



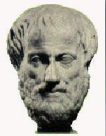
A Method for the Hierarchy of CALMO Tests

Euripides Avgoustoglou

Hellenic National Meteorological Service

18th COSMO General Meeting, Offenbach

September 6 2016



MOTIVATION

- ⊕ Towards the effort of model optimization and upon gauging the model sensitivity, when the number n of considered model parameters increases, the number of their *pair* combinations regarding their min and max values *vastly* increases [$O(2n)^2$].
- ⊕ In CALMO project , considering an extensive period of COSMO Model testing of order 1 year, the number of tests becomes of $O[10^2(2n)^2]$ and upon the accounting of *intermediate* parameter values regarding the use of a metamodel, the number of tests rises to $O[10^3(2n)^2]$ or the equivalent of runs for 5 centuries for $n=7$!!
- ⊕ An efficient methodology to constrain the number of tests should be to indicate their impact according to some quantitative criteria and decide upon the resulting *priority*.
- ⊕ The methodology is expected to be of practical value if two goals could be accomplished:
 - ◆ The tests get a *priority number* and are performed according to it.
 - ◆ If the number of tests becomes too extensive, the method should be *flexible* enough to be terminated at the priority that suits the available computational resources. The recommended truncation, however, needs to be supported by valid scientific arguments regarding the relative importance of the tests that will be included against those that will be omitted.



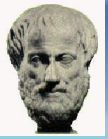
The proposed methodology will be presented in *STEPS* and *in reference* to my web presentation in the 17th COSMO General Meeting in Wrocław:

“Design and Evaluation of Sensitivity Tests for CALMO Project”

available at:

<http://www.cosmo-model.org/content/consortium/generalMeetings/general2015/parallel.htm>

and referred as **EA_COSMO_GM_2015** herein.



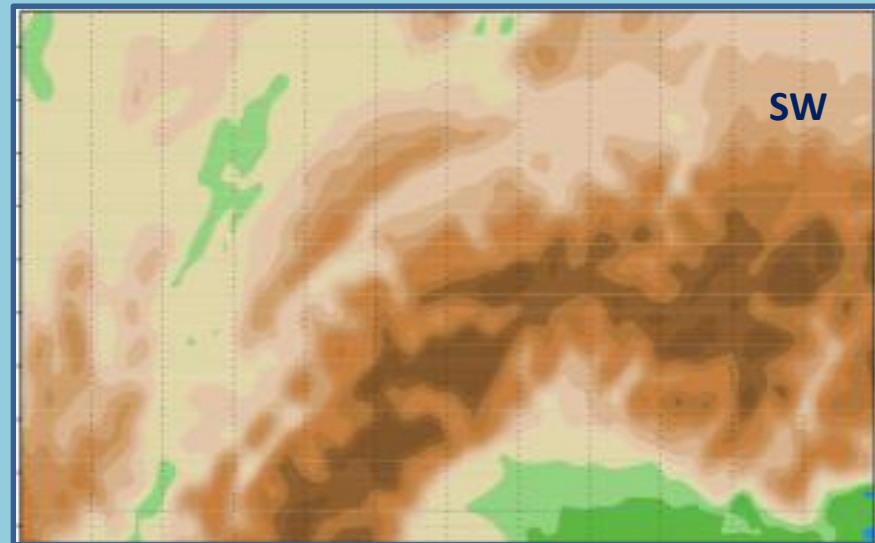
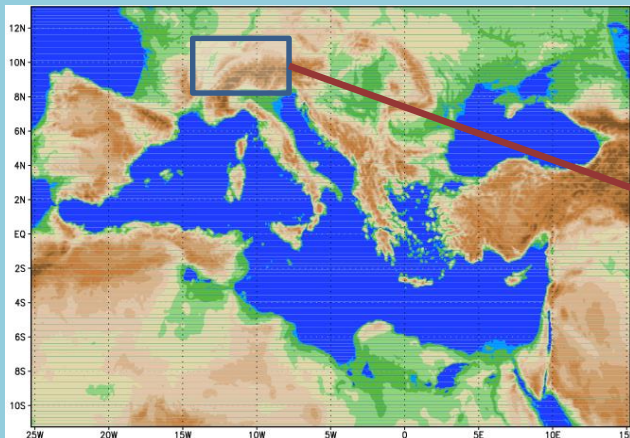
STEPS OF THE PROPOSED METHOD



STEP 1

Decide the domain where the metamodel will be used for optimization.

In the proposed application of the method, a domain that approximates the Swiss area in reference to the current COSMO.GR domain is chosen , i.e.:



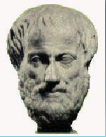


STEP 2

Choose the parameters that will be optimized by the meta model.

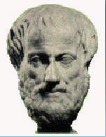
PARAMETER	INTERPRETATION	RANGE	TEST VALUES (default)
crsmin	minimum stomatal resistance	50-300	50, 150, 300
c_soil	surface area index of evaporative soil surfaces (dependent on surface area density of the roughness elements over land , c_lnd)	0-c_lnd(2.0)*	0, 1, 2
entr_sc	mean entrainment rate for shallow convection	5.0E-5 – 2.0E-3	5.0E-5, 3.0E-4, 2.0E-3
rlam_heat	scaling factor of the laminar boundary layer for heat	0.1 – 2.0	0.1, 1.0, 2.0
tkhmin tkmmin	minimal value of diffusion coefficient for heat and momentum (kept equal)	0.1-1.0	0.1, 0.4, 1.0
tur_len	asymptotic maximal turbulent length scale (m)	100 – 1000	100, 150, 1000
v0snow	factor in the terminal velocity for snow	10.0 – 30.0	10, 20, 30

* c_lnd: Surface-area index of gridpoints over land (excluding leaf-area index).



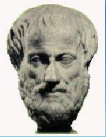
STEP 3: Create the *Priority Board Of Terms* (PBOT).

	c_soil min	c_soil max	v0_sn min	v0_sn max	crsmin min	crsmin max	entr min	entr max	tkhm min	tkhm max	tur_l min	tur_l max	rlam_h min	rlam_h max
c_soil min														
c_soil max														
v0_sn min														
v0_sn max														
crsmin min														
crsmin max														
entr min														
entr max														
tkhm min														
tkhm max														
tur_l min														
tur_l max														
rlam_h min														
rlam_h max														



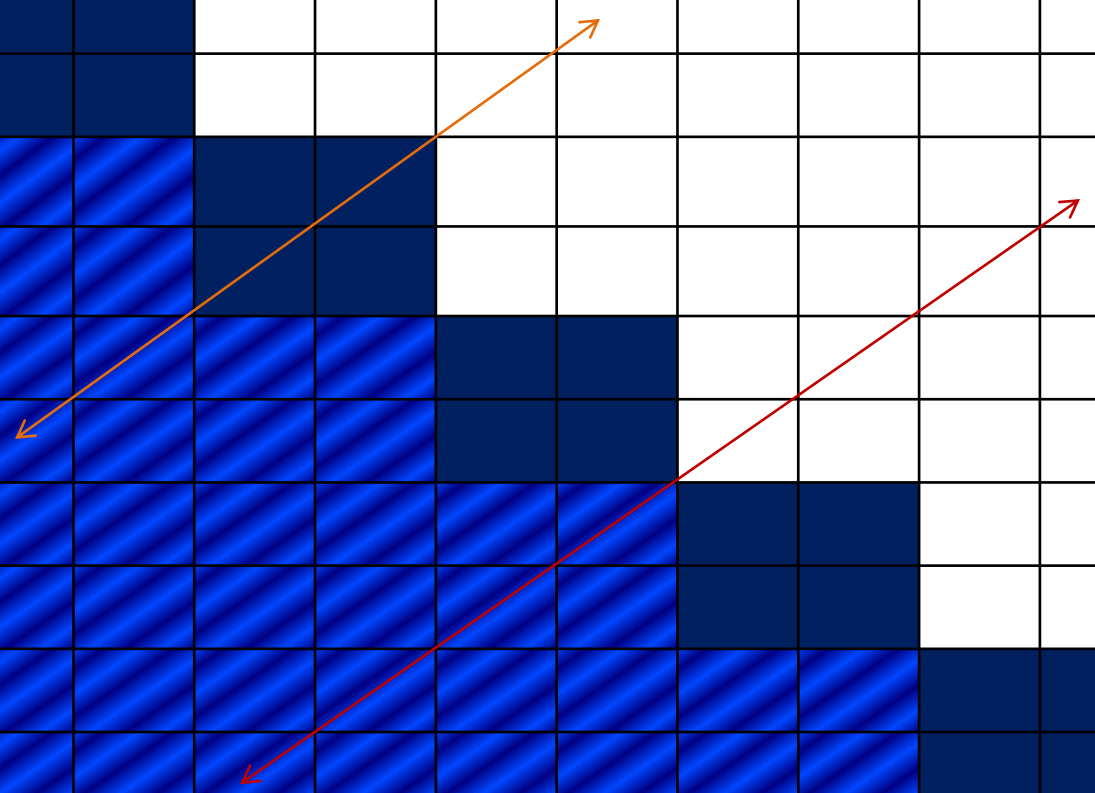
STEP 3: Create the *Priority Board Of Terms (PBOT)*.

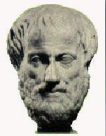
	c_soil min	c_soil max	v0_sn min	v0_sn max	crsmin min	crsmin max	entr min	entr max	tkhm min	tkhm max	tur_l min	tur_l max	rlam_h min	rlam_h max
c_soil min	Blue	Blue												
c_soil max	Blue	Blue												
v0_sn min	Blue	Blue	Blue	Blue										
v0_sn max	Blue	Blue	Blue	Blue										
crsmin min	Blue	Blue	Blue	Blue	Blue	Blue								
crsmin max	Blue	Blue	Blue	Blue	Blue	Blue								
entr min	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue						
entr max	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue						
tkhm min	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue				
tkhm max	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue				
tur_l min	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue		
tur_l max	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue		
rlam_h min	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue
rlam_h max	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue	Blue




STEP 3: Create the *Priority Board Of Terms (PBOT)*.

	c_soil_min	c_soil_max	v0_sn_min	v0_sn_max	crsmin_min	crsmin_max	entr_min	entr_max	tkhm_min	tkhm_max	tur_l_min	tur_l_max	rlam_h_min	rlam_h_max
c_soil_min														
c_soil_max														
v0_sn_min														
v0_sn_max														
crsmin_min														
crsmin_max														
entr_min														
entr_max														
tkhm_min														
tkhm_max														
tur_l_min														
tur_l_max														
rlam_h_min														
rlam_h_max														





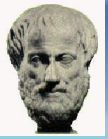
Remarks on STEP 3

- ⊕ The name of the parameters has been changed a little from the standard model documentation in order for the PBOT simply to match the page space.
- ⊕ It can easily be shown that the pair combinations for 7 parameters is 84, if we include their min and max values, i.e. $2n(n-1)$ where n is the number of model parameters.
- ⊕ Consequently, the 84 empty white cells will be filled with priority numbers 1 to 84
- ⊕ Every empty white shell refers to a 2-parameter combination.
- ⊕ The empty dark blue cells will not take any number due to [single parameter](#) assignment and due to [double counting](#) in every test.
- ⊕ Every number refers to the *priority* of the sensitivity run. For example, if the *second* empty white shell of the first line gets the number 5, the model runs with the combination $(c_soil_min, v0_snow_max)$ will have priority 5, so 4 other parameter combinations have to be performed first, i.e: 



STEP 3: Assign a priority number to PBOT.

	c_soil min	c_soil max	v0_sn min	v0_sn max	crsmin min	crsmin max	entr min	entr max	tkhm min	tkhm max	tur_l min	tur_l max	rlam_h min	rlam_h max
c_soil min				5										
c_soil max														
v0_sn min														
v0_sn max														
crsmin min														
crsmin max														
entr min														
entr max														
tkhm min														
tkhm max														
tur_l min														
tur_l max														
rlam_h min														
rlam_h max														



But how the priority numbers are going to be assigned ?

This is presented in the following steps ...



STEP 4:

- ⊕ Decide on the *model variables* that will be used and tabulate them according to their importance denoted as *class* .
- ⊕ At this stage of the work, this step is subjective but in my opinion, it can be done on meteorological arguments based on the internal knowledge of the model. The variables that are recommended below are as given in **EA_COSMO_GM_2015** but with some rearrangement according to their subjectively estimated importance.
- ⊕ The subjective criteria on making this choice are not of importance for the presentation of the method and can be discussed at a later stage.

1. < [TOTPREC](#) >: 0-24 hr period accumulated precipitation (kg m⁻²). TOTPREC
2. < [TMIN2m](#) >: Minimum 2m temperature 0-24 hr periods. TMIN2m
3. < [TMAX2m](#) >: Maximum 2m temperature for 0-24 hr periods. TMAX2m
4. < **CLCM** >: Medium cloud cover (%) average of 3hr time steps 03-24 hs.
5. < **SNOW_GSP** >: 0-24 hr periods accumulated grid-scale snow (kg m⁻²)
6. < **CLCL** >: Low cloud cover (%) average of 3hr time steps 03-24 hs.
7. < **CLCH** > : High cloud cover (%) average of 3hr time steps 03-24 hs.



STEP 5:

Define the sensitivities (S) of these parameters. This is also a matter of choice. I follow the choice made in EA_COSMO_GM_2015.

$$S_{\langle P \rangle} (\%) = \frac{\langle P \rangle_{TEST} - \langle P \rangle_{DEFAULT}}{\langle P \rangle_{DEFAULT}} \bullet 100$$

$\langle P \rangle$ stands for $\langle SNOWGSP \rangle$ or $\langle TOTPREC \rangle$ or $\langle CLCL \rangle$ or $\langle CLCM \rangle$ or $\langle CLCH \rangle$

$$S_{\begin{bmatrix} TMIN\ 2m \\ TMAX\ 2m \end{bmatrix}} = \begin{bmatrix} \langle TMIN\ 2m \rangle \\ \langle TMAX\ 2m \rangle \end{bmatrix}_{TEST} - \begin{bmatrix} \langle TMIN\ 2m \rangle \\ \langle TMAX\ 2m \rangle \end{bmatrix}_{DEFAULT}$$



STEP 6:

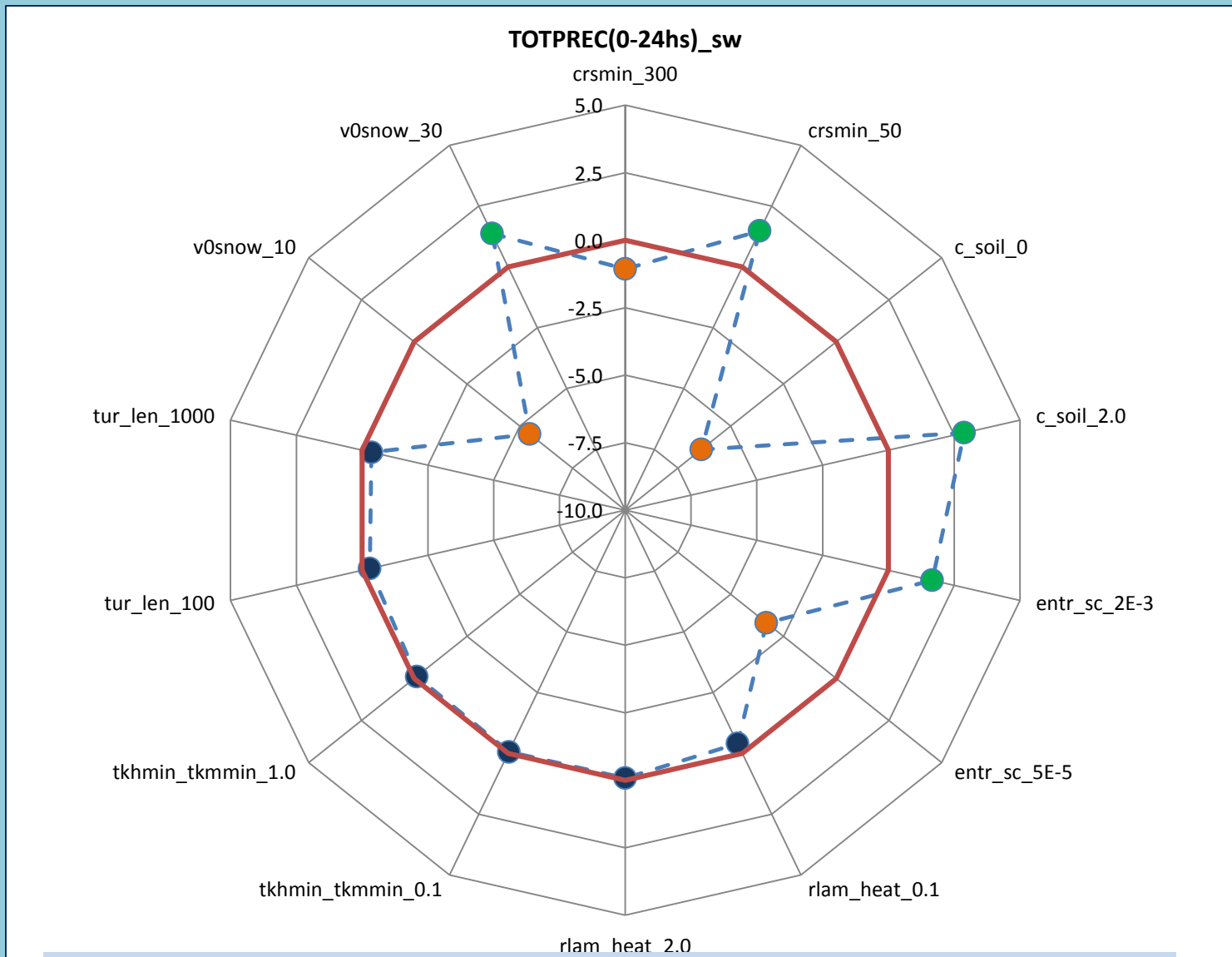
- ⊕ Perform *all* the model runs for the dates chosen for all the *default* values of the parameters chosen and produce *at least all* the considered variables mentioned before.
- ⊕ Perform *all* the model runs for the dates chosen for all the *max* values of the parameters chosen and produce *at least all* the considered variables mentioned before.
- ⊕ Perform *all* the model runs for the dates chosen for all the *min* values of all the parameters chosen and produce *at least all* the considered variables mentioned before.

In what follows, please note that for demonstration purposes, I use the results from EA_COSMO_GM_2015.



STEP 7:

Present the sensitivity of the variable of the first priority (STEP 4) in a **spider-type graph**, i.e. TOTPREC.





Analysis of STEP 7

- ⊕ The **red polygon** refers to the *zero sensitivity* “axis”.
- ⊕ The sensitivities *close to zero* are depicted with *blue bullets* ● . The parameters related to these bullets *will not* provide any priority number to the board presented in STEP3 regarding TOTPREC.
- ⊕ The *negative* sensitivities well *below* the **red polygon** are depicted with *orange bullets* ● .
- ⊕ The *positive* sensitivities well *above* the **red polygon** are depicted with *green bullets* ● .
- ⊕ The dashed polygon line that connects the dots, although not necessary in the method denotes optically the overall sensitivity for the considered meteorological variable especially to the degree that it is *convex/concave and mainly in reference to the zero sensitivity red polygon*.
- ⊕ The *first set* of priority numbers will be assigned to PBOT according to the *radial* distances between the **orange** and the **green** group of bullets and can be denoted as |orange-green| referring to the sensitivities opposite to the zero sensitivity axis.
- ⊕ The *second set* of priority numbers will be assigned to PBOT according to the *difference of the radial distances* between the *same* group of bullets and can be denoted as |orange-orange| and |green-green|. You may note that the criterion for the second set is *only the distance* and not the color, i.e. a priority number may be assigned to an |orange-orange| distance but the next priority number may be assigned to a |green-green| distance.
- ⊕ However, the priority numbers for |orange-green| distances should be *always smaller* than any same color distance. That means that |orange-green| distances will have always *higher* priorities than any *same color* distances even if their values might be smaller. To my understanding, this is recommended to keep an internal balance when the meta-model is applied.

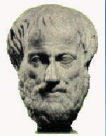


According to the previous comments PBOT becomes:



STEP 8: Update of PBOT for TOTPREC .

	c_soil min	c_soil max	v0_sn min	v0_sn max	crsmin min	crsmin max	entr min	entr max	tkhm min	tkhm max	tur_l min	tur_l max	rlam_h min	rlam_h max
c_soil min			14 	2 	1 	16 	13 	3 						
c_soil max			5 	19 	21 	10 	4 	20 						
v0_sn min					6 	18 	15 	7 						
v0_sn max					22 	11 	8 	23 						
crsmin min							9 	24 						
crsmin max							17 	12 						
entr min														
entr max														
tkhm min														
tkhm max														
tur_l min														
tur_l max														
rlam_h min														
rlam_h max														

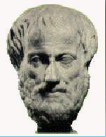


STEP 9: Update of PBOT by coloring the boxes referring to TOTPREC .

This is important to depict the *class* of the meteorological variable.

TOTPREC (class 1)

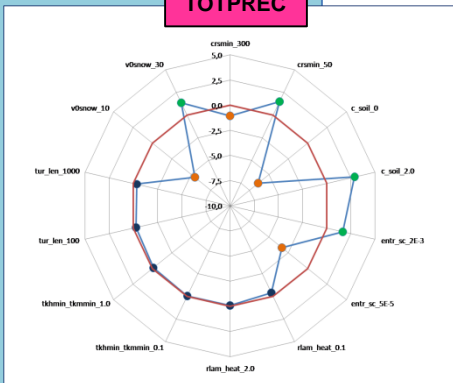
	c_soil min	c_soil max	v0_sn min	v0_sn max	crsmin min	crsmin max	entr min	entr max	tkhm min	tkhm max	tur_l min	tur_l max	rlam_h min	rlam_h max
c_soil min			14 	2 	1 	16 	13 	3 						
c_soil max			5 	19 	21 	10 	4 	20 						
v0_sn min					6 	18 	15 	7 						
v0_sn max					22 	11 	8 	23 						
crsmin min							9 	24 						
crsmin max							17 	12 						
entr min														
entr max														
tkhm min														
tkhm max														
tur_l min														
tur_l max														
rlam_h min														
rlam_h max														



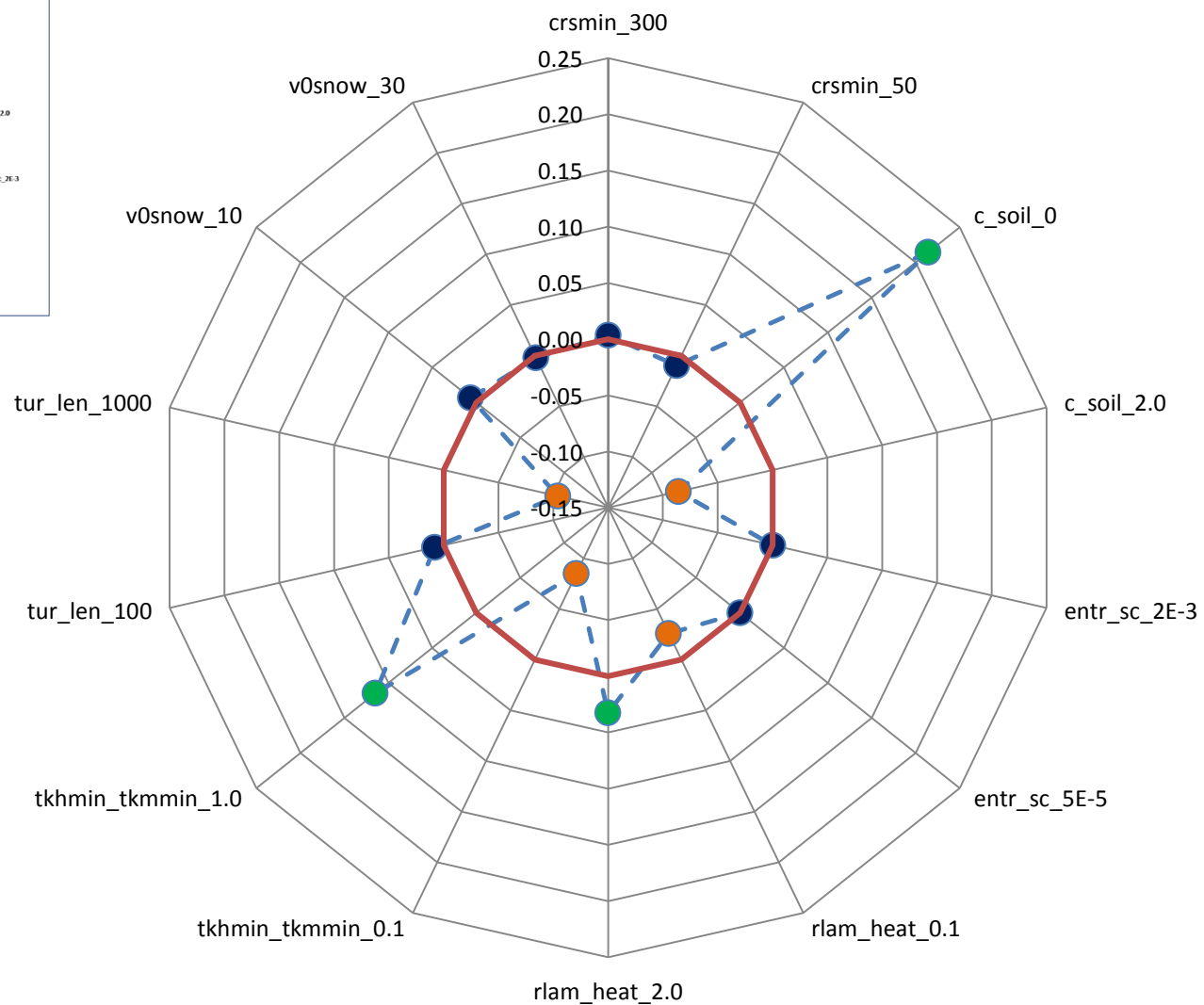
STEP 10:

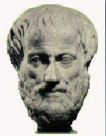
Present the sensitivity of the variable of second priority in a spider-type graph, i.e. TMIN2m in reference to the spider-graph of TOTPREC.

TOTPREC



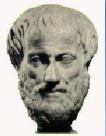
TMIN2m(0_24hs)_sw





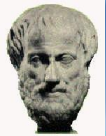
Comments on the graph of STEP 10

- ⊕ The same process is followed as in [STEP 7](#) but with an *important difference*:
- ⊕ The process is followed for all the combinations of orange and green bullets whenever *at least one* of these bullets was *blue* in the meteorological variable of [class 1](#) (TOTPREC).
- ⊕ It should be noted that the process now becomes a little cumbersome and there might be some mistakes in the following update of the PBOT but the main purpose is to focus on the methodology.
- ⊕ At this stage, the update is performed manually, but the writing of a computer code of the method is rather straightforward (but tricky !).
- ⊕ It was decided to present the method as soon as it was formulated because the task of running the experiments was really huge and it might be the case that this method could provide some help since it is of no extra computational cost but has the potential to cut or modify the cost in a scientifically consistent way.



STEP 11: Update of PBOT for TMIN2m.

	c_soil min	c_soil max	v0_sn min	v0_sn max	crsmin min	crsmin max	entr min	entr max	tkhm min	tkhm max	tur_l min	tur_l max	rlam_h min	rlam_h max
c_soil min			14	2	1	16	13	3	26	35		25	27	34
c_soil max			5	19	21	10	4	20	41	29		39	40	32
v0_sn min					6	18	15	7						
v0_sn max					22	11	8	23						
crsmin min							9	24						
crsmin max							17	12						
entr min														
entr max														
tkhm min												38	42	33
tkhm max												28	30	36
tur_l min														
tur_l max													37	31
rlam_h min														
rlam_h max														

**STEP 12: Update of PBOT by coloring the boxes referring to TMIN2m.**

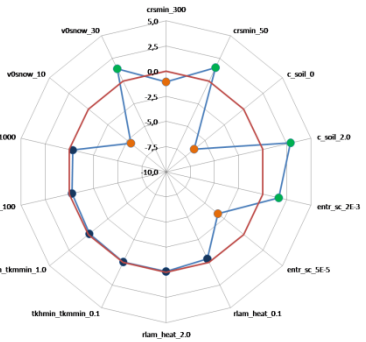
	c_soil min	c_soil max	v0_sn min	v0_sn max	crsmin min	crsmin max	entr min	entr max	tkhm min	tkhm max	tur_l min	tur_l max	rlam_h min	rlam_h max
c_soil min			14	2	1	16	13	3	26	35		25	27	34
c_soil max			5	19	21	10	4	20	41	29		39	40	32
v0_sn min					6	18	15	7						
v0_sn max					22	11	8	23						
crsmin min							9	24						
crsmin max							17	12						
entr min														
entr max														
tkhm min												38	42	33
tkhm max												28	30	36
tur_l min														
tur_l max													37	31
rlam_h min														
rlam_h max														



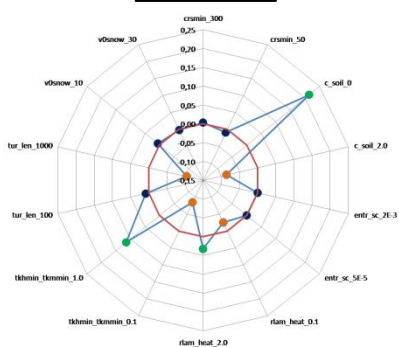
STEP 13:

Present the sensitivity of the variable of third priority in a spider-type graph, i.e. TMAX2m in reference to the spider-graph of TOTPREC and TMIN2m.

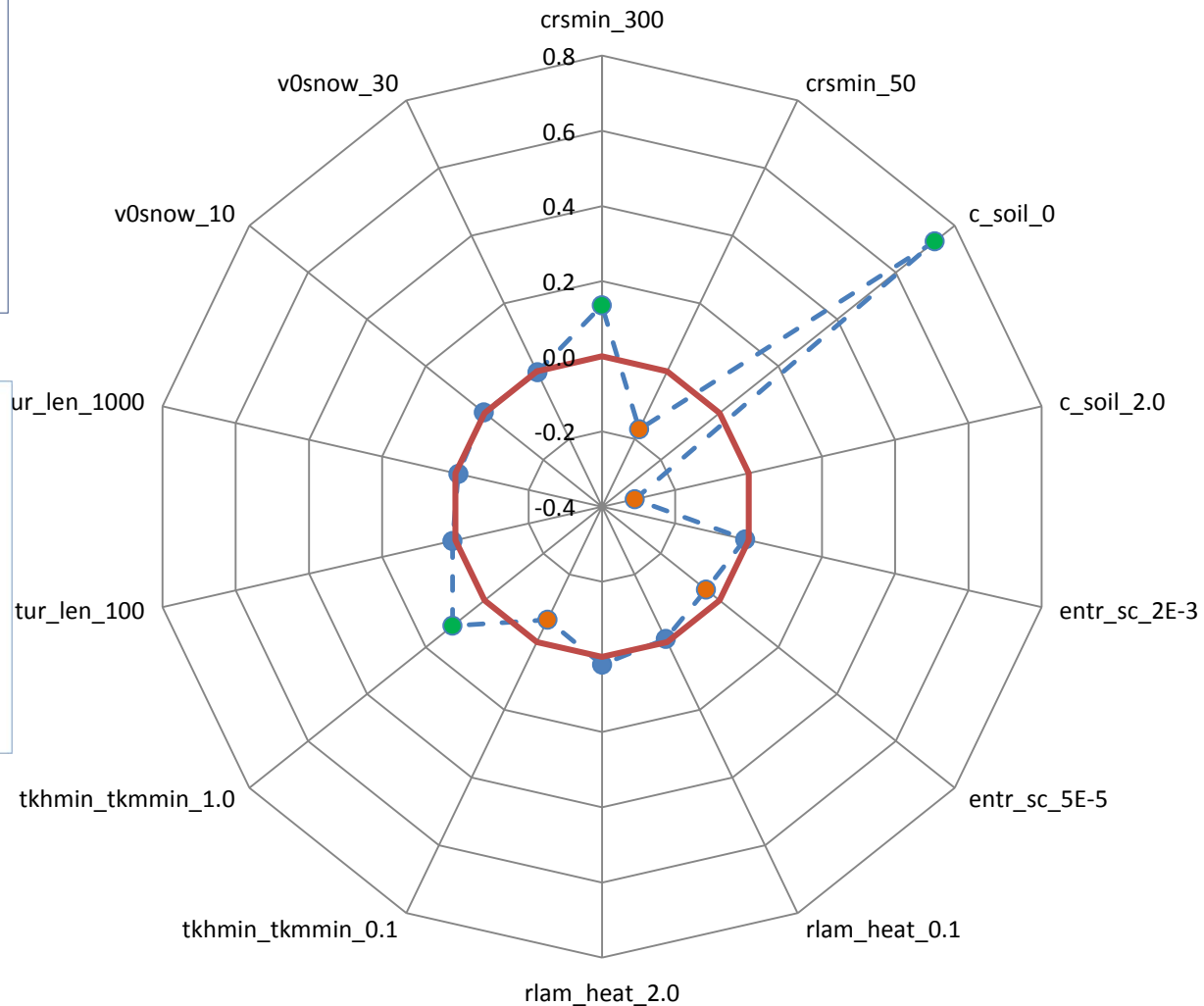
TOTPREC

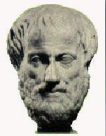


TMIN2m



TMAX2m(0-24hs)_sw





STEP 14: Update of PBOT for TMAX2m.

	c_soil min	c_soil max	v0_sn min	v0_sn max	crsmin min	crsmin max	entr min	entr max	tkhm min	tkhm max	tur_l min	tur_l max	rlam_h min	rlam_h max
c_soil min			14	2	1	16	13	3	26	35		25	27	34
c_soil max			5	19	21	10	4	20	41	29		39	40	32
v0_sn min					6	18	15	7						
v0_sn max					22	11	8	23						
crsmin min							9	24	45	44				47
crsmin max							17	12	43	46				48
entr min														
entr max														
tkhm min												38	42	33
tkhm max												28	30	36
tur_l min														
tur_l max													37	31
rlam_h min														
rlam_h max														



STEP 15: Update of PBOT by coloring the boxes referring to TMAX2m.

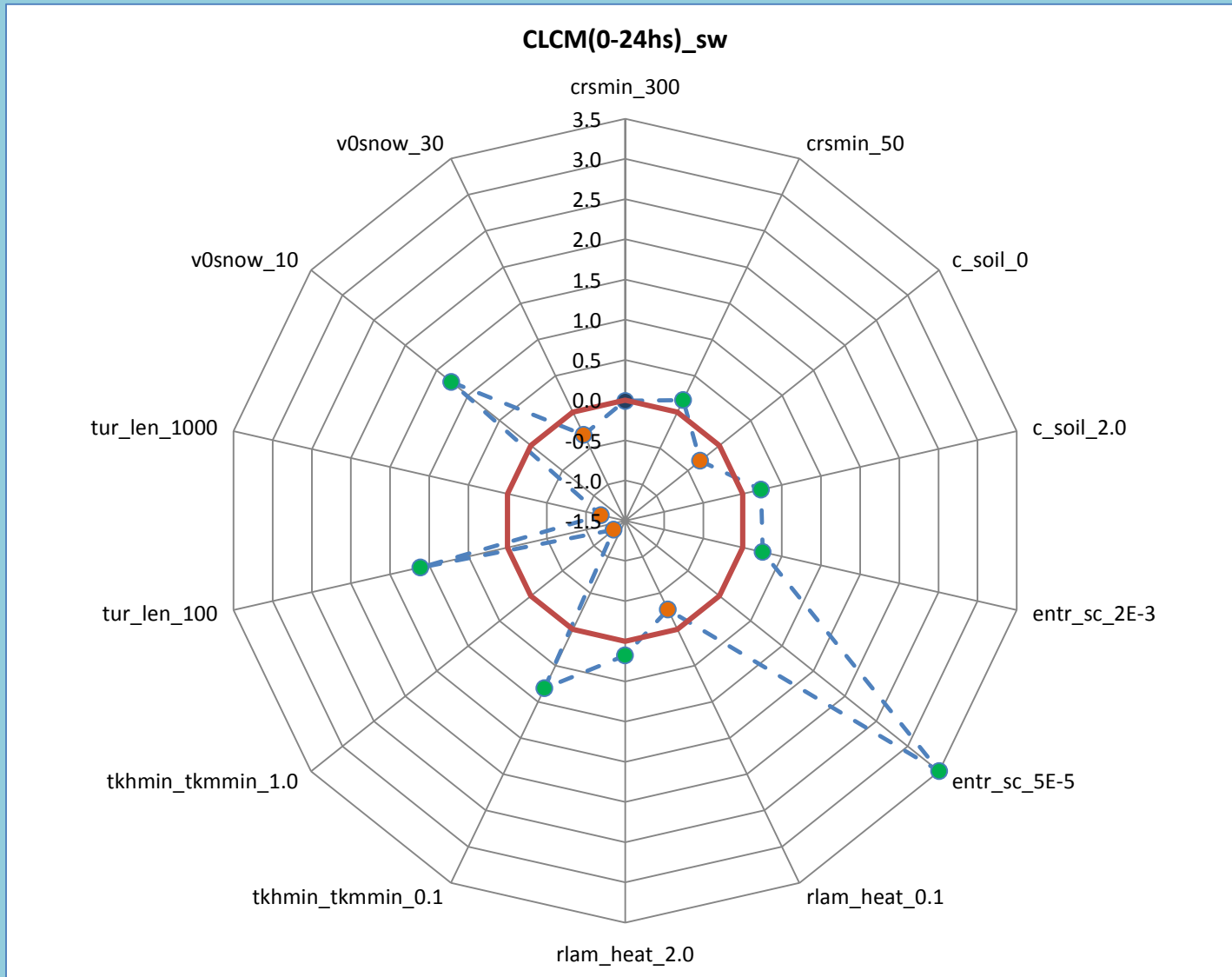
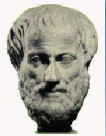
- TOTPREC class 1
- TMIN2m class 2
- TMAX2m class 3

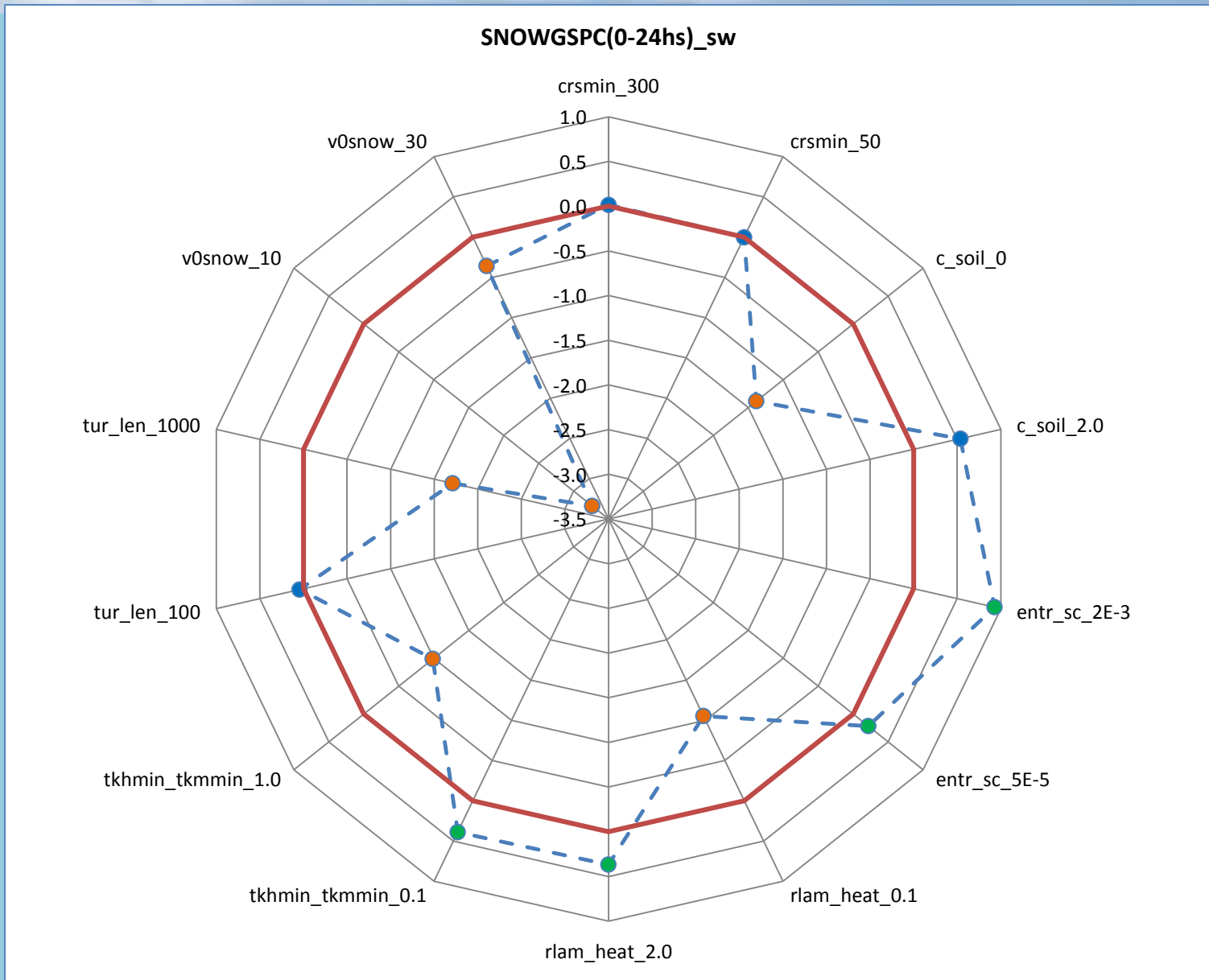
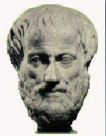
	c_soil min	c_soil max	v0_sn min	v0_sn max	crsmin min	crsmin max	entr min	entr max	tkhm min	tkhm max	tur_l min	tur_l max	rlam_h min	rlam_h max
c_soil min			14	2	1	16	13	3	26	35		25	27	34
c_soil max			5	19	21	10	4	20	41	29		39	40	32
v0_sn min					6	18	15	7						
v0_sn max					22	11	8	23						
crsmin min							9	24	45	44				47
crsmin max							17	12	43	46				48
entr min														
entr max														
tkhm min												38	42	33
tkhm max												28	30	36
tur_l min														
tur_l max													37	31
rlam_h min														
rlam_h max														

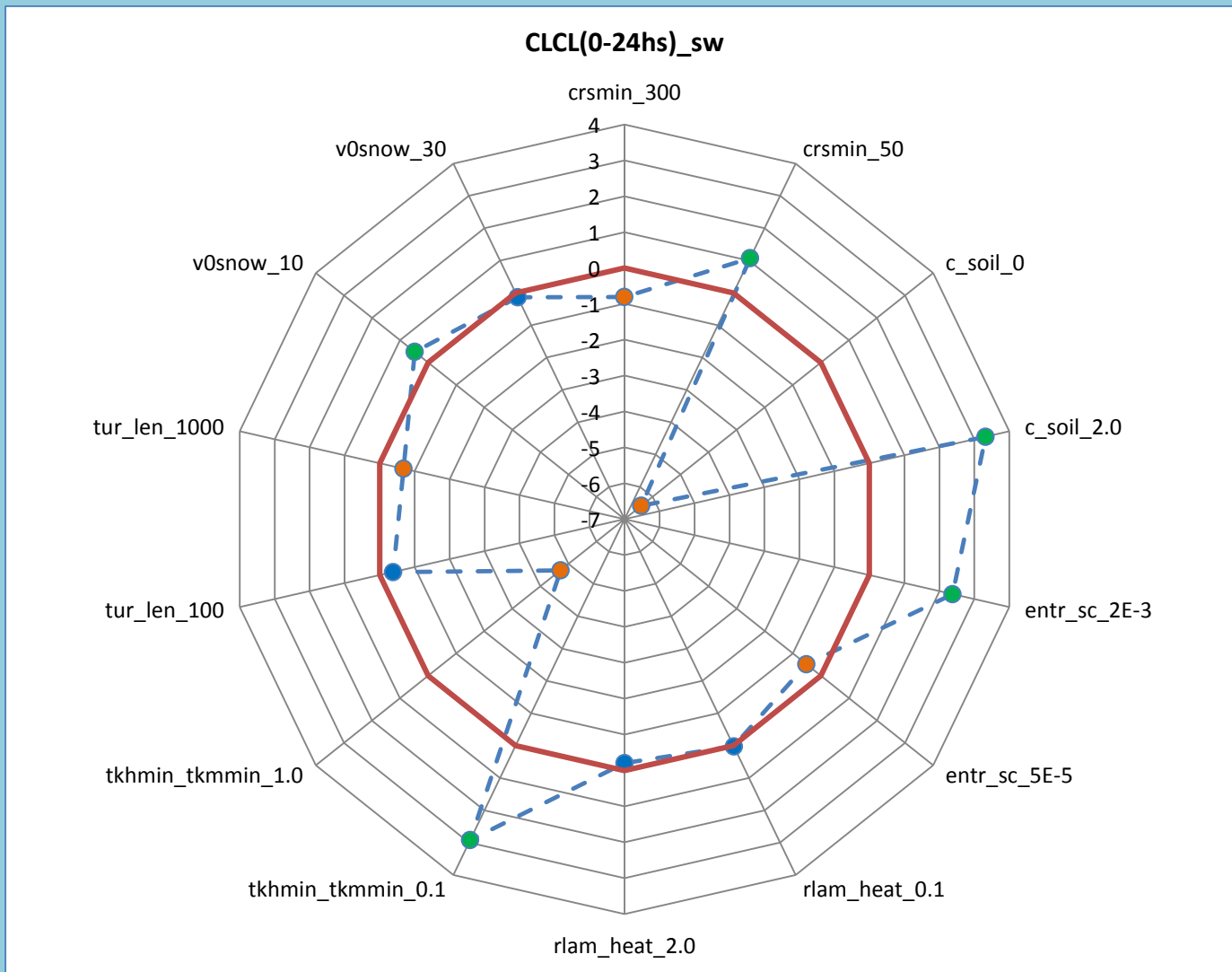
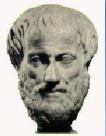


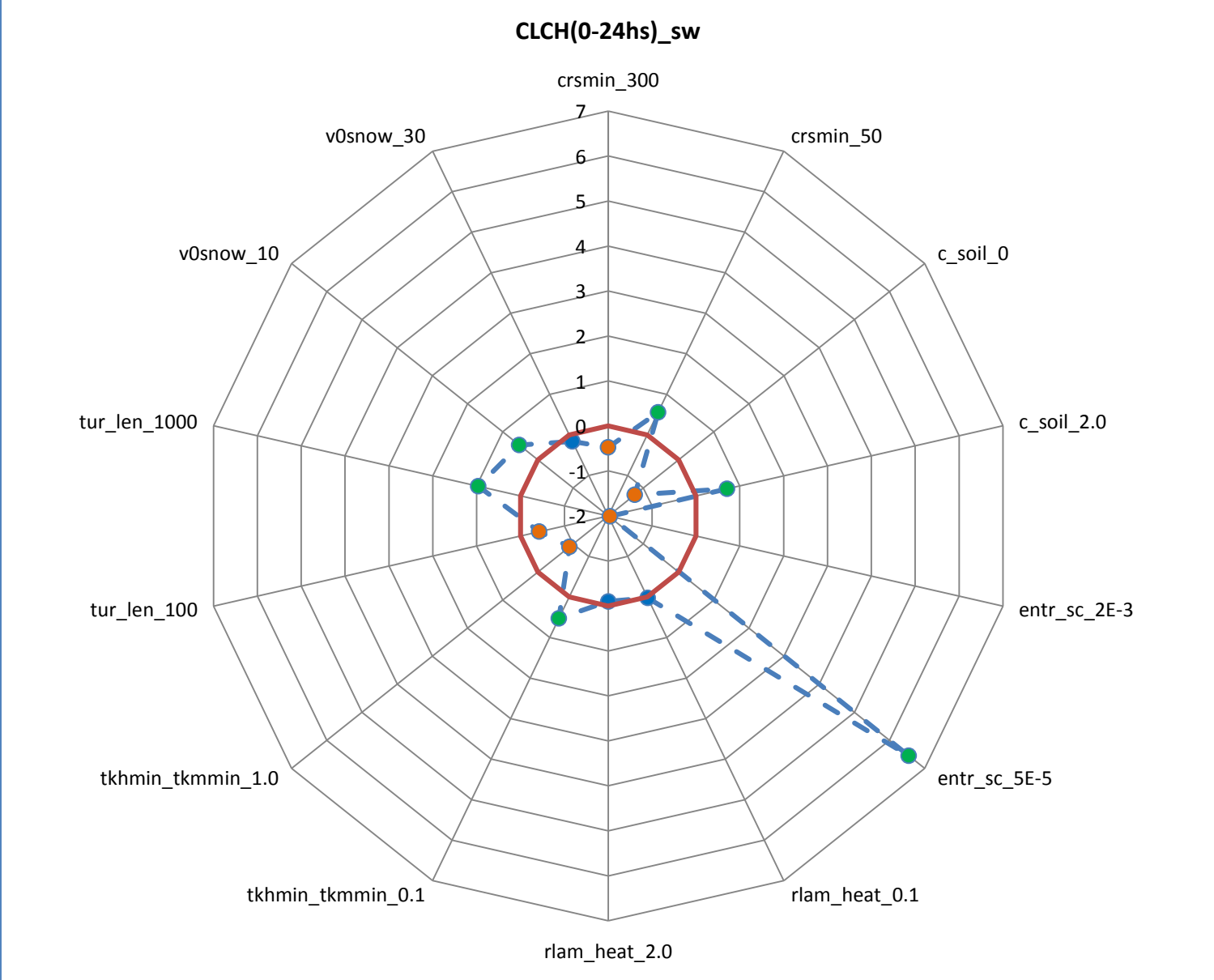
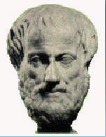
NEXT STEPS ...

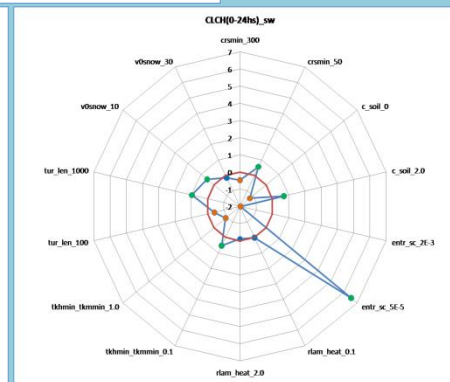
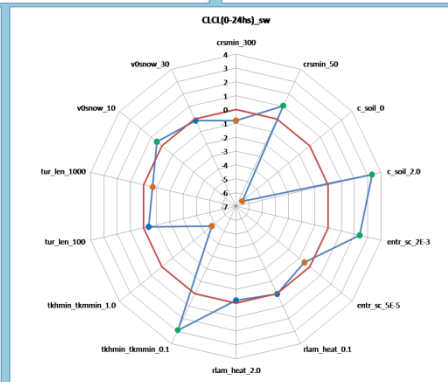
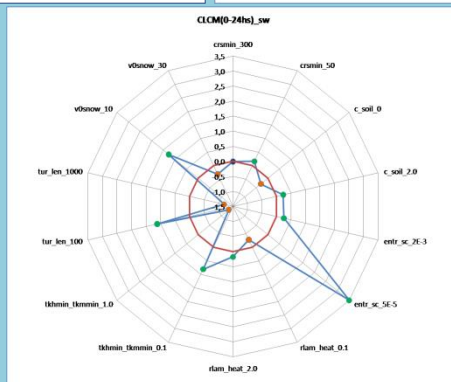
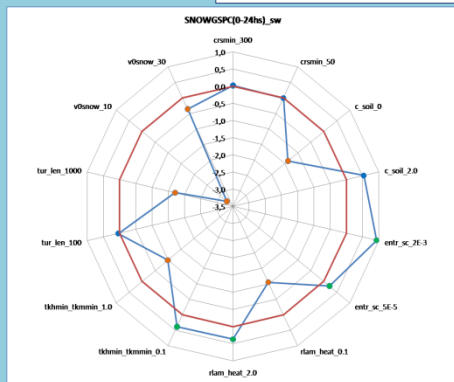
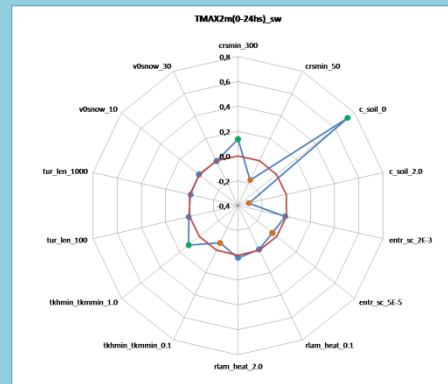
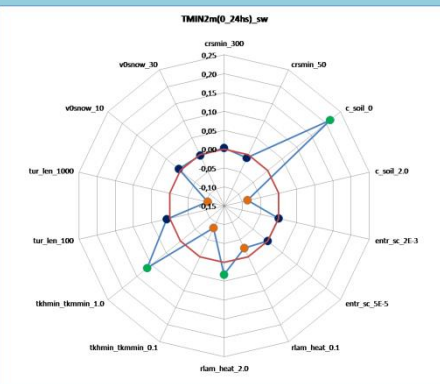
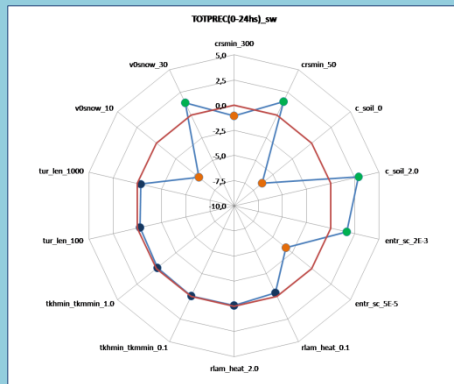
The next steps are similar and are based on the spider graphs for the remaining model variables, CLCM, SNOWGSP, CLCL, CLCH, i.e:











Interesting features regarding the sensitivity of the model can also be drawn by placing all the spider graphs together especially due to the relative changes in their convexity.



NEXT STEPS AND SOME MUSINGS

- ⊕ As the steps progress upon the remaining spidered graphs, the board can be filled (if necessary).
- ⊕ If any cells remain empty after this process, most probably these parameter pairs will not contribute anything to the optimization upon the use of a metamodel.
- ⊕ Upon the results coming upon the successive consideration of the meteorological variables, the metamodel can be run to trace if there are any progressive improvements and also have some hints where to place any intermediate values if needed.
- ⊕ The method can be readily generalized to include groups of three, four or any parameters. In fact it could develop into a (scientific) game (in the spirit of **Reiner Knizia**, a tiny part of his work was humbly used as an inspiration).
- ⊕ The method can also be modified upon the choice of the:
 - ◆ [Geographical Domain\(s\)](#)
 - ◆ Choice of the Tested Time Period(s)/Dates according to the availability of computer resources, e.g:
 - ⊕ Continuous period (like present CALMO Project)
 - ⊕ A set of extreme events, possibly over several years.
 - ⊕ A set of random dates where seasonal dependence is taken into account.
 - ⊕ An interesting option could be to use a small number of *specially selected* dates and apply the methodology and the metamodel with the largest possible number of parameters.
- ⊕ Spider (Arachne in greek) is supported by a sad, as well as cryptic, myth but it is hoped that the methodology will be of fair value to the exceptionally complex «web» of designing and applying the tests of model optimization.



Eisen, Charles (1720-1778)



Velázquez, Diego 1599-1660

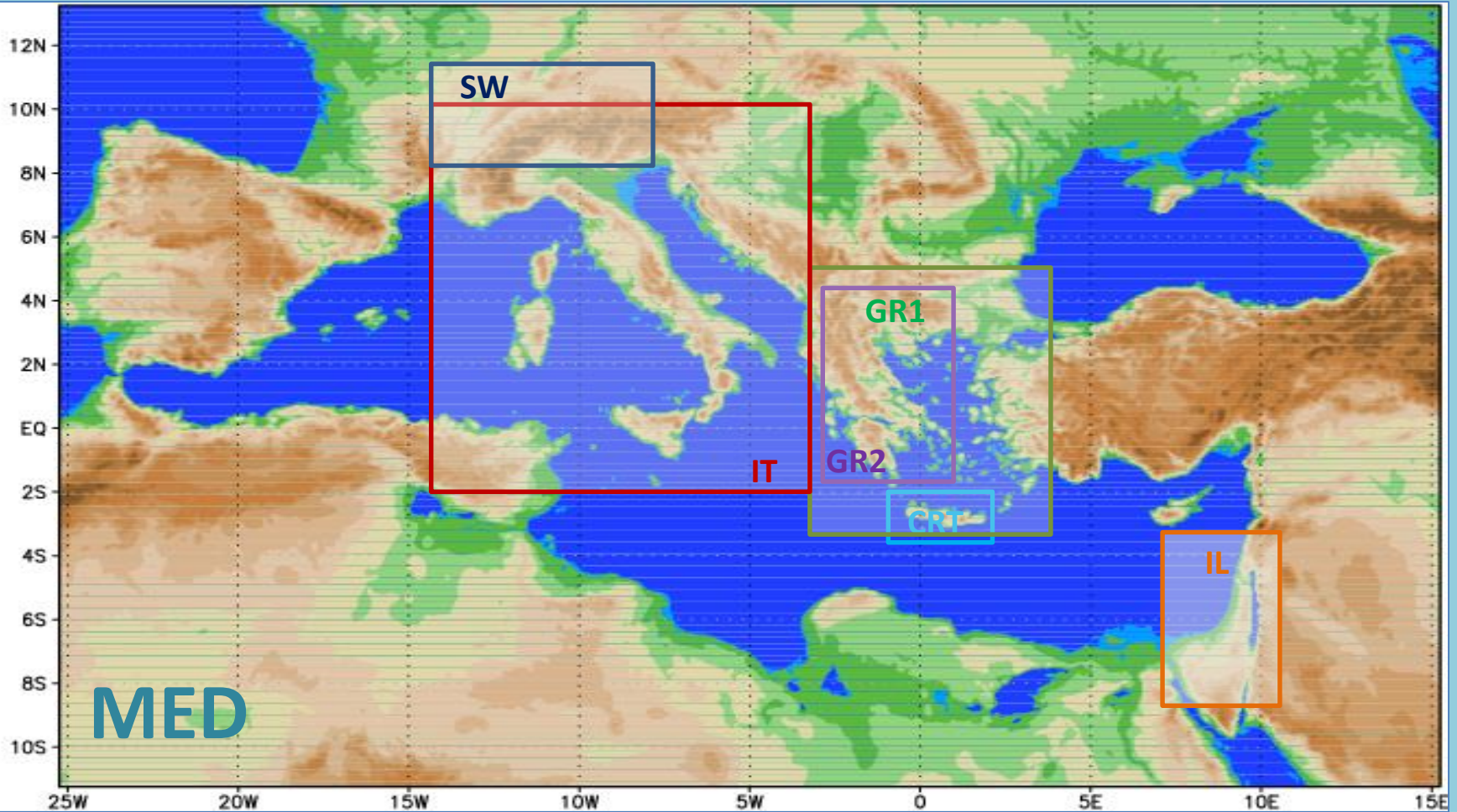
Thank You !



Questions ?



SELECTED DOMAINS



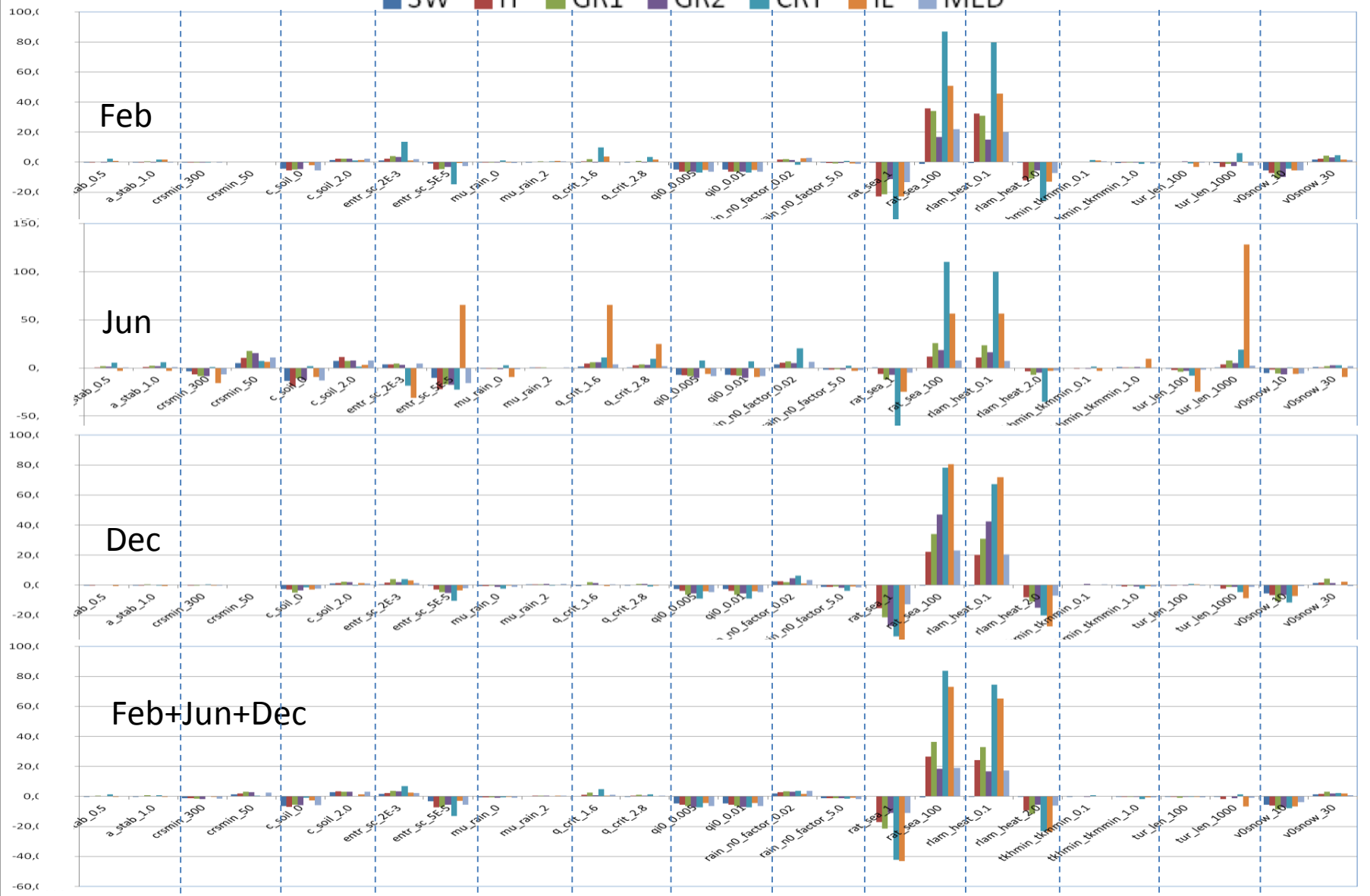
Selection based on COSMO countries location from the SouthEast (domain IL) to the NorthWest (domain SW) in order to reveal any changes to the [parameter sensitivities](#) due to the climatic differences of the domains.

- ◆ Marine vs Continental areas (GR1 vs. SW or GR2 vs CRT(!))
- ◆ Significant longitudinal difference (GR vs IL)



Sensitivity (%) for TOTPREC

Legend: SW (dark blue), IT (red), GR1 (green), GR2 (purple), CRT (teal), IL (orange), MED (light blue)



Params:	a_stab	crsmin	c_soil	entr_sc	mu_rain	q_crit	qi0	rain_n0_factor	rat_sea	lam_heat	tkhmin_tkmmmin	tur_len	v0snow
Lims:	0.0, 1.0	50, 300	0, 2	5E-5, 2E-3	0, 2	1.6, 2.8	0, 0.01	0.02, 5.0	10, 100	0.1, 2.0	0.1, 1.0	100, 1000	10, 30
Def:	0.5	150	1	3E-4	1	4.0	0.005	1.0	20	1.0	0.4	150	20