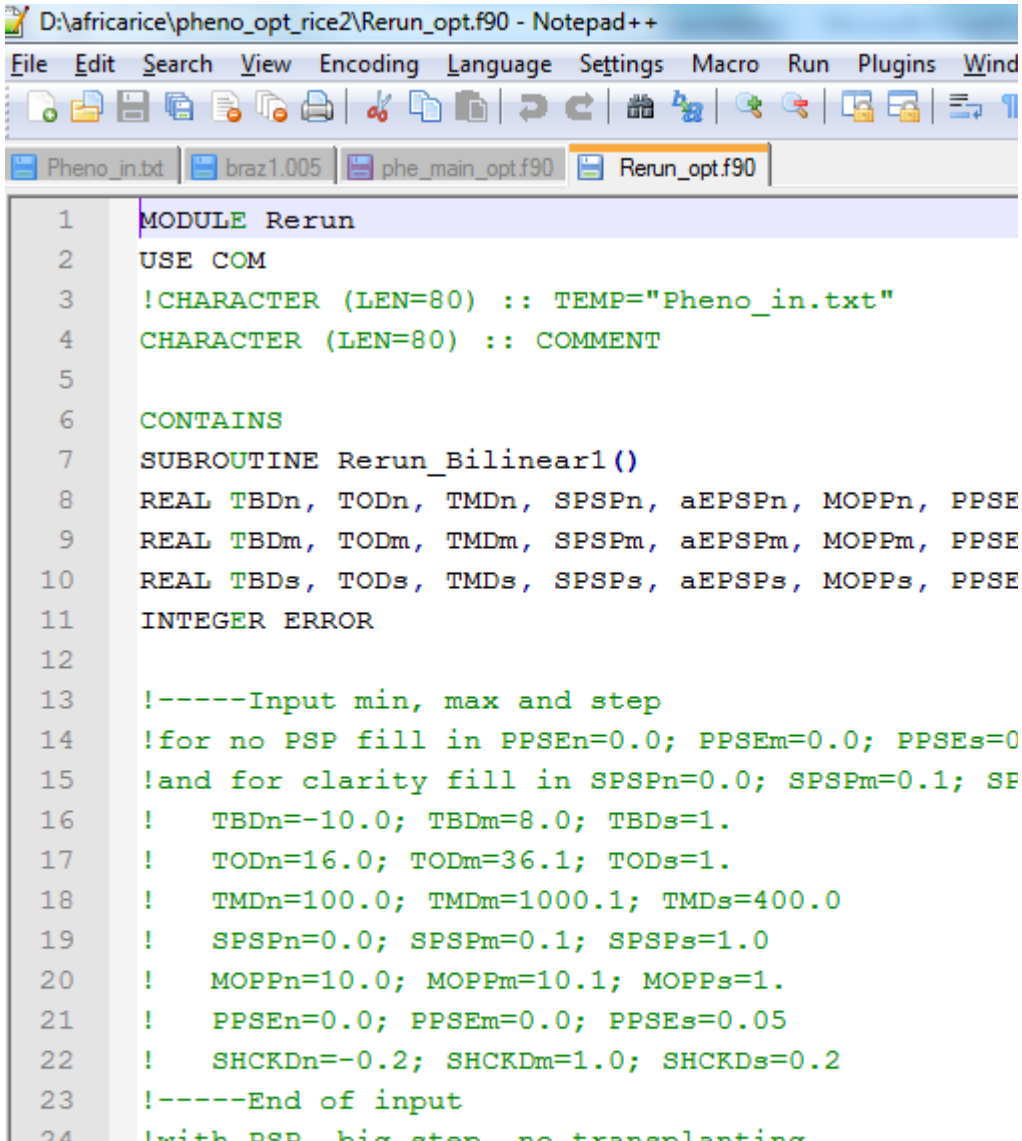
```
CHARACTER (LEN=100) :: WeaPath = "D:\weather\“
```

Select phenology model(s)



- Open file **phe_main_opt3.f90**
- Exclamation mark ! Means model is NOT run. In this example we are running model bilinear1 and not bilinear2
- You can run multiple models at same time

Set parameter range (1)



- Open file **rerun_opt3.f90**
- Go to subroutine of your model. In this case: **Rerun_Bilinear1()**
- Exclamation mark ! Means parameter set is not used.
- Lowest set is used

Set parameter range (2)

```
!no PSP, big step, no transplanting
    TBDn=0.0; TBDm=20.0; TBDs=4.
    TODn=22.0; TODm=40.1; TODs=4.
    TMDn=42.0; TMDm=1000.1; TMDs=479.0
    SPSPn=30.0; SPSPm=70.1; SPSPs=10.0
    MOPPn=11.0; MOPPm=11.1; MOPPs=2.
    PPSEn=0.0; PPSEm=0.01; PPSEs=0.4
    SHCKDn=0.; SHCKDm=0.01; SHCKDs=0.4
!-----End of input
! Short example run
    TBDn=0.0; TBDm=16.1; TBDs=8.
    TODn=30.0; TODm=30.1; TODs=1.
    TMDn=100.0; TMDm=900.1; TMDs=400.0
    SPSPn=0.0; SPSPm=0.1; SPSPs=1.0
    MOPPn=10.0; MOPPm=10.1; MOPPs=1.
    PPSEn=0.0; PPSEm=0.0; PPSEs=0.05
    SHCKDn=0.0; SHCKDm=0.81; SHCKDs=0.4


ALLOCATE (TBD_set( INT((TBDm-TBDn)/TBDs)+1 ))
ALLOCATE (TOD_set( INT((TODm-TODn)/TODs)+1 ))
ALLOCATE (TMD_set( INT((TMDm-TMDn)/TMDs)+1 ))
ALLOCATE (SPSP_set( INT((SPSPm-SPSPn)/SPSPs)+1
ALLOCATE (MOPP_set( INT((MOPPm-MOPPn)/MOPPs)+1
ALLOCATE (PPSE_set( INT((PPSEm-PPSEn)/PPSEs)+1
ALLOCATE (SHCKD_set( INT((SHCKDm-SHCKDn)/SHCKDs
```

- TBDn = minimim value for parameter TBD
- TBDm = maximum value for TBD
- TBDs = stepsize

For example with
TBDn=0.0; TBDm=16.1; TBDs=8.

We simulate
TBD = 0,8,16oC

Set parameter range (3)

- TBDn = minimim value for parameter TBD
- TBDm = maximum value for TBD
- TBDs = stepsize

For example with

TBDn=0.0; TBDm=16.1; TBDs=8. we simulate TBD = 0,8,16oC

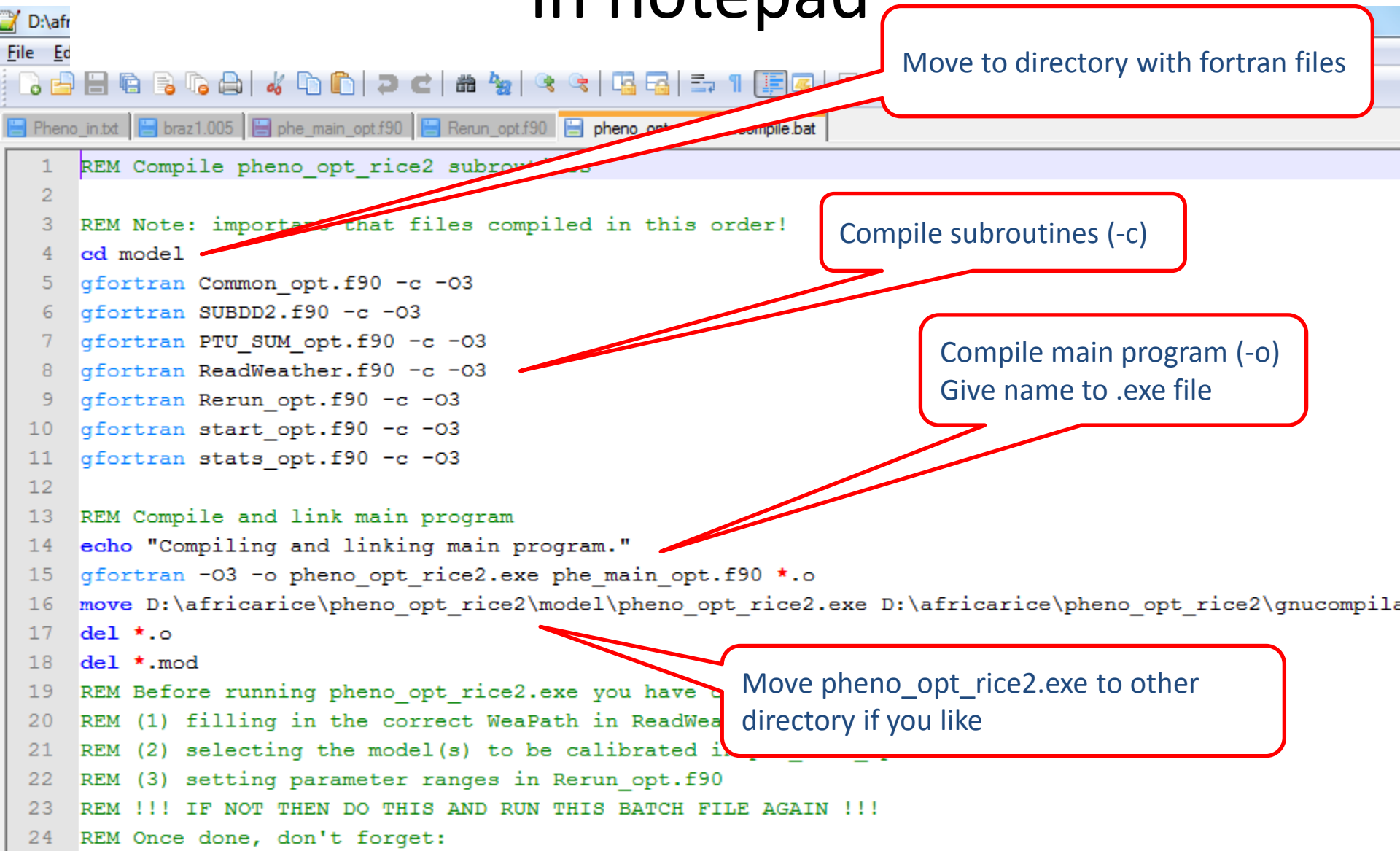
TODn=25.0; TODm=40.1; TODs=5. we simulate TOD = 25,30,35,40oC

Total simulations becomes: 3*4 = 12

Pheno_opt3 simulates ALL parameter combinations

So number of simulations can become very large !

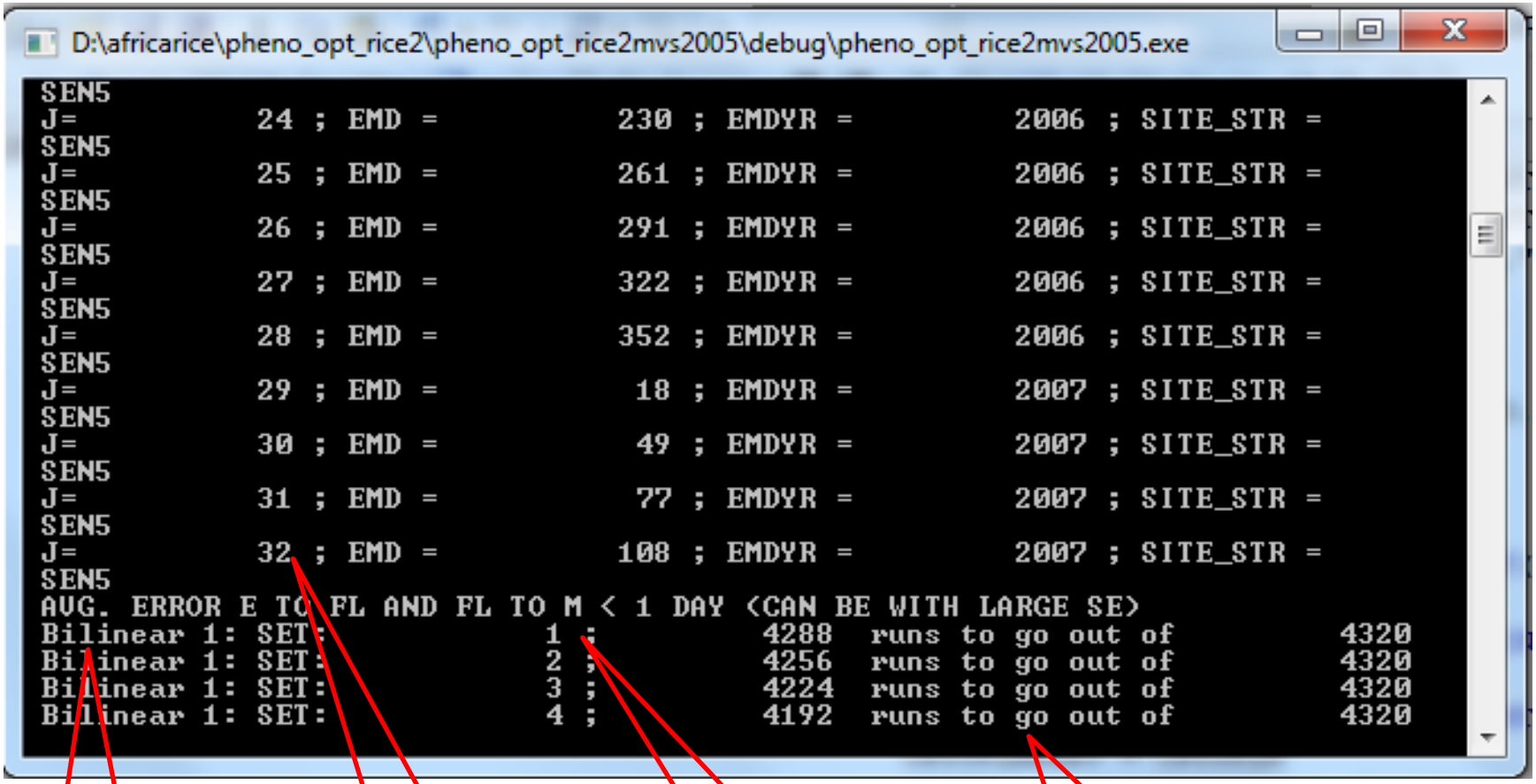
Open file pheno_opt3_gnucompile.bat in notepad



Compile & run (1)

- Double click with left mouse button on pheno_opt3_gnucompile.bat
- Check if pheno_opt3.exe was created
- Place file pheno_in.txt in same directory as pheno_opt3.exe
- Create backup of your pheno_in.txt and give it a logical name, example pheno_in_IR64COTSEN.txt
- Double click pheno_opt3.exe

Compile & run (2) – DOS window



Model

32 observations used

Counter for parameter set

How much more runs to go

Compile & run (3)

- Pheno_opt3 simulates ALL parameter combinations
- So number of simulations can become very large !
- Start with large step size, then zoom in to promising parameter range (refine)

Output files (1)

- SlopeCV_bilinear1.txt = summary statistics and development rates (DVR) for input in ORYZA2000. One line for each parameter set
- output_Bilinear1.txt = observed and simulated values for each parameter set
- Suggested to rename these files to a name that makes sense to you, for back-up

Output files (2)

- Paste pheno_in file to Excell file for reference
- Paste parameter range to Excell file for reference
- Open SlopeCV_bilinear1.txt in notepad
- Paste to Excel file, for example to sheet BILINEAR1_SLOPECV_NOPSP in file **pheno_opt_rice2training1.xlsx**
- **Sort by RMSE_EM (small to large)**

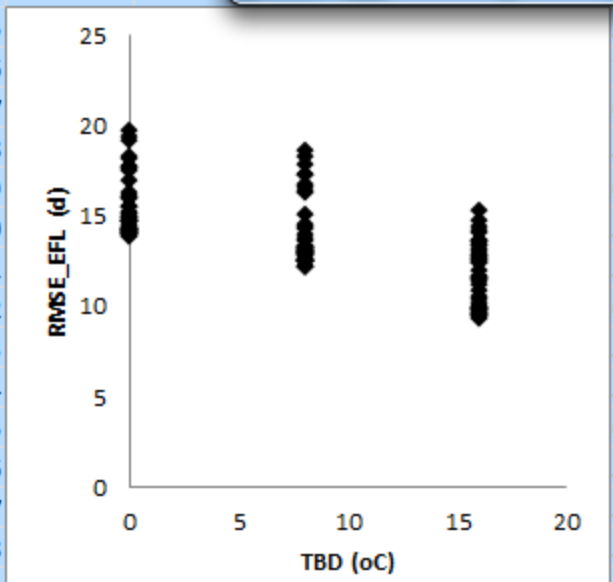
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	1	8	30	42	0	47	10	0	0.4	65	20	-4.437	14.137	0.222	4.
2	*SET	TBD	TOD	TMD	SPSP	EPSP	MOPP	PPSE	SHCKD	CPTUFL	CPTUFLM	SLOPEEPI	RMSE_EPI	SLOPEPIFL	RMSE
3	1	0											17.054	-0.538	4.
4	2	0											17.386	-0.481	4.
5	3	0											17.788	-0.458	4.
6	4	0											15.795	-0.342	4.
7	5	0											16.224	-0.172	3.
8	6	0											16.603	-0.274	3.
9	7	0											15.627	-0.284	3.
10	8	0											15.99	-0.242	4.
11	9	0											16.49	-0.223	3.
12	10	0											14.662	-0.125	3.
13	11	0											15.013	-0.179	3.
14	12	0											15.709	-0.154	3.
15	13	0											13.976	0.009	3.
16	14	0											14.377	-0.017	3.
17	15														
18	16														
19	17														
20	18														
21	19														
22	20														
23	21														
24	22														
25	23														
26	24														
27	25														
28	26														
29	27														
30	28														
31	29														

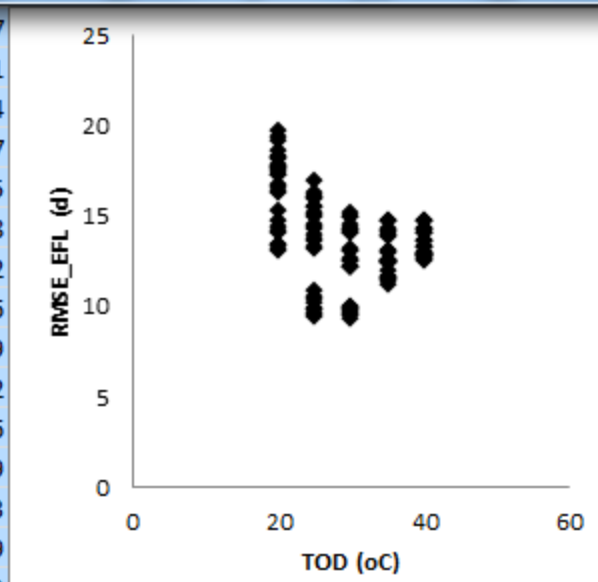
Sort

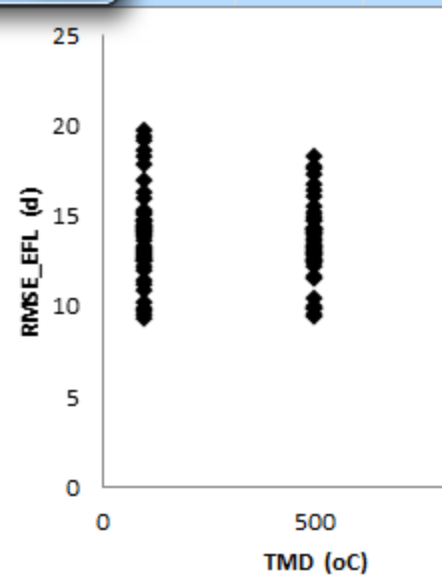
Add Level | Delete Level | Copy Level | Options... | My data has headers

Column	Sort On	Order
Sort by RMSE_EM	Values	Smallest to Largest

OK | Cancel







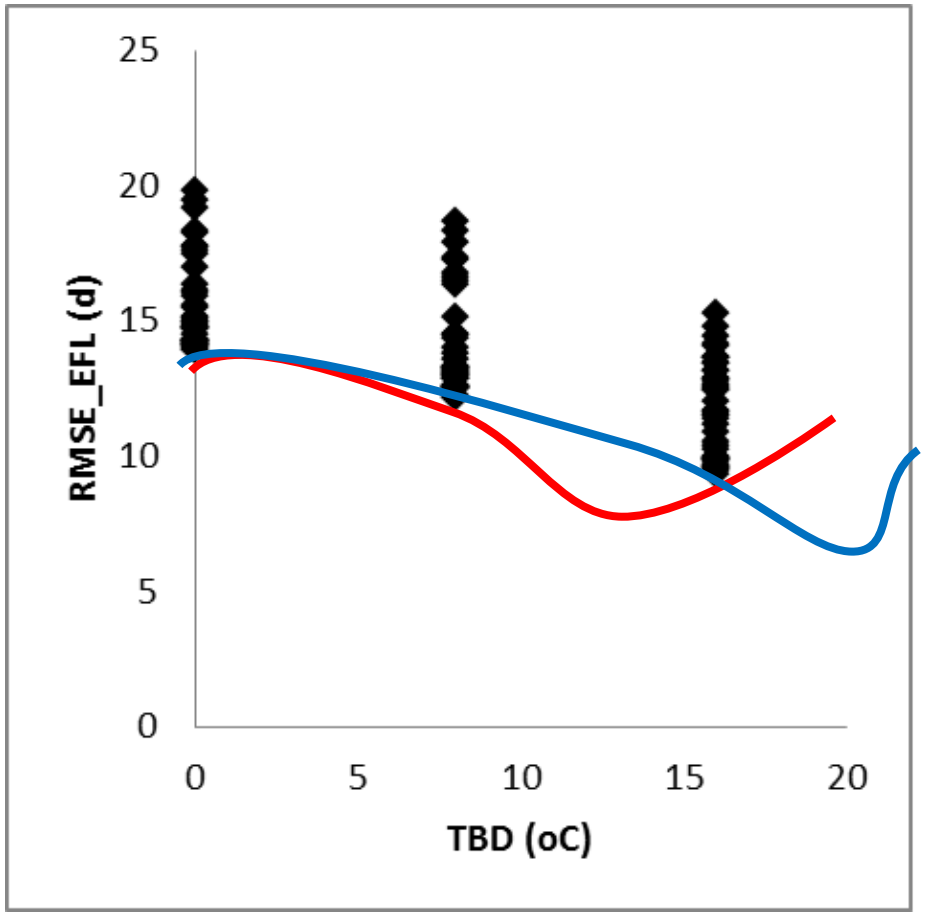
BILINEAR1_SLOPECV | BILINEAR1_OUTPUT | pheno_in

Interpretation (1)

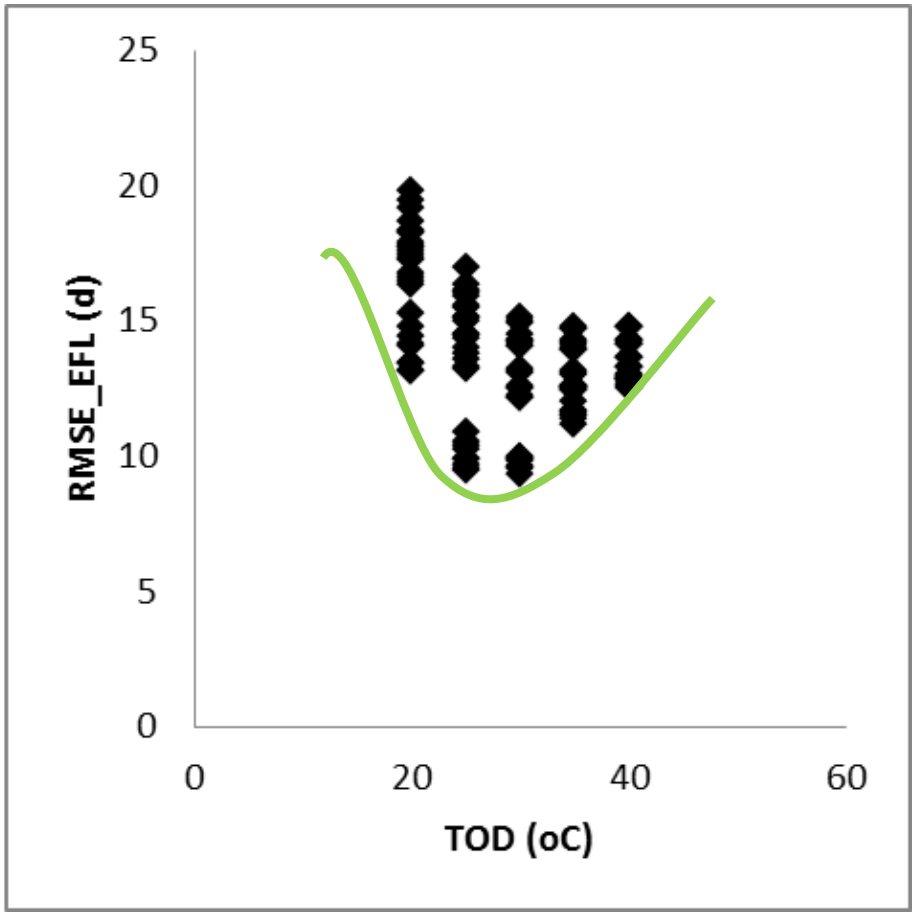
$$RMSE_{EFL} = \sqrt{\sum_{j=1}^{EXPNUM} (IEFL - IEFLSIM)^2/(EXPNUM)}$$

- In example **pheno_opt_rice2training1.xlsx** best fit was parameter *SET = 70
- Best parameter set: TBD = 16oC, TOD = 30oC, TMD = 1000oC, SHCKD = 0., EPSP = 55d,CPTUFL=76d, CPTUFLM = 22d (calibration assuming no photoperiod sensititivity
- Best root mean square error for emergence to flowering (EFL): RMSE_EFL = 6.095 days
- Best root mean square error for flowering to maturity (FLM): RMSE_FLM = 4.988 days
- Best root mean square error for emergence to maturity (EM): RMSE_EM = 6.565 days

Interpretation (2)

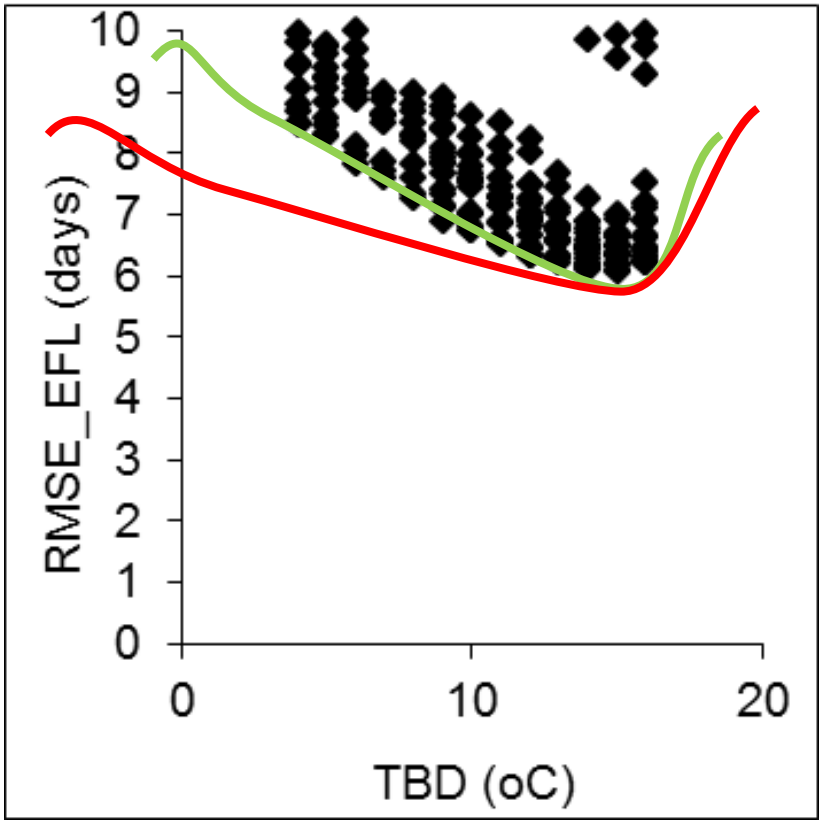


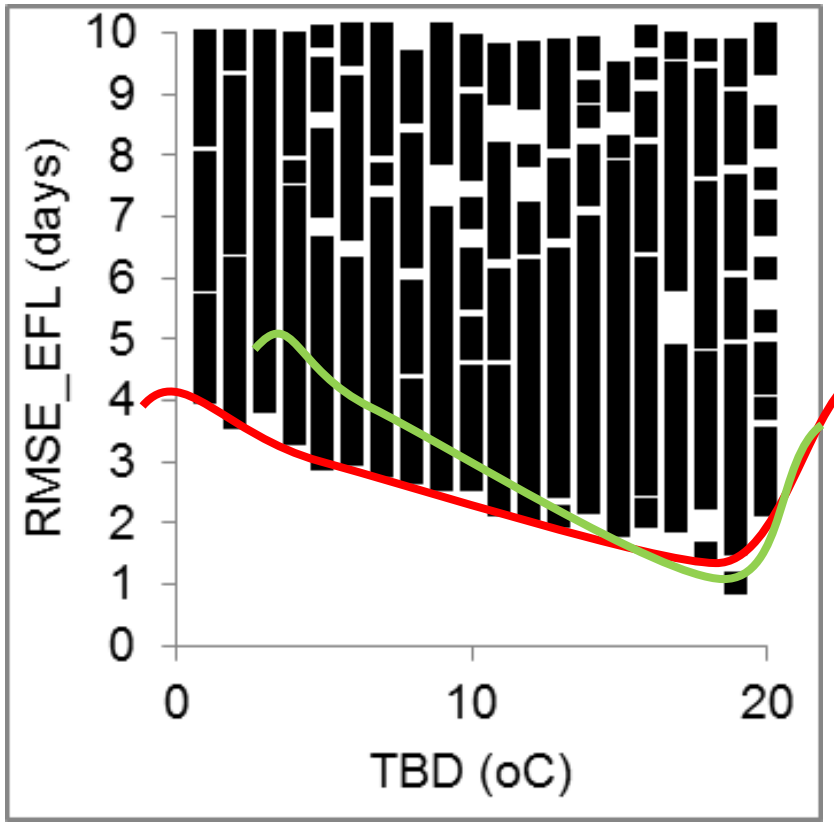
Optimum can be between 8 and 16 oC (red) or > 16oC (blue)



Good: optimum parameter between minimum and maximum

Interpretation (3)





Green has stronger convergence.

Many parameters give good fit -> enough experiments?

Output files (3)

- Open Output_bilinear1.txt in notepad
- Look up parameter set 109. Paste to Excel file, for example to sheet BILINEAR1_OUTPUT_NOPSP in file **pheno_opt_rice2training1.xlsx**
- From this file you can calculate R2 and maximum errors (instead of RMSE, which was already provided)

Output files (4) **Exercises**

1. Open file **pheno_opt_rice2training1.xlsx**
2. With the default ORYZA2000 parameters, is the error in the period from emergence to flowering correlated with average temperature during that period?
3. Calculate RMSE_EFL from sheet BILINEAR1_OUTPUT_NOPSP and compare with value shown in BILINEAR1_SLOPECV_NOPSP
4. Calculate SLOPEEFL_TMEFL from sheet BILINEAR1_OUTPUT_NOPSP and compare with value shown in BILINEAR1_SLOPECV_NOPSP
5. Compare the two sites, is simulated duration always too short or too long in one of the sites?

Refine (1)

- A good strategy is to first do a short run with large step size, interpret results, then set up new runs to refine
- If you find strange results, probably there is errors in the input files (weather & phenology). Check again!

Refine (2)

- General rule: optimum parameter must always be in middle of parameter range
- In example **pheno_opt_rice2training1.xlsx** best fit was with TBD = 16oC
- This is in the middle of the range (8,16,24) which we tried. Next step could be to refine around 16, for example with 8 to 24 oC stepsize 1oC
- If best fit had been at 24oC then next step would be TBDn = 16, TBDm=32, TBDs = 8 (different range, same large step)
- In this example we assumed not photoperiod sensitive. In a next run, we could try with photoperiod sensitive:

```
!with PSP, big step, no transplanting
    SPSPn=5.0; SPSPm=50.1; SPSPs=5.0
    MOPPn=10.0; MOPPm=13.1; MOPPs=1.
    PPSEn=0.1; PPSEm=0.4; PPSEs=0.1
    SHCKDn=0.; SHCKDm=0.01; SHCKDs=0.4
```