CALMO – recent achievements and open questions

P. Khain, A. Voudouri, J. M. Bettems

COSMO General Meeting, Eretria, September 2014

Outline

1. Removing built-in temperature biases in the areas of complex topography

2. COSMO-7 sensibility to the change of parameters

3. Further development of the Meta-Model

4. Additional important tasks

Outline

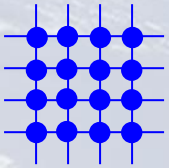
1. Removing built-in temperature biases in the areas of complex topography
2. COSMO-7 sensibility to the change of parameters
3. Further development of the Meta-Model
4. Additional important tasks

Built-in temperature biases in the areas of complex topography

Consider 2m-temperature observations from nearby located stations

or:

"Observations grid"
Ex: 2m-temperature observations grid over Switzerland (C. Frei)
resolution: 2km, on real terrain



Goal: verify coarse grid model (ex: COSMO-7km)

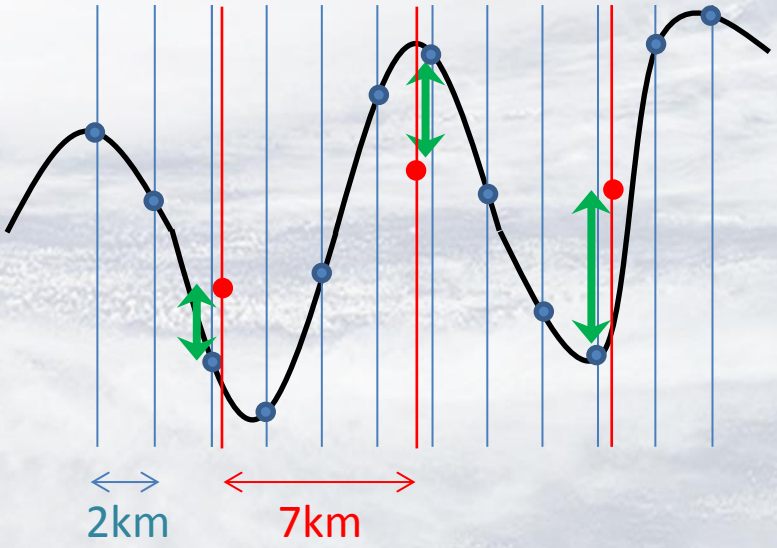
But: there are no observations located exactly on the coarse grid

Option: **linearly interpolate** the observations to the coarse grid points. Bad…

The coarse grid has smoothed terrain!
The model will not be able to correctly predict the observations

2m-temperature very much depends on height. Usually: higher = colder

If the coarse grid point is too low
→ temperature too high
If the coarse grid point is too high
→ temperature too low



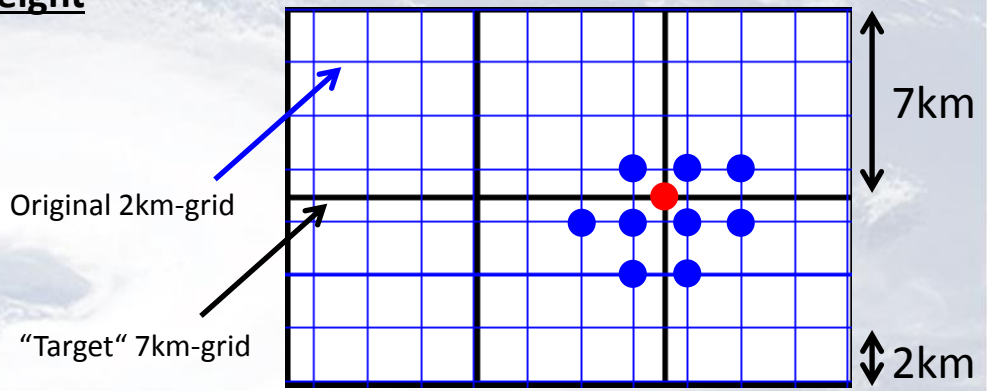
Must adjust the observations to the model smoothed grid **before** verification! But … ***How?***

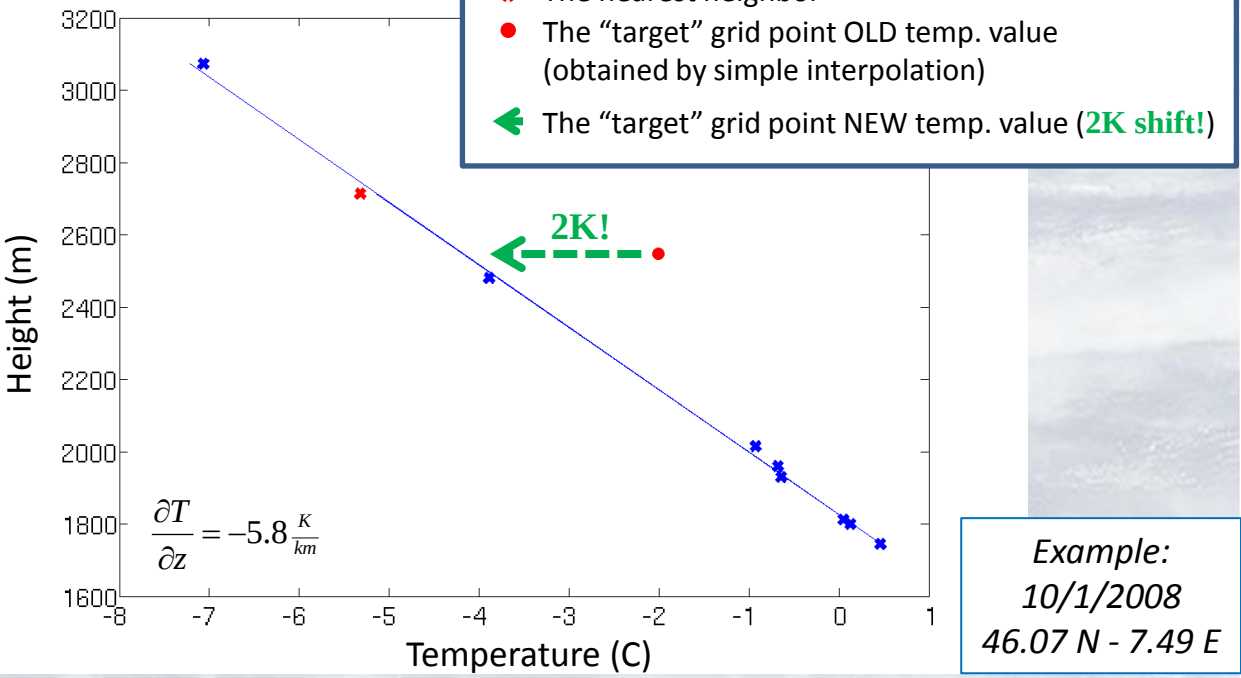
2m-temperature adjustment to the (smoothed) model grid

- C. Frei: Interpolate considering the "neighbors" **height**

1. For every grid point in the "target" grid (red dot), find the nearest 9 neighbors on the original 2km-grid (blue dots)
2. Plot the 2m-temperature values of these neighbors vs. their altitude (blue and red "x").
3. Perform a linear fit of the data, which will be the ***local 2m-temperature profile***
4. Having the altitude of the "target" grid point, use the linear regression, to calculate its 2m-temperature
5. Perform this operation (1-4) for every "target" grid point, for every day

Local profile:
works even for inversions!





Example:
10/1/2008
46.07 N - 7.49 E

Outline

1. Removing built-in temperature biases in the areas of complex topography

2. COSMO-7 sensibility to the change of parameters

3. Further development of the Meta-Model

4. Additional important tasks

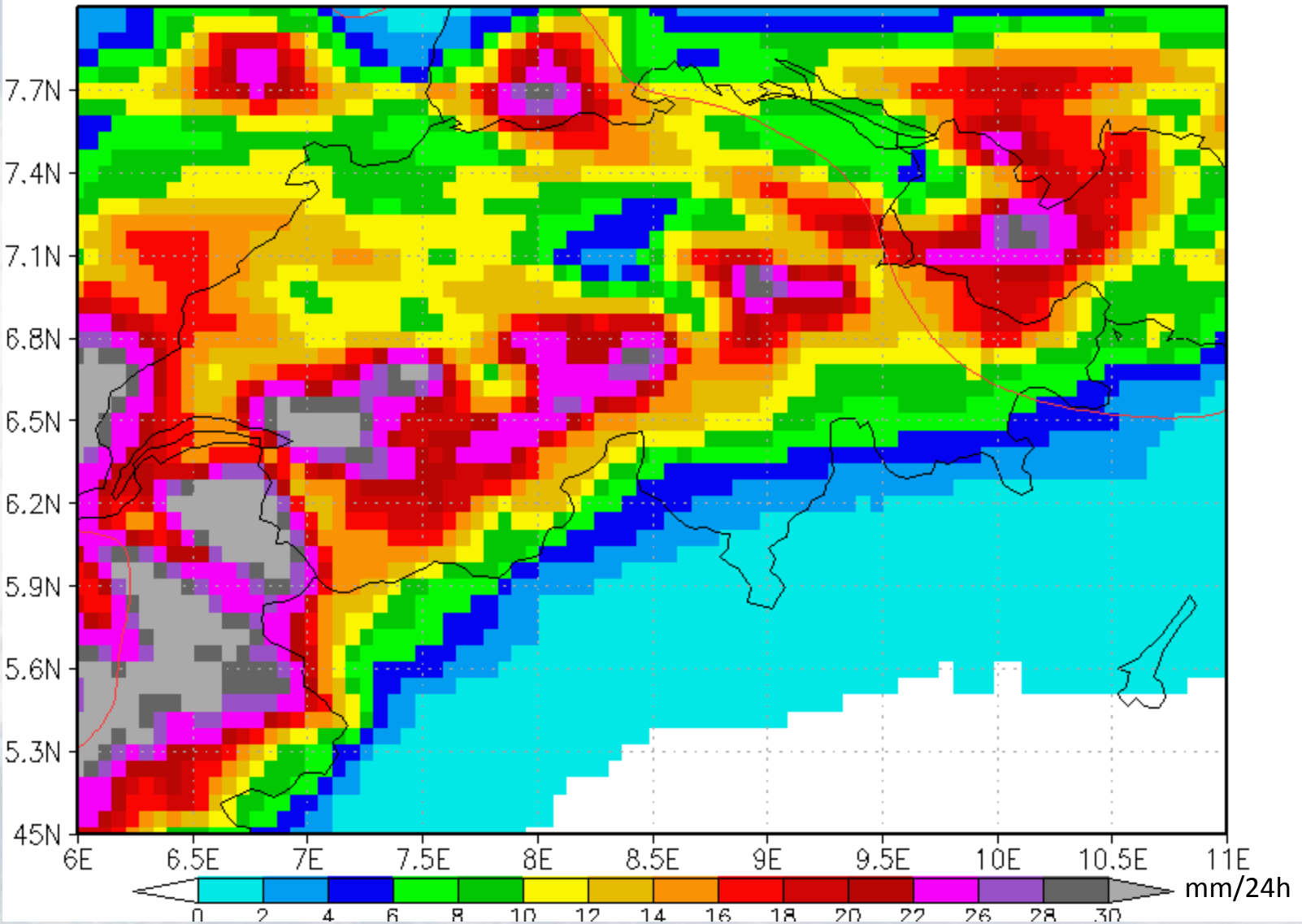
A. Precipitation sensibility to specific parameter highly depends on the region

Example:

Sensibility of 24h-precipitation (during 7/1/2008)
to the change of “rat_sea” (evaporation over sea)

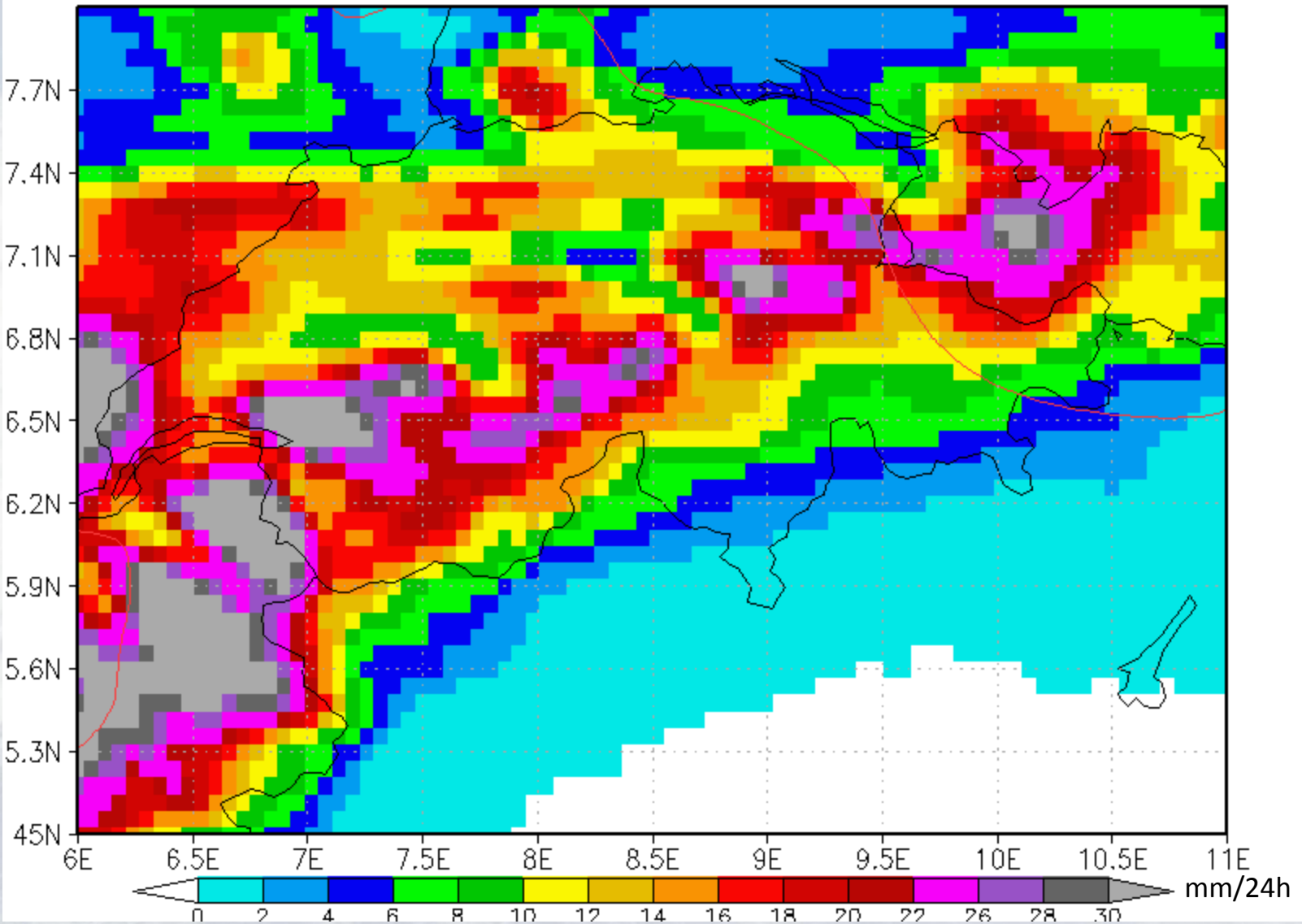
Zoom over Switzerland

24h accum. precipitation during 7/1/2008
Pr(rat_sea=100)



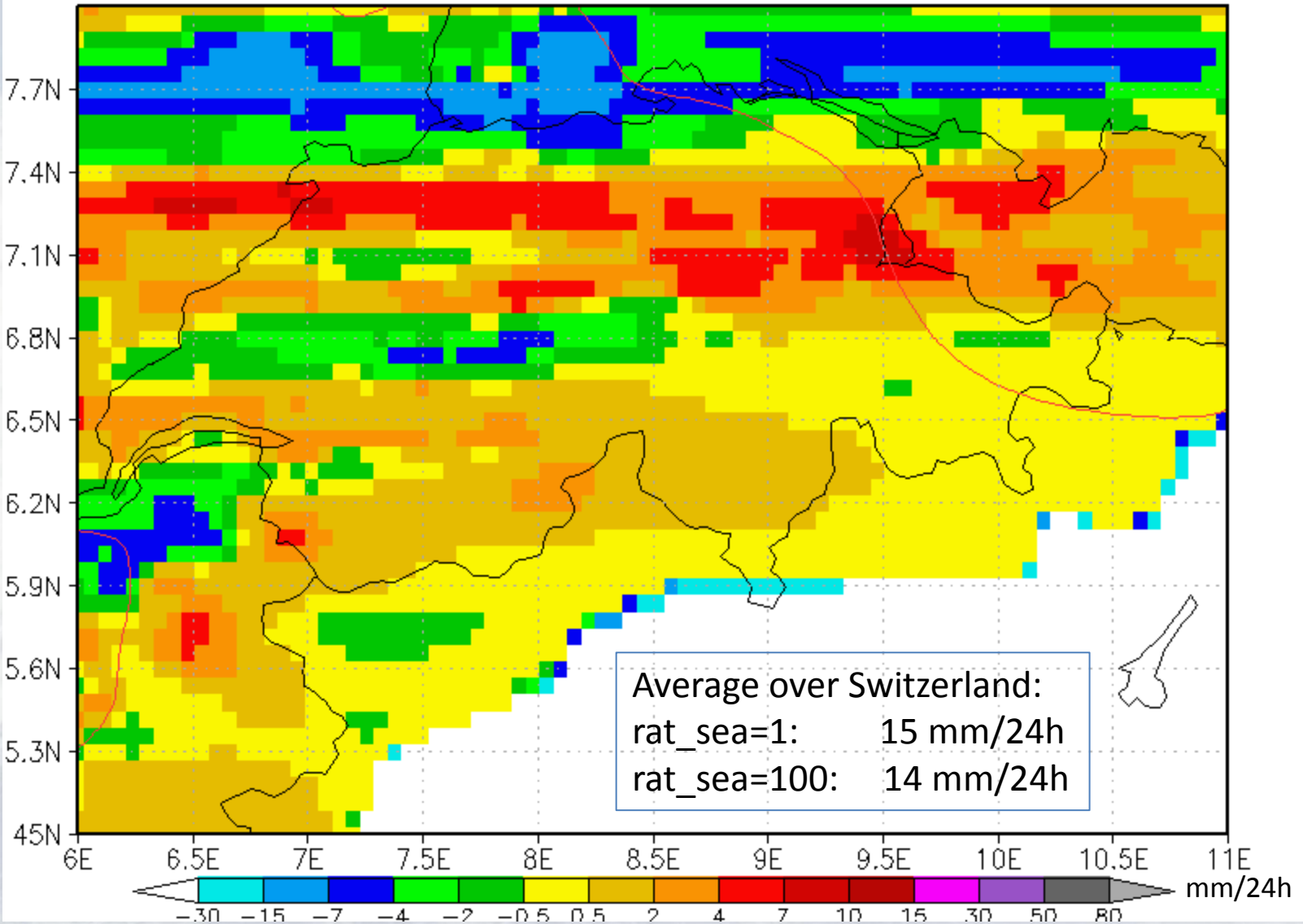
White areas = no rain

24h accum. precipitation during 7/1/2008
Pr(rat_sea=1)



White areas = no rain

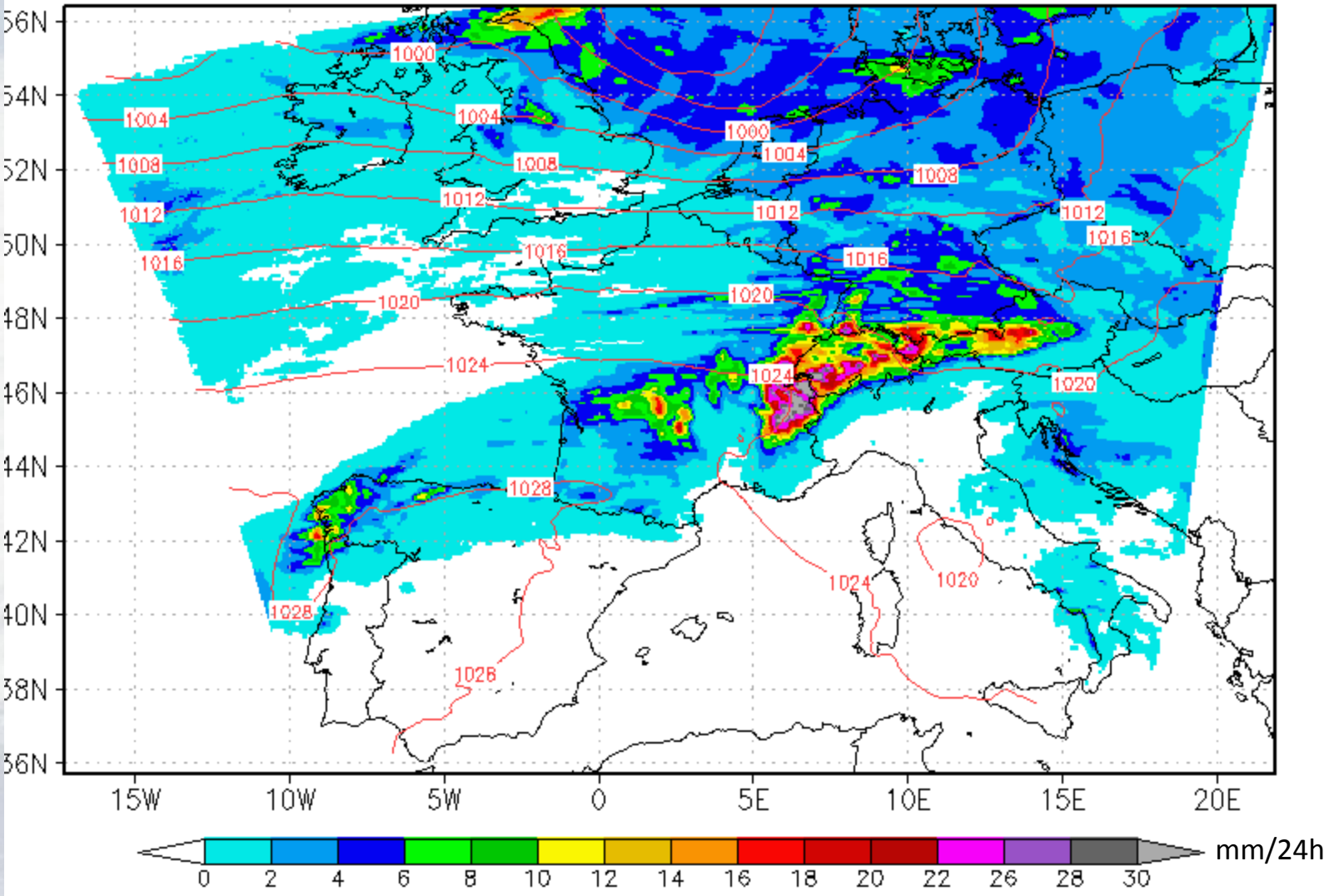
Difference of 24h accum. precipitation during 7/1/2008
Pr(rat_sea=1)-Pr(rat_sea=100)



White areas = no rain

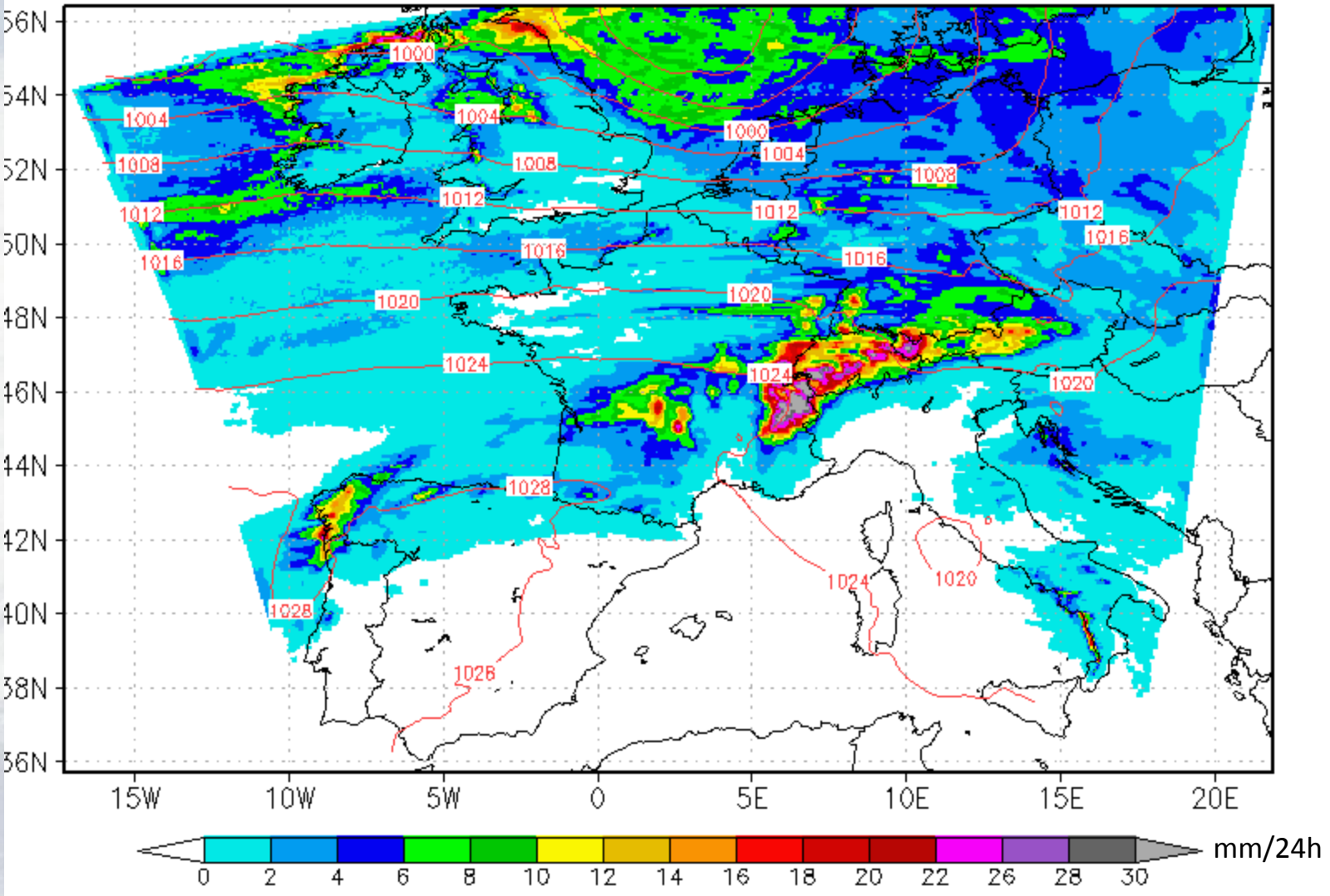
- Low precipitation sensibility for “rat_sea” over Switzerland.
- Precipitation over Switzerland is very much orographic, and not significantly affected by the Mediterranean Sea or the Atlantic Ocean.
- But, when looking on the whole domain, the picture is different…

24h accum. precipitation during 7/1/2008
Pr(rat_sea=100)



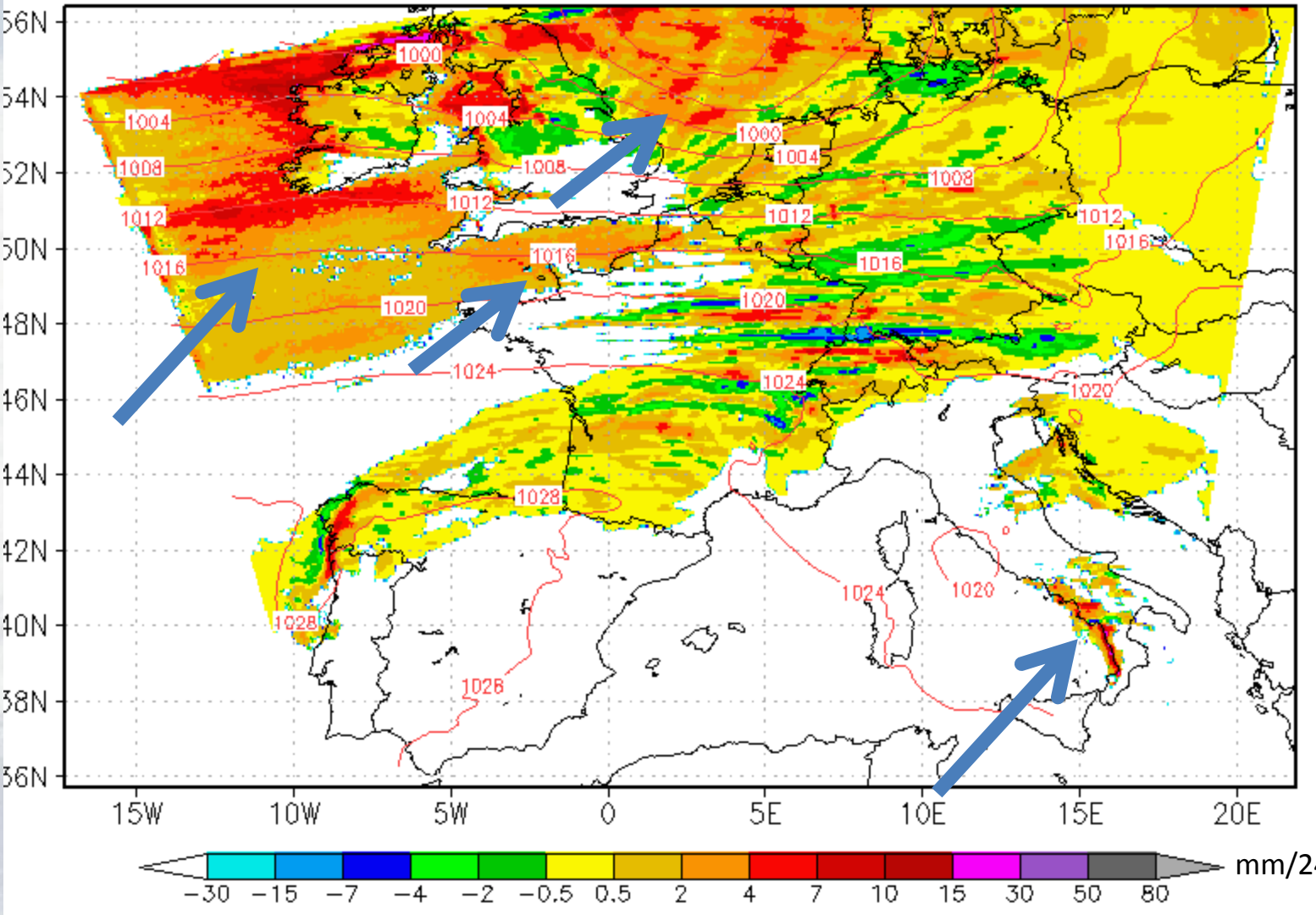
White areas = no rain

24h accum. precipitation during 7/1/2008
Pr(rat_sea=1)



White areas = no rain

Difference of 24h accum. precipitation during 7/1/2008
Pr(rat_sea=1)-Pr(rat_sea=100)

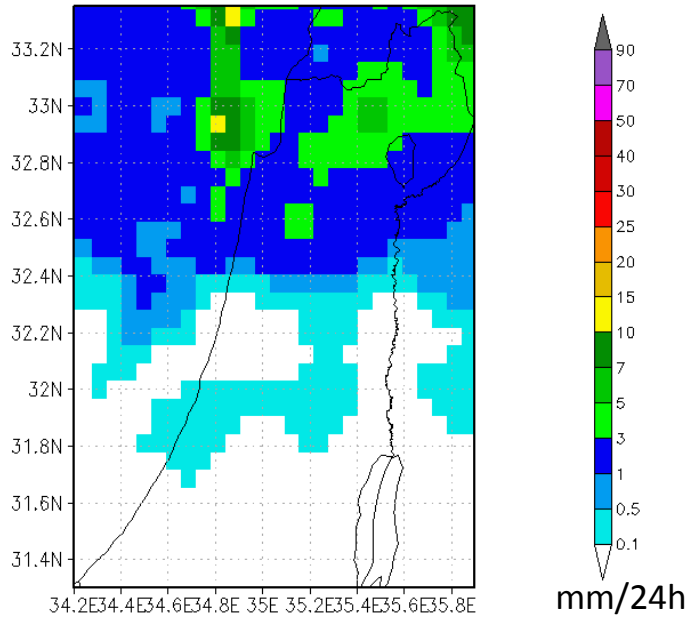


White areas = no rain

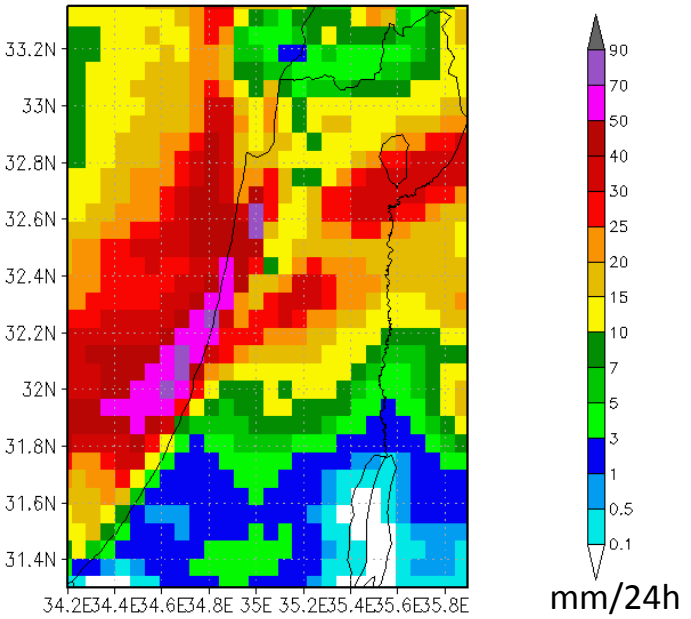
Example:

Heavy rain event over the Eastern Mediterranean during 10-14.12.2013

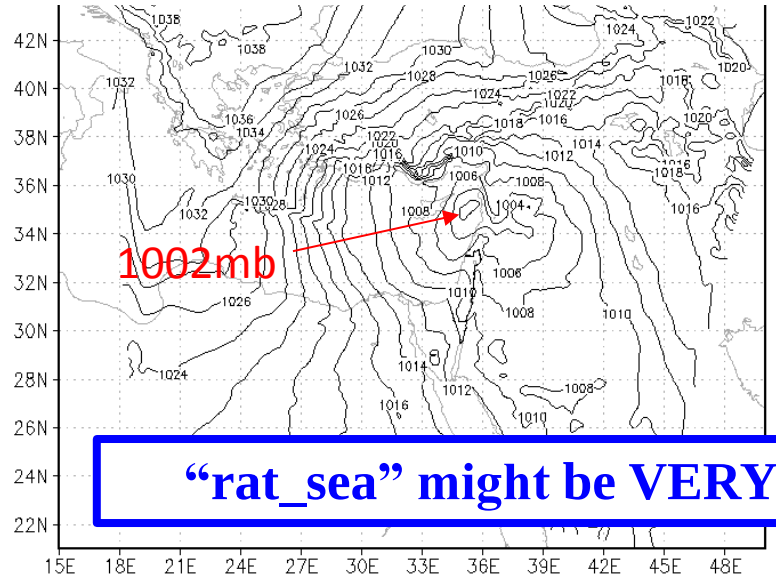
24h-precip. (10.12.2013) over Israel. rat_sea=100



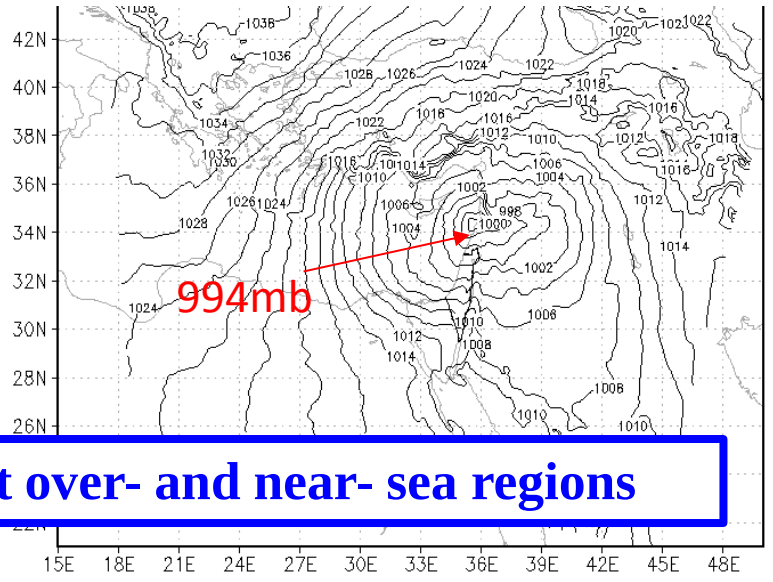
24h-precip. (10.12.2013) over Israel. rat_sea=1



MSL (11.12.2013 12UTC) over Eastern Med. rat_sea=100



MSL (11.12.2013 12UTC) over Eastern Med. rat_sea=1

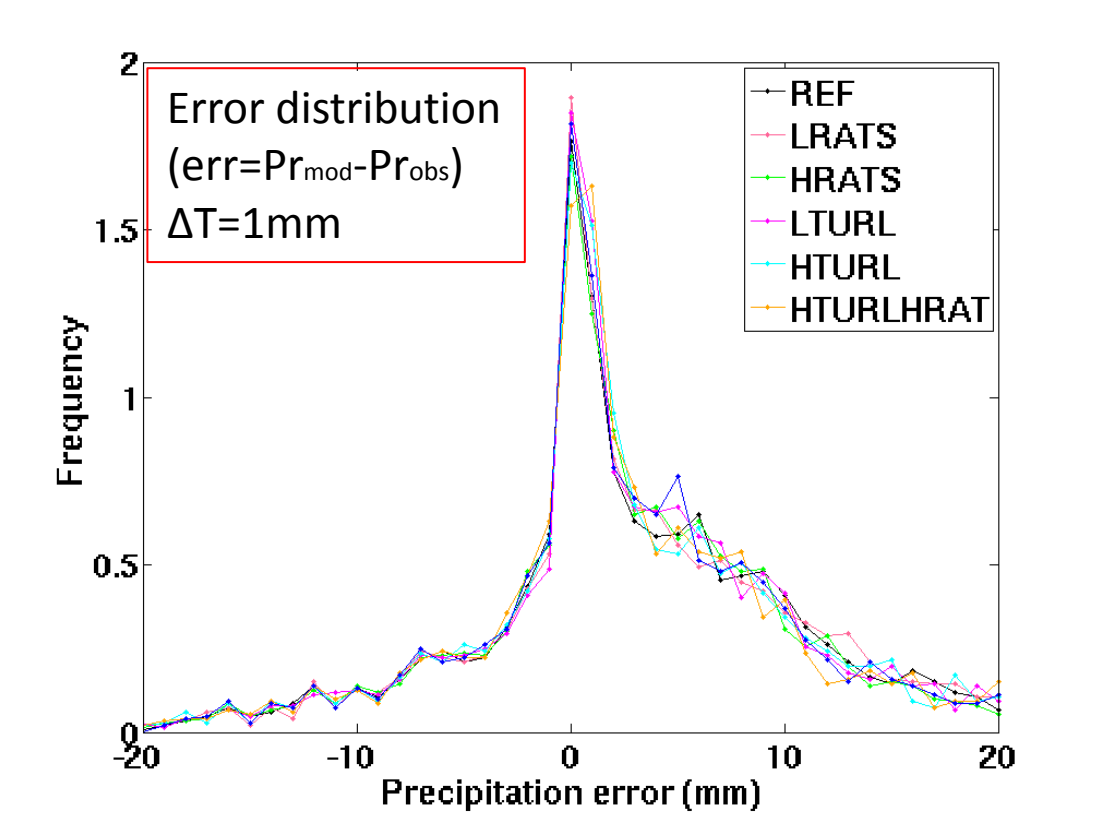


"rat_sea" might be VERY important over- and near- sea regions

B. Precipitation is a very noisy field

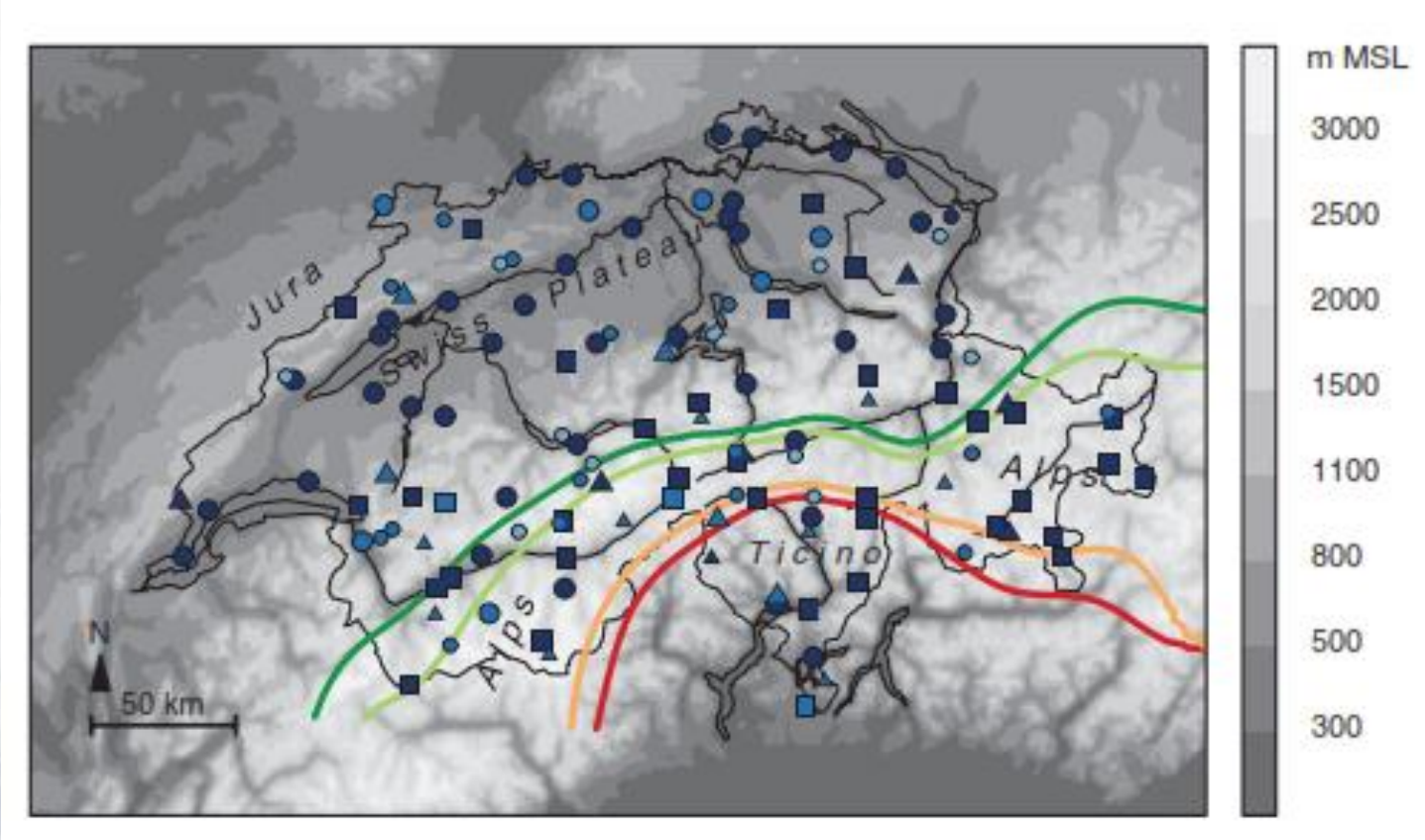
Omar Bellpart:

Even in my application (climate) using larger spatial averages and monthly means, the precipitation remained the hardest variable to predict due to the level of noise. A smoother statistic would help, yet considering that it should remain a verification of an NWP application.



Have to consider: working with smoother precipitation fields, which can be obtained after bigger temporal averaging (several days) and bigger **areas averaging.**

Instead of 152 regions, we can divide Switzerland to several “physically” selected regions:



C. Frei, 2013

C. Precipitation sensibility to parameters depends on the season

Jean-Marie Bettems:

"The insensitivity of the precipitation … could be due to the type of precipitation at this time of the year (mainly frontal systems in winter/spring); I expect more sensitivity for convective precipitation".

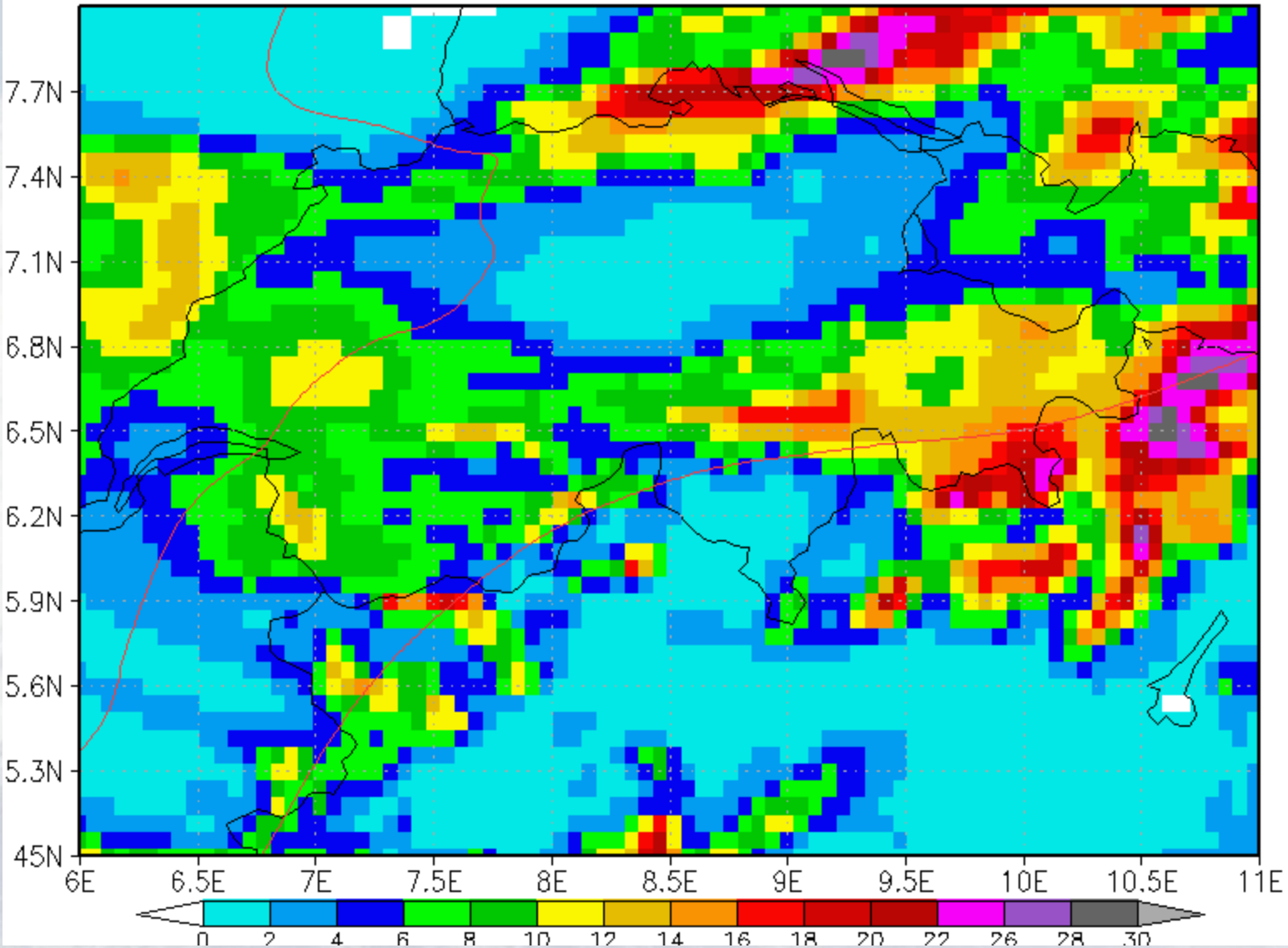
We indeed have seen some sensitivity of precipitation (convection) … over Switzerland with respect to the turbulence length scale in summer, ..."

Example:

Sensibility of 24h-precipitation (during 7/**6**/2008)
to the change of "tur_len" (turbulent length)

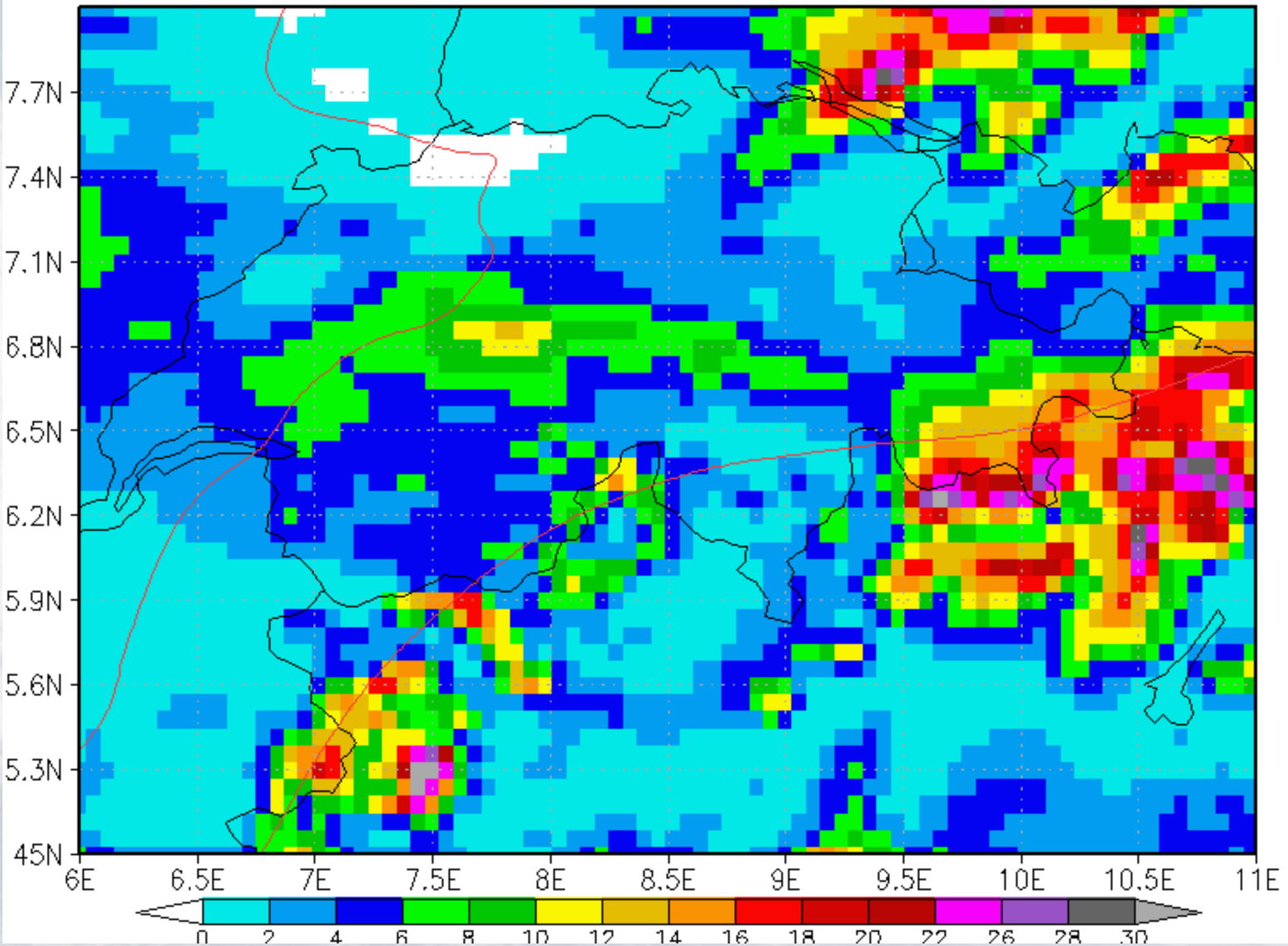
Zoom over Switzerland

24h accum. precipitation during 7/6/2008
Pr(tur_len=10000)



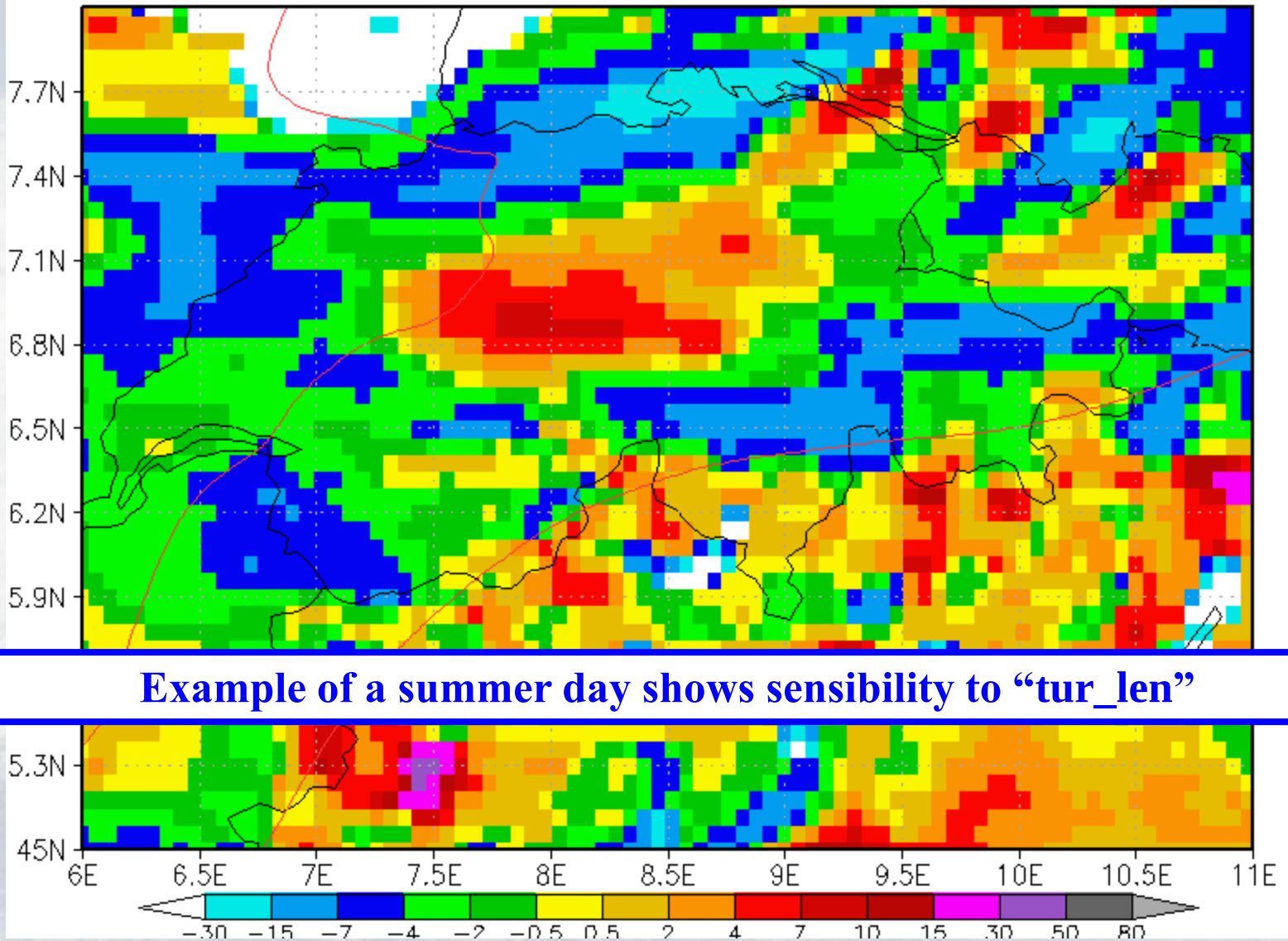
White areas = no rain

24h accum. precipitation during 7/6/2008
Pr(tur_len=100)



White areas = no rain

Difference of 24h accum. precipitation during 7/6/2008
Pr(tur_len=100)-Pr(tur_len=10000)

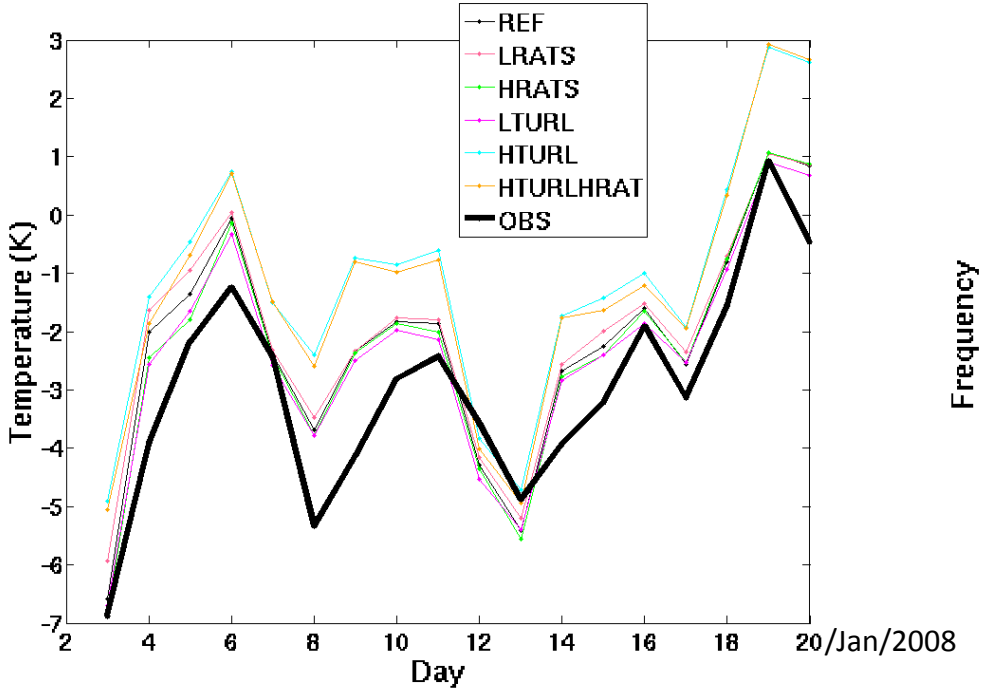


Example of a summer day shows sensibility to "tur_len"

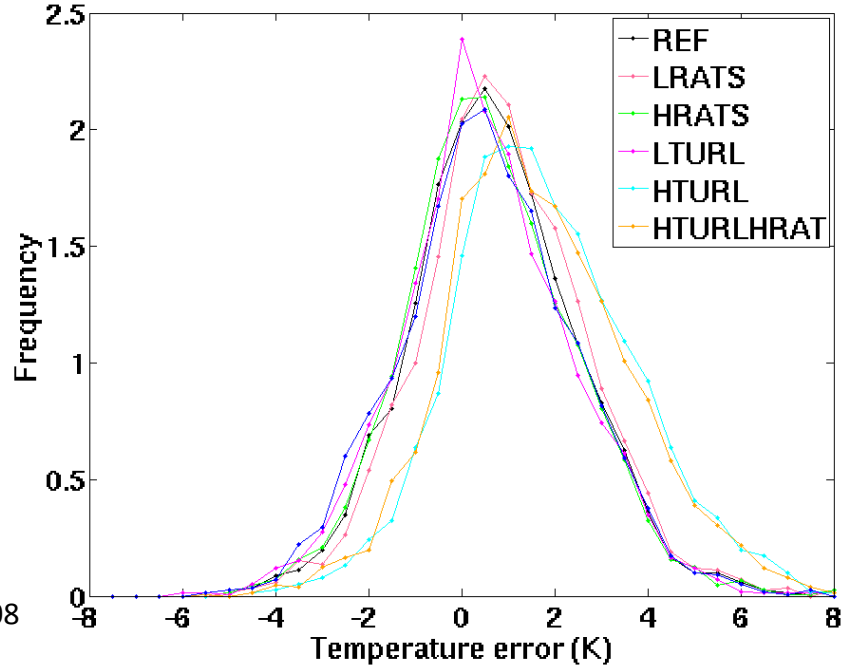
White areas = no rain

D. Minimum 2m-temperature sensibility to parameters

Tmin- averaged over whole Switzerland

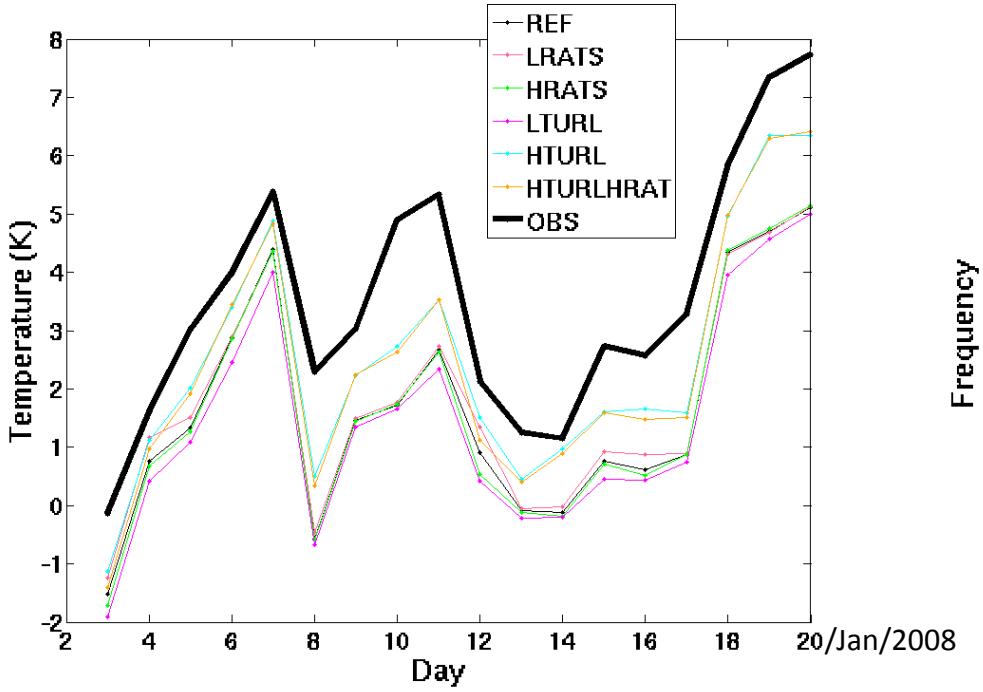


Tmin error distribution (err=T_{mod}-T_{obs}, ΔT=0.5K)

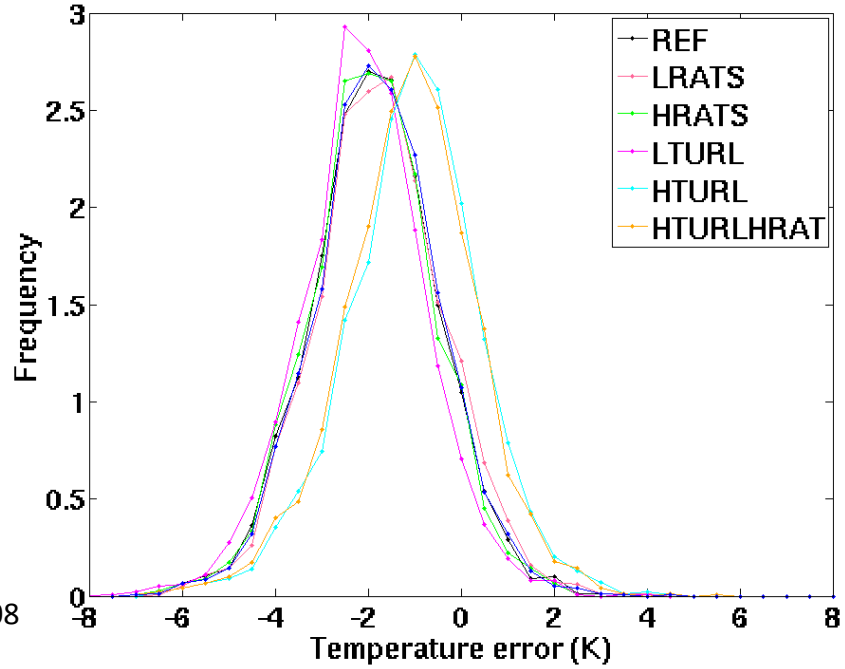


E. Maximum 2m-temperature sensibility to parameters

Tmax - averaged over whole Switzerland



Tmax error distribution (err=T_{mod}-T_{obs}, ΔT=0.5K)



- COSMO gives too high Tmin and too low Tmax
- Tmax is better predicted then Tmin (narrower error distribution)
- Apparently, higher "tur_len" yields higher temperature forecast - **Sensibility!**

Outline

1. Removing built-in temperature biases in the areas of complex topography

2. COSMO-7 sensibility to the change of parameters

3. Further development of the Meta-Model

4. Additional important tasks

Further development of the Meta-Model

1. Consideration of simulations noise due to internal variability. Particularly important when using the "interaction terms" to build the Meta-Model.

Omar Bellpart:

In some applications I realized that considering the interaction terms actually worsened the accuracy of the metamodel, quiet dramatically in some cases. This also occurred even if the interaction terms were very small and I'm still not sure if I understand this correctly. Unfortunately I never had time to test this properly. This occurs only if only one simulation is used to determine the interaction terms, so I eventually performed more than one for each parameter combination (as in Neelin 2010).

…

The interaction is very small but large interaction terms are estimated due to large noise of the data.

…

The data of the simulations are noisy because of the internal variability of the RCM. This level of internal variability corresponds to the noise level that I considered in the code. In case of CALMO we actually never tested what the sensitivity to the perturbations of the initial conditions ...

In the current code there is no proper consideration of this noise in the regression of the metamodel. This can be problematic in particular for the interaction terms since before estimating these, several differences of simulations are computed - which increases the noise level.

…

… which increases the noise level by a factor of 12! Not 5 in fact realize now, so this fact is vastly more important than I thought previously and that's a challenge. This explains why we need multiple simulations to determine a robust signal for the interaction term...

Unfortunately I'm unable to work further on this at the moment, but feel free to develop further. I guess there are two options, either estimate the metamodel differently or considering interaction terms only if they surpass the noise level of a factor of 12 if only one simulation is considered. Hope this helps, feel free to ask again if this is not clear!

…

… And I realized that you get a factor of 12 times the standard deviation, so increase of variance of around 48. But I didn't use much time testing it, perhaps would be useful if you do the same test..

…

I would consider to use 2-3 interaction simulations if there is indication that the interaction seems to be large, otherwise omit the interaction. With such a high noise level this is, however, hard to test. *What I tried is to consider the interaction only if the term dint_p1p2 if the signal is larger than the noise. Since it seems to be 12 times the internal variability this would rule out most of the terms, not sure if that's the best way to go.*

To test the internal variability in CALMO I would consider a random perturbation, not sure if COSMO has something like singular vectors for perturbations. I'm not sure about that, since I have little experience with NWP. We discussed this issue with Jean-Marie in the beginning of the project and he said that the internal variability would be very small since lead time is short, but it might be important considering this inflation of noise for the interaction terms.

Further development of the Meta-Model

2. Build the Score Function considering the uncertainty of different variables

Christoph Schär:

Some info might be needed regarding the performance score, as the one used by Omar is not appropriate for your purpose. The importance of the score is evident and will strongly affect the results. You will not need to provide a full description of the score in the proposal, but some indication about the directions you intend to take. Make sure the score is geographically balanced, ***use the uncertainty of different variables (i.e. T, P, etc.) to weight the different contributions****, and use only those observations that are consistent with the modeling framework and resolution (e.g. no point observations at extreme mountain or valley stations).*

3. Adaptations to COSMO-2.2 and COSMO-1 grids, including observations adjustments

4. Adaptations of the Meta-Model when additional fields will be analyzed

Outline

1. Removing built-in temperature biases in the areas of complex topography

2. COSMO-7 sensibility to the change of parameters

3. Further development of the Meta-Model

4. Additional important tasks

Additional important tasks

1. Data thinning

Pavel Khain:

A. First of all, we have to be "objective", so instead of choosing the year 2008, it makes better sense to choose representing months. **Using Swiss observations database**, we can choose climatologically representative months out of recent several years. For example, the most representative January may be taken from 2013, the most representative February may be taken from 2009, and so on. We can save computer resources by choosing the representative month for **entire season**, and run our simulation over 4 month only. For example, the most representative January may be taken from 2013, meaning that its averages resemble the averages of all the winter months (Dec, Jan, Feb) during recent several years. This is the way people usually prepare so called "wind atlases" for wind energy projects.

B. However, (A) is still a "climatological" approach. Our goal is to improve forecasts and not the climatological averages (as for RCM's). When some "expert" tries to tune parameters, he chooses few case studies, when the forecast errors were large.

   Therefore I could suggest, again using Swiss database, to pick out (objectively!) 2 groups of days:

   1. Days when the original COSMO 7km errors where the largest.
   2. Days when the original COSMO 7km errors where of standard quality (mainly due to biases), may be using approach (A).

   Then we could run our simulations on these two groups of days. The first one would represent our goal to improve the model in "difficult" situations, and the second one - not to spoil the climatological averages. Then, using the "CALMO score function" we can decide which weight to give to each of these groups. This approach would, from one hand, deal with the real forecast problems, and from the other hand save computational time.

Jean-Marie:

These are useful input for a second phase of the project where one tries to minimize the cost of the method (and to make the method less arbitrary with respect to the choice of the tested period)

Additional important tasks

2. Additional fields

Christoph Schär: … More recent work of Omar (see some added text in the proposal) has resulted in an additional very substantial improvement in the case of COSMO-CLM, beyond that described in the paper. The key parameter was the **hydraulic soil conductivity**. It is important that this parameter is also considered in your calibration. You are probably aware of this work and anyway considered to do so (I just mention this to make sure).

Omar Bellpart: … **hydraulic soil conductivity strongly affects the summer climate** …

Jean-Marie Bettems: … One sees that the definition of the quality score is a difficult problem. In my opinion, one should add a third type of gridded observations, representative of the PBL and catching the daily cycle; maybe **GPS derived integrated water vapor** could be an option (although I do not know if gridded values are available).

Jean-François Mahfouf: "Address the uncertainties associated with the **look-up tables, especially for the SVAT model**": This aspect is particularly important; objective parameter optimization should be considered.

J. Helmert: The automatic parameter optimization procedure for tuning of look-up tables could be integrated into the COSMO project PP CALMO.
... could one integrate in some way an objective calibration of some of the **look-up tables used in the derivation of the external parameters for the soil and surface**? This is worth considering ...

Jean-Marie Bettems: I still think it is important to find a third type of observations, possibly **related to BL humidity**, in the definition of the quality measure.
A parameter describing the **hydraulic conductivity of the soil should be introduced** in the set of tunable parameters.
A method to represent the uncertainties introduced by the **external parameter look-up tables for the vegetation** should be evaluated.

Additional important tasks

2. Additional fields

Jean-Marie Bettems:

Integrated water vapor is an interesting candidate, being essentially representative of the PBL and having a high variability.

So my questions:

- Ø Do you know if gridded fields of observed GPS ZTD over central Europe exist?
- Ø If this is the case, what is the quality and the temporal resolution of these data?
- Ø If this is the case, do you know if these data are available for the year 2008?
- Ø If this is the case, do you know who to contact to access these data?

Henrik:

…Depending on type of terrain and inter GNSS site distances the quality of the gridded product will vary a lot from region to region. Using the actual ZTDs reduces the risk that variations you might see with respect to COSMO are due to this inhomogeneity, not due to COSMO. The longer you go back in time, the more important this becomes, since the GNSS ZTD coverage was much more patchy in the past than now. And still now it varies a lot.

Elmar:

I support Henriks advice to use real GPS stations instead of gridded information.
Might be the easiest for you and your colleagues is to follow a similar strategy as the NWP colleagues do already now (might be with our GNSS ZTD estimates which of course does not cover all Europe): using the COST files from the EGVAP server for which also a combined products should exist in 2008 (makes it easier and more independent of the contributing analysis center).
A second option is to use e.g. troposphere estimates of a global reprocessing (but attention: this series are generated in postprocessing and will have better quality and consistency over longer periods than estimating ZTDs in near-realtime; station availability is global and therefore sparse in Europe). Which density you are looking for? Very soon, also European repro results will become available.

Rosa:

Yes we have Gridded GNSS ZTD data over Europe… It is running on a routine basis but unfortunately **we don't have any gridded data for 2008**.

Additional important tasks

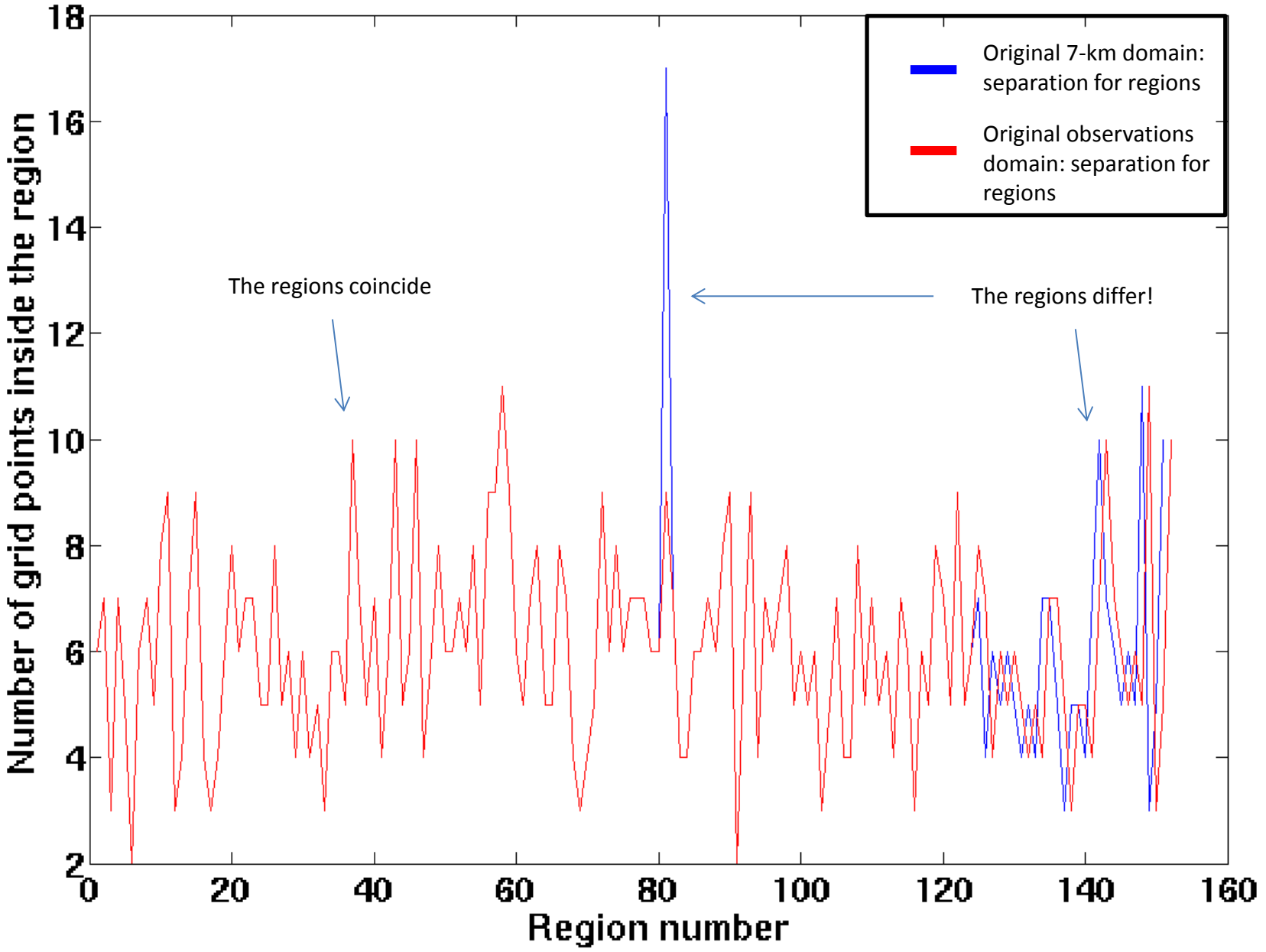
3. Soil initialization

Jean-Marie Bettems: … `A strategy for the initialization of the soil should be fixed…`


THANK YOU !

Additional slides …

Found discrepancy in few regions between the COSMO domain and the observations domain



Dear Pavel,

I'm back from my travels, it was a very inspiring journey, particularly Iran was full of surprises for me. Now back and trying to clear my mailbox :)

So to your questions:

1. Honestly can't recall exactly what I did in that great hurry, probably not a very meaningful thing. In principle I had defined two matlab files containing the indices for the regions, one for the small COSMO domain (regions_small.mat) and one for the large (regions_large.mat). Since for the observations are mapped for the small domain I used the regions_small.mat for those, so that should be correct.

2. I think the difference is because the interaction terms were switched off in the first example. There is a switch in "neelin_e.m" termed "intest=true/false" which defines if the interaction terms are estimated. I introduced this because I realised that in case that only one simulation is used to estimate the interaction term the prediction can get worse, as in this case from 0.87 to 0.83. I saw this for a few parameters and I have a theory why this occurs, I'll need more time however to explain you this, remains in my inbox then :)

3. Indeed validating high altitude regions can be problem. Even though we perform regional averages, the regions are quiet small and the topography of the model is much smoother than in the observations. The observations are, however, also estimated using spatial statistics (kriging, you can check the paper of Frei et al. 2013) so they are not point observations. But you could test this in checking if there is a relation between altitude and bias of the model.

Having a low sensitivity of the parameter is troubling though, what was the feedback at the seminar? Generally, I'm very curious to know how it went, I didn't hear anything sofar, what was your impression?

And let me know if you come up with more problems, I'll have very limited time but I'll try to help you. In about a month it will be better.

Cheers,
Omar

Pavel:

A. First of all, we have to be "objective", so instead of choosing the year 2008, it makes better sense to choose representing months. **Using Swiss observations database**, we can choose climatologically representative months out of recent several years. For example, the most representative January may be taken from 2013, the most representative February may be taken from 2009, and so on. We can save computer resources by choosing the representative month for **entire season**, and run our simulation over 4 month only. For example, the most representative January may be taken from 2013, meaning that its averages resemble the averages of all the winter months (Dec, Jan, Feb) during recent several years. This is the way people usually prepare so called "wind atlases" for wind energy projects.

B. However, (A) is still a "climatological" approach. Our goal is to improve forecasts and not the climatological averages (as for RCM's). When some "expert" tries to tune parameters, he chooses few case studies, when the forecast errors were large.

   Therefore I could suggest, again using Swiss database, to pick out (objectively!) 2 groups of days:

   1. Days when the original COSMO 7km errors where the largest.
   2. Days when the original COSMO 7km errors where of standard quality (mainly due to biases), may be using approach (A).

Then we could run our simulations on these two groups of days. The first one would represent our goal to improve the model in "difficult" situations, and the second one - not to spoil the climatological averages. Then, using the "CALMO score function" we can decide which weight to give to each of these groups. This approach would, from one hand, deal with the real forecast problems, and from the other hand save computational time.

C. Regarding the "CALMO score function", in my point of view, the most important (regarding temperature) is to evaluate min and max temperatures, and not the daily averages (2 temperature sinus functions with different amplitudes would give the same 24-h averages, giving an illusion that the forecast was good).

Points A and B of Pavel mail

- These are useful input for a second phase of the project where one tries to minimize the cost of the method (and to make the method less arbitrary with respect to the choice of the tested period)

Point C of Pavel mail

- As stated in my previous mail, I fully agree (in fact I thought that Tmin and Tmax were already used in the performance function)

Dear Jean-Marie and colleagues

Thanks for sending me the proposal, which I had not seen previously. In the attached version I have provided some additional inputs. There are a few issues which I suggest to consider:

(1) More recent work of Omar (see some added text in the proposal) has resulted in an additional very substantial improvement in the case of COSMO-CLM, beyond that described in the paper. The key parameter was the **hydraulic soil conductivity**. It is important that this parameter is also considered in your calibration. You are probably aware of this work and anyway considered to do so (I just mention this to make sure).

(2) A spcific suggestion: Apply the same calibration also to an 2.2 km version that should otherwise be as close as feasible to the to-be-calibrated version of COSMO-1. This would yield an objective inter-comparison between the two model versions, and provide an objective assessment of the added value of COSMO-1. In computational terms the COSMO-2 calibration would be much cheaper, and it is also be well suited as a start-up excercise. Also, if you intend to publish this project (which I strongly encourage to do), the intercomparison of two model versions using calibration standards should make publication rather easy (I believe it would be the first study of its type, at least for a full model).

(3) Some info might be needed regarding the performance score, as the one used by Omar is not appropriate for your purpose. The importance of the score is evident and will strongly affect the results. You will not need to provide a full description of the score in the proposal, but some indication about the directions you intend to take. Make sure the score is geographically balanced, **use the uncertainty of different variables** (i.e. T, P, etc) to weight the different contributions, and use only those observations that are consistent with the modeling framework and resolution (e.g. no point observations at extreme mountain or valley stations). If you would like to document the improvement of COSMO-1 at stations like Saentis or Jungfrau, do this offline but not as part of the calibration.

(4) I find this a very interesting proposal and suspect independent reviewers will do so as well. However, I definitively should not be a co-PI.

I have added Omar to the distribution list (although he will likely be too busy to react quickly).

Cheers from Washington DC
Christoph

Dear Christoph Frei,

As you know, we are working on calibrating the COSMO model parameters, comparing 2m-temperature forecasts (in 7km-resolution) with the observations.

Since your previous email, I was thinking of the optimal way to adjust the observed 2m-temperature, because of elevation differences between model and observation grid points. I hope to find a better approach, rather than using moist adiabat.

I have carefully learned your paper describing the procedure you used for creating the 2m-temperature observations grid over Switzerland (http://onlinelibrary.wiley.com/doi/10.1002/joc.3786/abstract).

It is very important for me to hear your opinion regarding the following 3 questions, related to your paper:

1. As part of your interpolation procedure you obtained, in addition to the 2m-temperature $T_{2km}(x)$ at 2km-grid point x (located at elevation $h_{2km}(x)$), **also the background temperature profile $T_{bgd}(z,x)$ at every grid point x at every day**. When trying to guess, what would be the 2m-temperature $T_{7km}(x)$ at 7km-grid point x (located at elevation $h_{7km}(x)$), I can fix $T_{2km}(x)$ to obtain $T_{7km}(x)$ **following the profile $T_{bgd}(z,x)$**! In other words the correction will be: $T_{bgd}(h_{7km}(x)) - T_{bgd}(h_{2km}(x))$. In my point of view, this correction is much better than following a simple (and usually wrong) moist adiabat.

   Do you think it is reasonable?

2. There is a discrepancy between the 2km-resolution grid elevation and the real elevation of the stations. I don't fully understand how do you overcome the problem of interpolating observations from a real elevation to the 2km-resolution grid elevation?

   When you calculate the background temperature at a given grid point x, you probably pick up the temperature $T_{bgd}(h(x))$ from the background temperatureprofile $T_{bgd}(z,x)$ at an exact elevation of the grid point z=h(x).

   However, when you calculate the residual $T_r(x)$ at a grid point x, you calculate the generalized distance $D\lambda(x_{obs},x)$, but I don't understand which elevation do you use at x_{obs} (the real one or from the 2km-resolution grid), and at x (the real one or from the 2km-resolution grid)? If you use the elevation from the 2km-resolution grid in both cases, do you correct the temperature observation at x_{obs} for the elevation difference?

3. As I understand, the up-to-date 2m-temperature observations data set on a grid is now available only for daily means, but not for T_{max} or T_{min}. In your paper, you mention the work of Hasenauer et al. (2003) for interpolation of T_{max} and T_{min} (DAYMET approach), but it was performed on part of Switzerland only and it is not up-to-date. Is this really the situation so that there is no data set of measured T_{max} and T_{min} on a grid over Switzerland for the year of 2008?

With great appreciation,

Pavel

Hi Pavel,

Thank you for your interest in temperature interpolation. Here are quick answers on your questions below:

3. Daily Tmin and Tmax are available as gridded fields over Switzerland too. (Product names are TminD and TmaxD.) You can get access to these results - Guy, Jean-Marie and Oliver know how. They were determined with a slightly modified version of the published algorithm that ensures Tmax-Tmin is positive everywhere.

2. The two steps of the procedure deal with real temperatures (background) and anomalies from the background (residuals). The residuals at the stations are determined with respect to the background at the height of the stations.

1. In principle the use of the background field for correction would be preferable over a constant lapse rate. However I expect the gain in accuracy to be relatively small, considering that there is interpolation uncertainty on top. We do not store the background profiles, they would have to be recalculated. An alternative would be to estimate local linear temperature profiles and correct with those rather than with a pre-specified lapse rate. Eg. estimate a 7-km grid point value by linear regression fitting from 2-km grid observation grid points in the surrounding.

Kind regards,

Christoph

Dear Pavel,

the original 2 km resolution Dataset is as you say the file "TabsD...". The 7 km file is a linear interpolation of the 2 km grid without the consideration of the topography. We discussed the uncertainty we introduce this way at the beginning of the project and assumed that it would play a minor role when computing spatial averages, but it is surely a good thing to use a more reasonable interpolation considering the results you have sent.

Cheers,
Omar

On 30 April 2014 12:23, Pavel Khain <pavelkh_il@yahoo.com> wrote:
Dear Antigoni and Omar,

I got confused with the observation files:

"t2m_meteoswiss_2008_c7mch_daily.nc" is the temperature observations file you used, but it has 7 km resolution. How this file was created? What are the values in its grid points? Are they interpolated from the original 2km resolution file?
There is another observations file: "TabsD_ch02.lonlat_200801010000_200812310000.nc" which is located in: "/workspace/omarb/calmo/observations/t2mmch/Tabs", which has 2km-resolution, and, as I understand, it is the original Christoph Frei's result.

Best regards,
Pavel

Dear Pavel,

some considerations of the results you've sent.

Using a height correction for the interpolation of the observations will surely improve the magnitude of biases, overall they seem however to be very large in both directions. I still wonder if the level of these biases is realistic for this type of setting, perhaps Francis Schubiger can comment on this.

Regarding the precipitation part, it seems also to me that the fields are very noisy which explains also that there is no skill in reproducing the experiments with the interpolation. Even in my application using larger spatial averages and monthly means the precipitation remained the hardest variable to predict due to the level of noise. A smoother statistic would help yet considering that it should remain a verification of an NWP application.

The fact that you get boundary solutions for the optimal parameters is typical when only one variable is considered which can also be seen in Neelin 2010. From experience adding more variables will constrain the solution inside the parameter space due to competing biases, however as long there is no skill in predicting the precipitation errors there is no point adding these to the optimization**.**

Hope this is help, best regards,
Omar

Dear all,

Its encouraging to hear that the project is reaching new dimensions. As Christoph said we've been working on an extension of the calibration for COSMO-CLM with more parameters, a first draft will be available soon.

Small detail with this respect, not 2 but 3 additional parameters are being considered in this analysis. These parameters are tkhmin and rad_fac (new parameter) which have already been in the discussion in CALMO and a factor for the hydraulic conductivity which strongly affects the summer climate.

Other than that I have not much to add. A comparison with the 2.2 km version would be interesting and probably favorable to demonstrate the added value of COSMO-1 when unmasking tuning efforts that have gone into COSMO-2. I think there is much potential in this direction and I'm course happy to see the methodology being applied to this new high-end GPU COSMO version!

Good luck with the proposal and best regards from Barcelona,
Omar

Dear Pavel,

I see I forgot to introduce "nl" in the rest of the Code, it was an Idea I implemented at one point but still not very happy with it. I need to explain a bit more:
In some applications I realized that considering the interaction terms actually worsened the accuracy of the metamodel, quiet dramatically in some cases. This also occurred even if the interaction terms were very small and I'm still not sure if I understand this correctly. Unfortunately I never had time to test this properly. This occurs only if only one simulation is used to determine the interaction terms, so I eventually performed more than one for each parameter combination (as in Neelin 2010).
I created the calmo_toymodel.m (which remains under development) to find out what caused for this behavior, but I'm still not clear if I got the correct answer. So I think there are two hypothesis why this happens:
(1) The interaction of the parameters is important but not well captured by the term b_xy*p_x*p_y
(2) The interaction is very small but large interaction terms are estimated due to large noise of the data.
(In fact (3) would be: there is a bug in the code, but this is what I checked most, but nevertheless worth thinking it through by someone else)
I mainly followed second Idea. The data of the simulations are noisy because of the internal variability of the RCM. This level of internal variability corresponds to the noise level ("nl") that I considered in the code. In case of CALMO we actually never tested what the sensitivity to the perturbations of the intial coniditions i.e. what the value of "nl" is for each data point.
In the current code there is no proper consideration of this noise in the regression of the metamodel. This can be problematic in particular for the interaction terms since before estimating these several differences of simulations are computed - which increases the noise level.
Say we consider a simulation as a time series sim with mean value y and Gaussian white noise N(0,sig), where sig is the standard deviation of noise level (nl aka internal variability)
(1) sim=y+N(0,sig^2).
The estimation of the linear and quadratic terms is performed on the difference of the axial simulation and the reference simulation (difference of parameter, say dsim_p)
(2) dsim_p = sim_p-sim_ref = y_p-y_ref+N(0,1.4*sig)
Difference of this noisy data increases the variability by the sum of the variances (2*sig^2)=1.4*sig , so the difference (dsim_p) is actually more noisy than the original simulation. To estimate the interaction multiple differences are computed,
(3) dint_p1p2=int_p1p2-sim_ref-dsim_p1-dsim_p2 + N(0,12*sig)
which increases the noise level by a factor of 12! Not 5 in fact realize now, so this fact is vastly more important than I thought previously and that's a challenge. This explains why we need multiple simulations to determine a robust signal for the interaction term...
Unfortunately I'm unable to work further on this at the moment, but feel free to develop further. I guess there are two options, either estimate the metamodel differently or considering interaction terms only if they surpass the noise level of a factor of 12 if only one simulation is considered.
Hope this helps, feel free to ask again if this is not clear!
Saludos,
Omar

On 20 May 2014 13:04, Pavel Khain <pavelkh_il@yahoo.com> wrote:

Dear Omar,
How are you? Hope everything is fine.
I am trying to understand the changes you recently introduced in the calibration code, and particularly in the file "**neelin_e.m**". I can see, that you deal with the question, whether to use the interaction term or not.
You put a comment there:
"If interaction signal smaller than noise level, no fit to it. Since interaction term considers several differences of noisy data total noise level is 5 times the noise level".
Can you please explain me, what is going on there? (what is this noise level? what is the parameter "nl", which is not defined anywhere? More general - in which cases would you recommend to ignore the interaction term?).
Thank you,
Pavel

Dear Omar,

Thank you very much for the detailed explanation. Few more questions:

1. Can you please comment - how did you obtain the number "12" in formula (3) below?
Somehow I get 2*sig^2, and not 12*sig^2:
dint_p1p2=
=int_p1p2-sim_ref-dsim_p1-dsim_p2=
=int_p1p2-sim_ref-(sim_p1-sim_ref)-(sim_p2-sim_ref)=
=int_p1p2-sim_p1-sim_p2+sim_ref
so we have 4 terms, where each one has a noise of sig^2, so the total is 4*sig^2, which means an equivalent noise of 2*sig. Where is my mistake?

2. Following your explanation, it is better to run more interaction terms (instead of one "corner", we can run 2,3 or even 4 "corners" for each pair of parameters). How many, in your opinion?

3. How would you suggest calculating the internal variability in CALMO? We can't shift the initialization time, as in the CLM case. I am trying to think, what could we disturb in REF simulation for calculating the internal variability?

Thank you very much,
Pavel

Dear Pavel,

here some answer to your questions

(1) as you write the equation for dint_p1p2 the terms dsim_p1 and dsim_p2 have a variance of 2*sig^2 whereas the other two have a variance of sig^2, so the total variance of the term dint_p1p2 would be 6*sig^2 or sqrt(6)*sig

However, the noise of the different terms is no longer independent since in the terms dsim_p1/2 we subtract the reference with the same noise. If you now subtract both terms the variance is no longer additive since it is correlated. I'm unable though to write the equation right now, in fact I realized in the last mail that the total increase in standard deviation is 12 because I tested it with Matlab with Gaussian noise :) And I realized that you get a factor of 12 times the standard deviation, so increase of variance of around 48. But I didn't use much time testing it, perhaps would be useful if you do the same test..

(2) **I would consider to use 2-3 interaction simulations if there is indication that the interaction seems to be large, otherwise omit the interaction**. With such a high noise level this is, however, hard to test. What I tried is to consider the interaction only if the term dint_p1p2 if the signal is larger than the noise. Since it seems to be 12 times the internal variability this would rule out most of the terms, not sure if that's the best way to go.

(3) To test the internal variability in CALMO I would consider a random perturbation, not sure if COSMO has something like singular vectors for perturbations. I'm not sure about that, since I have little experience with NWP. We discussed this issue with Jean-Marie in the beginning of the project and he said that the internal variability would be very small since lead time is short, but it might be important considering this inflation of noise for the interaction terms.

Hope this of help! Best regards,
Omar

Dear Pavel,
COSMO, as all co-operations is based on mutual trust, so we do not need evidence.
However I have to admit that the brochure is very enlightening!!
We should have a web conference to discuss about the next steps of CALMO, and mainly about the progress of the work with the meta-model and the achievements gained so far. More specifically, I would also like to discuss about the report you have send us, your suggestion and question. In addition any feedback on the tasks proposed for you to undertake is welcomed.
I would also appreciate it if you could upload the modified version of the meta-model, as it is one of the main deliverables for CALMO.
Regarding the LA_RING directory it contains laf* files generated with data assimilation on lnudge=.true and then these laf* are used for all simulations.

So assuming that Jean-Marie is on-line on Thusday 0800UTC we could have a web conference then with Jean-Marie, and arrange a new web next week Tuesday or Thursday 0800 UTC with Pavel.

Best regards
Antigoni

----- Original Message -----**From:** Pavel Khain
To: Voudouri Antigoni
Sent: Friday, May 30, 2014 9:44 AM
Subject: Re: Web. Conf. and FTE's

Dear Antigoni,

Thank you for the detailed answer regarding the data assimilation. **So just to make it clear to myself - do you use assimilation cycle at LA_RING in order to create the ANALYSES for initializing ALL the other simulations?**

Regarding the web.conf. - on Thursday we an Open day in our IMS, and I have duties there, so I will be free starting from 11 UTC.
(I attach the brochure, just in case you don't believe me :)).

Basically I am at home now (Friday, 6:40 UTC), and can simply call you by phone to your office at any time suitable for you.

Best regards,

Pavel

Dear Antigoni, dear Pavel,

Some miscellaneous remarks after having read the many mails exchanged during my holidays (-;

- Great work on the temperature scores! I totally support the use of T2min and T2max in the quality score.

- The insensitivity of the precipitation with respect to rlam_heat and tkhmin could be due to the type of precipitation at this time of the year (mainly frontal systems in winter/spring); I expect more sensitivity for convective precipitation.

- **One sees that the definition of the quality score is a difficult problem. In my opinion, one should add a third type of gridded observations, representative of the PBL and catching the daily cycle; maybe GPS derived integrated water vapor could be an option (although I do not know if gridded values are available).**

- I have read the Daint proposal, and this is a very good document (if not a little bit over-optimistic about the potential of the method).
- I need an updated PP plan till June 12th (for the SMC members). I will send my feedback to Antigoni in another mail.
- 15' CALMO talk is foreseen in the COSMO GM plenary session (Tuesday afternoon).
- I will soon prepare a draft for the COSMO GM parallel session.

Friendly greetings,
Jean-Marie

Dear Antigoni,

Generally I agree with the plan presented in the proposal. However, I have few comments regarding my part in CALMO:

Following my short experience working with CALMO, I spend a lot of time further developing the meta-model code, adding new scripts (and sometimes improving current ones). For example, during the last 3 month I have solved bugs related to segmentation of Switzerland to regions, added scripts which allow constructing various error distribution functions (which are important for visualizing the sensibility of the model to the parameters changes), and this work will definitely continue. **For example, in the near future I have to deal (according Omar's recommendation) also with the correct way of implementation of the interaction term into the meta-model, which is a difficult task**. Therefore I suggest adding in subtask "1.3 **Technical Infrastructure"** (or at any other place you decide), something like: further development of the Neelin meta-model, and add there about 0.05-0.1 for myself.

In addition, during the last 3 month, I have performed a successful work on proper implementation of the temperature observations. I think that such type of work will be needed in the future as well (with temperature when we will deal with 2 and 1 km resolutions, and even before that, with other COSMO fields, as Tmin, Tmax, and so on). Therefore I suggest adding in subtask "1.3 **Technical Infrastructure"** (or at any other place you decide), something like: Proper adaptation of the observation gridded data for the meta-model implementation, and add there at least 0.05 for myself.

In summary, I think that further investigation is expected with the meta-model, that I am happy to perform. I think that this work has to be declared in the PP plan, and at least 0.1-0.15 FTE more have to be intended for myself for that purpose.

Jean-François Mahfouf:
"**Address the uncertainties associated with the look-up tables, especially for the SVAT model**":
This aspect is particularly important; objective parameter optimization should be considered.

Answer(J. Helmert)): The automatic parameter optimization procedure for tuning of look-up tables could be integrated into the COSMO project PP CALMO.
... could one integrate in some way an objective calibration of some of the look-up tables used in the derivation of the **external parameters for the soil and surface**? This is worth considering ...

Cheers,
Jean-Marie

Dear Pavel,

The documentation of TERRA: COSMO doc. (http://www.cosmo-model.org/content/model/documentation/core/default.htm#p2) , part II, chap. 11

The documentation of external parameters and look-up tables: COSMO doc. Part II, chap 14 (in particular table 14.3), or http://www.cosmo-model.org/content/model/modules/externalParams/default.htm

Cheers,
Jean-Marie

From: Pavel Khain [mailto:pavelkh_il@yahoo.com]
Sent: Donnerstag, 3. Juli 2014 19:45
To: Bettems Jean-Marie
Cc: Voudouri Antigoni
Subject: Re: [cosmo-calmo] An interesting input from Jean-François (in relation with COSMO SP)

Dear Jean-Marie,

Could you please provide me with more information regarding SVAT (Soil-Vegetation-Atmosphere) model and the look-up tables mentioned in your email below.

Thank you,

Pavel

Dear Antigoni,

Please find in attachment my proposal for a consolidated PP plan for the CALMO extension.
Of course, feel free to make further modifications.
Once you are happy with the status of this document, please send it to Michal with CC: to me.

The main changes I made:

- I removed the MeteoSwiss resources)-; As I told you, I spoke with Philippe, and he stated again that our involvement is limited to 'coaching' the project team. But this is not so bad.... We (myself and my colleagues) are ready to help you whenever needed... so do not hesitate to ask! Furthermore, we will share (most of) the computing environment on Daint (and Tödi).
- I modified the risk assessment chapter, taking into consideration the new possibility to have access to Tödi.
- I tried to emphasize that the goal of task 3 is not tuning COSMO-1, but showing that the method is able to improve the quality of 'non-standard' model configurations.
- Otherwise I tried to improve the clarity of the text.

Some miscellaneous points, in particular to summarize my previous mails

- Please follow the advice of André to improve the turnover of experiments on lema (adapt time limit, launch more jobs in parallel).
- With a little bit of luck, you should get the gridded T2m min / max next week.
- **I still think it is important to find a third type of observations, possibly related to BL humidity, in the definition of the quality measure**. Any progress on this topic?
- **A parameter describing the hydraulic conductivity of the soil should be introduced in the set of tunable parameters**.
- **A method to represent the uncertainties introduced by the external parameter look-up tables for the vegetation should be evaluated**.
- **A strategy for the initialization of the soil should be fixed.**
- Have you received any feedback from the CSCS? In principle a technical feedback from the CSCS was due till the end of June, and a scientific feedback till today.
- Even if the Daint proposal is rejected, or if the allocated resources are significantly reduced, Tödi offers a reasonable alternative.
- Next web conference is planned on Tuesday 15th, 8h UTC
- Antigoni will chair a CALMO parallel session at the GM, on Monday 14h - 16h.
- Antigoni gives a CALMO talk in Plenum, on Tuesday 15h30 - 15h45.
- A small workshop is envisaged at Athens after the GM (Thursday 12h - Friday 12h).
- It is important to be ready to start the 'intensive experimenting period' as soon as Daint will be available (if the project is accepted): in particular the choice of the model parameters, and the strategy for the soil initialization.

All the best,
Jean-Marie

Dear colleagues,

In the frame of a COSMO priority project, one tries to develop a method for automatic re-calibration of tuning parameters in NWP models.
For that purpose one needs gridded observations to automatically assess the quality of the model.
We are already using T2m analysis from Ch. Frei, and radar derived precipitation, but we would like to complement these data with a third type of observations.
Integrated water vapor is an interesting candidate, being essentially representative of the PBL and having a high variability.

So my questions:

- Do you know if gridded fields of observed **GPS ZTD** over central Europe exist?
- If this is the case, what is the quality and the temporal resolution of these data?
- If this is the case, do you know if these data are available for the year 2008?
- If this is the case, do you know who to contact to access these data?

Thanks for your time,
Jean-Marie

Dear Antigoni,
dear Pavel,

Some information I have collected since our web conference:

- We indeed have seen some sensitivity of precipitation (convection) and T2m over Switzerland with respect to the turbulence length scale, in summer, with a typical impact on T2m of about 0.3K at mid-day.
- We also have gridded sun shine duration and some satellite products at MeteoSwiss; besides that we have some high resolution re-analysis. If you need some of these I can ask for more details.
- We have no experience with the use of satellite derived radiances for model validation; these observations could be tricky to use because radiances may be very dependent on soil properties
- Daniel also share the opinion that integrated water vapor is a good candidate for the CALMO quality score.
- Integrated water vapor can be produced in model output, in form of ZTD (Zenith Total Delay).
- **There is a European action E-GVAP to collect GPS derived integrated water vapor**; web site of E-GVAP program is http://egvap.dmi.dk/

Cheers,
Jean-Marie

3 Derivation of zenith delays from meteorological data

The ZTD is the integral of the refractivity over a vertical column of the neutral atmosphere, commonly written as

$$ZTD = 10^{-6} \int_{z_{site}}^{z_{top}} \{k_1 R_d \rho_d + (k_2 + k_3/T) R_w \rho_w\} \delta z, \quad (1)$$

(see e.g., Bevis et al. (1994)). Here ρ means density, z geometric height, R gas constant, whereas subscripts d and w mean dry and wet. The k's are empirically determined constants quantifying the refractivity of the lower neutral atmosphere for radio-waves in the GPS range, they may contribute to the uncertainty of derived ZTD's on the millimeter level. See Bevis et al. (1994) for estimates of the k's. In this article we shall be using $k_1 = 7.76\ 10^{-1}$ K/Pa , $k_2 = 7.04\ 10^{-1}$ K/Pa and $k_3 = 3.739\ 10^3$ K^2/Pa when deriving delays from meteorological data. Further one should

Dear Dominique,

I don't know about an existing gridded ZTD field. But I know that people have considered to make such fields, for example to experiment with augmentation of ZTDs for usage in certain GNSS data processing and for processing of INSAR data. I'll forward your mail to a wider of people working in the field of GNSS meteorology.

Having said that, I personally would prefer to validate directly against the ZTDs from actual GNSS sites, rather than using a gridded product.

Depending on type of terrain and inter GNSS site distances the quality of the gridded product will vary a lot from region to region. Using the actual ZTDs reduces the risk that variations you might see with respect to COSMO are due to this inhomogeneity, not due to COSMO. The longer you go back in time, the more important this becomes, since the GNSS ZTD coverage was much more patchy in the past than now. And still now it varies a lot.

Kind regards, Henrik

On 2014-07-15 13:37, Dominique.Ruffieux@meteoswiss.ch wrote:
> Dear Henrik,
>
> My MeteoSwiss colleagues from the NWP group would be very interested
> to use GPS info at the European scale for COSMO verification. Do you
> have an answer to the following questions:
>
> ØDo you know if gridded fields of observed GPS ZTD over central Europe
> exist?
>
> ØIf this is the case, what is the quality and the temporal resolution
> of these data?
>
> ØIf this is the case, do you know if these data are available for the
> year 2008?
>
> ØIf this is the case, do you know who to contact to access these data?
>
> Thank you very much for your help!
>
> With my best regards
>
> Dominique
>

Hi Henrik and Dominique,

Our NWP model produces ZTD fields at 4km over Europe and ~17km globally, with forecasts at hourly intervals out to 12 hours, updated every 6 hours. This is a fairly recent thing, so we don't have the fields for 2008. In terms of quality, I haven't had time to assess them yet, but they will be subject to biases compared to real ZTDs from GNSS receivers, as there will be a mis-match between the model surface and the real surface. This might be a fairly constant bias which could be corrected for, perhaps by adding some additional hydrostatic delay to the ZTD values.

If you are interested in these fields, I plan to start uploading them to the ftp site here in the next month or two. They would probably be in netcdf format, as that is the simplest conversion I can do at the moment from our own ridiculous internal field formats.

Best Wishes,
Gemma

Dear Dominique

I support Henriks advice to use real GPS stations instead of gridded information.
Might be the easiest for you and your colleagues is to follow a similar strategy as the NWP colleagues do already now (might be with our GNSS ZTD estimates which of course does not cover all Europe): using the COST files from the EGVAP server for which also a combined products should exist in 2008 (makes it easier and more independent of the contributing analysis center).

A second option is to use e.g. troposphere estimates of a global reprocessing (but attention: this series are generated in postprocessing and will have better quality and consistency over longer periods than estimating ZTDs in near-realtime; station availability is global and therefore sparse in Europe). Which density you are looking for? Very soon, also European repro results will become available.

Greetings
Elmar

Dear Dominique,

Yes we have Gridded GNSS ZTD data over Europe.

Based on Ordinary Kriging interpolation, we developed a GNSS tropo grid creator. Very briefly, it estimates ZTD corrections taking into account the residuals between GNSS-derived (the input is a GNSS ZTD solution in cost format) and model-computed ZTD at continuously operating GNSS stations. Then at a known user location, the correction to add to the modeled-ZTD value, is obtained through a bi-linear interpolation with the four nearest grid points surrounding it. Details in the attachment which has been accepted for publication in the IAG Symposia Series.
We have tested it in an European project and are willing to do other tests.

It is running on a routine basis but **unfortunately we don't have any gridded data for 2008**. What is your period of interest?

Best Regards
Rosa

-----Messaggio originale-----
Da: Henrik Vedel [mailto:hev@dmi.dk]
Inviato: venerdì 18 luglio 2014 10.56
A: egvap_processing_wg@dmi.dk; egvap_user_wg
Cc: dominique.ruffieux@meteoswiss.ch
Oggetto: Gridded ZTD data?

Dear colleagues,

I have got the attached mail from Dominique at MeteoSwiss.

I know that some of you have made proposals in which the use of gridded ZTD fields was a part. Have even been involved in such proposals myself, but they didn't obtain funding.

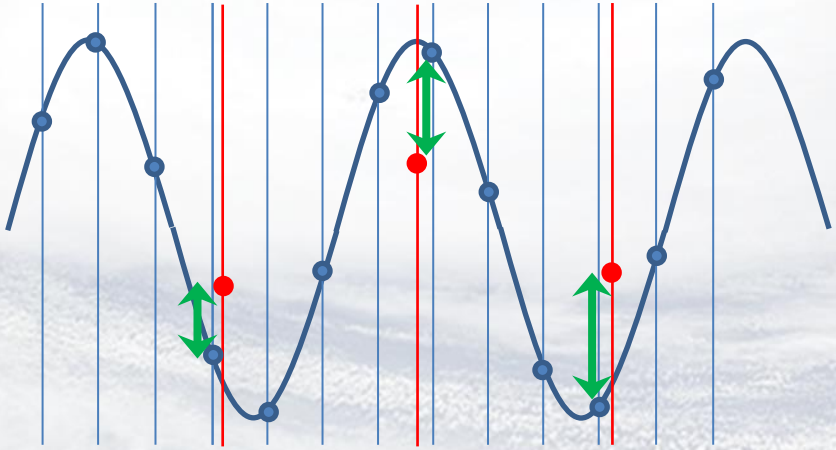
Do some of you have experience with or access to gridded ZTD fields that might help Dominique and his colleagues?

Thanks and regards,
Henrik

Built-in temperature biases in the areas of complex topography

- Observations grid (Christoph Frei) – resolution 2 km + real (not smoothed) topography
- Simulations grid – resolution 7km + smoothed topography
- Previously: temperature observations simply interpolated from 2km to 7km grid. Bad…

- 2m-temperature very much depends on height. Usually: higher = colder
- If the grid point is too low → temperature too high
- If the grid point is too high→ temperature too low



- Must adjust the observations to the model grid **before** calibration! But … ***How?***

- First Cristoph Frei's e-mail:

  "The most frequently used is the 6.5 degree per km moist adiabat. But constant lapse-rate corrections are of limited representatively sometimes. From knowing the elevation differences you can estimate a typical error coming from the correction by assuming the truth is somewhere between the moist and the dry adiabat".

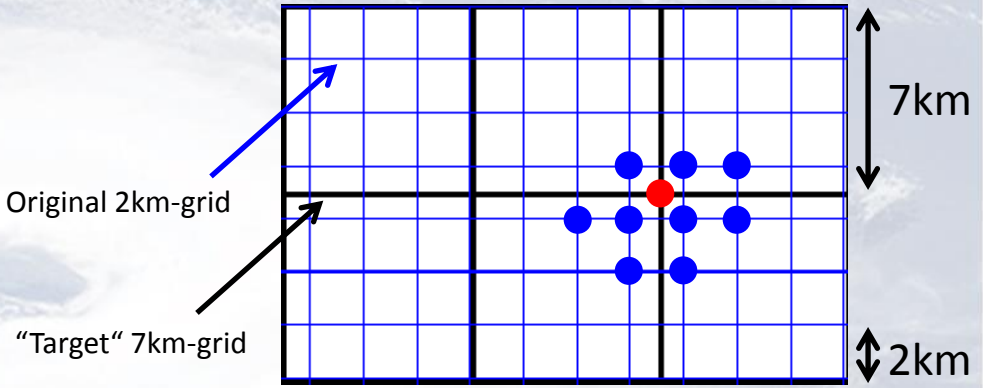
2m temperature adjustment to the (smoothed) model grid

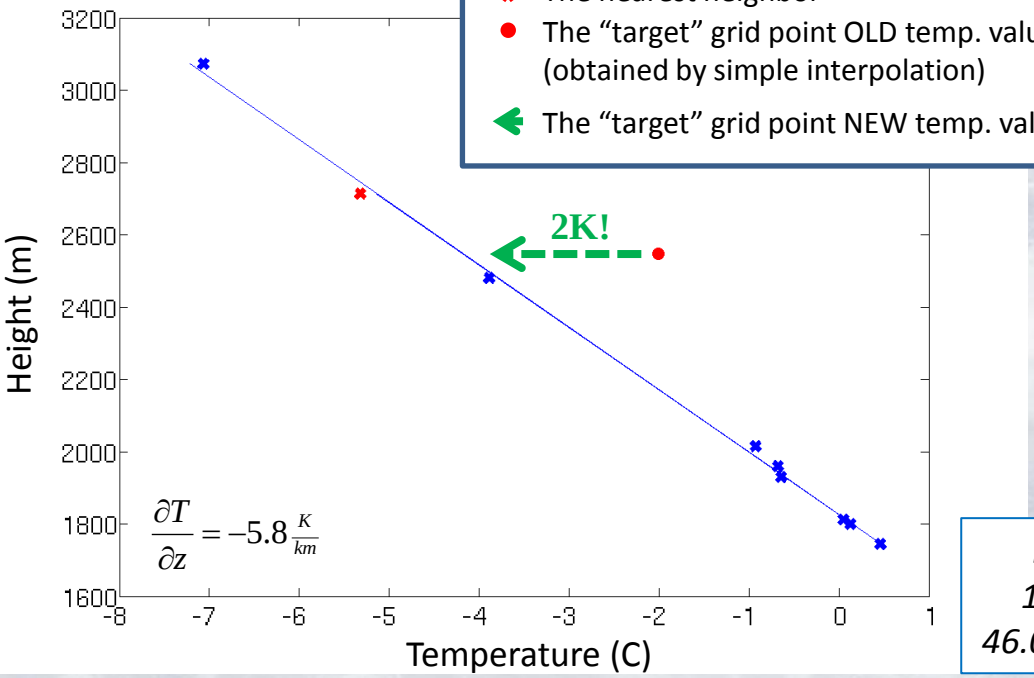
- Last Cristoph Frei's e-mail – much better idea !

1. For every grid point in the "target" grid (red dot), find the nearest 9 neighbors on the original 2km-grid (blue dots)
2. Plot the 2m-temperature values of these neighbors vs. their altitude (blue and red "x").
3. Perform a linear fit of the data, which will be the ***local 2m-temperature profile***
4. Having the altitude of the "target" grid point, use the linear regression, to calculate its 2m-temperature
5. Perform this operation (1-4) for every "target" grid point, for every day

Local profile:

works even for inversions!





Example:
10/1/2008
46.07 N - 7.49 E