



# Summary of the evaluation of COSMO and ICON model products at the ESSL Testbed 2015

ESSL Report 2015-02

Author: Pieter Groenemeijer

European Severe Storms Laboratory – Science and Training

Bräunlichgasse 6a / 6

2700 Wiener Neustadt, Austria

Tel: +49 151 59031839

Tel: +49 8153 281845

Fax: +49 8153 96599911

e-mail: [testbed@essl.org](mailto:testbed@essl.org)

# 1. Introduction

## **This summary**

This report is a summary of the main findings of the ESSL Testbed 2015 regarding the COSMO and ICON products. It is not exhaustive but an attempt has been made to condense the feedback from discussions at the Testbed and the questionnaires, and to illustrate them with a number of Testbed maps. Chapters 2 and 3 are the core of this summary. Chapter 4 is a collection of Testbed model output maps for a number of cases, to which reference is made.

## **Testbed Participants**

The ESSL Testbed 2015 took place from 25 May to 26 June 2015 at the ESSL Research and Training Centre in Wiener Neustadt, with a break from 8 to 12 of June. The Testbed was visited by four DWD R&D employees (Susanne Theis, Paul James, Thomas Hengstebeck, Kathrin Wapler), who gave presentations on site as part of their participation and, additionally, provided several online presentations. From MeteoSwiss, researcher and developer Jacques Ambühl attended in addition to three forecasters, Lionel Peyraud, Sabrina Lang and Daniel Cattani. Five DWD forecasters participated in the Testbed activities (Berndt Zeuschner, Robert Herrmann, Helge Tuschy, Thomas Schumann, Martin Jonas, Kathrin Hohmann). ESSL is very thankful for the contributions in form of presentations, contributions to discussions by all of these persons! In total, 41 persons from Europe and abroad took part in the Testbed. These persons originated from Germany, Switzerland, Austria, the Netherlands, Croatia, Poland, Greece, Romania, Portugal, Slovakia, the Czech Republic, Latvia, France, the U.S.A, Finland and Lithuania.

## **Evaluation of COSMO and ICON model products**

The focus of the evaluation activities for these NWP products were focused on the following points.

### COSMO-DE-EPS

- General assessment of model performance
- Probabilistic products for extreme precipitation, using the local, upscaled and fuzzy methods\*
- Probabilistic products for graupel and wind gusts, using the local, upscaled and fuzzy methods
- Probabilities of exceedance of CAPE and CIN\*

### COSMO-DE

- Study of vorticity and SDI within storms to identify supercells, and comparison with rotation (track) radar products\*

### ICON-EU

- General assessment of model performance, biases, comparison with COSMO-EU

\* these items were originally scheduled for the Testbed in 2014 and have been postponed until 2015

#### COSMO-E

- General assessment of model performance, with particular attention to the question how forecasters can distill occurring physical processes from the ensemble data; comparison with COSMO-DE-EPS
- Probabilistic products for extreme precipitation, using the local, upscaled and fuzzy methods
- Probabilistic products for graupel and wind gusts, using the local, upscaled and fuzzy methods
- Probabilities of exceedance of CAPE and CIN

#### COSMO-1

- General assessment of model performance, with particular attention to the added value of the very high resolution, as compared to the lower-resolution COSMO-E ensemble forecasts, as a function on forecast lead time; comparison with COSMO-DE
- Study of vorticity and SDI within storms to identify supercells and comparison with rotation (track) radar products

### **Testbed Resources**

The following online resources can be used to obtain more detailed information about the Testbed 2014 that is contained in this summary document:

The ***Testbed Data Interface*** showing all products and all data, is available online after the end of the Testbed at: <http://80.109.154.60/testbed/nowcast2015.php>

A ***Blog*** describing the daily activities at the Testbed can be found at:

<http://www.essl.org/testbed/blog>

***Background information*** and ***all presentations*** given at the Testbed can be accessed at:

<http://www.essl.org/testbed/info> password: tornado15

### **Feedback**

Feedback on the products was collected throughout the Testbed, partly in **direct discussion with the DWD R&D participants**, partly through the **documentation of answers to questionnaires** that was filled out jointly by the participants (contained in full as an Appendix to this document).

## 2. Description and evaluation results of probabilistic products of extreme precipitation, graupel and wind gusts, CAPE and CIN

In 2012 - 2014, ESSL has already evaluated several visualization of high-resolution ensemble data. This year, so-called 'fuzzy' and 'upscaled' probabilities were implemented using the methods described by Ben Bouallègue and Theis (2013).

The fuzzy method is a smoothing method that enlarges the ensemble sample size by including the forecasts in neighboring grid points, and the upscaled method modifies the reference area of the probabilities. Forecaster feedback was collected on each of these products, which were compared with the traditional pointwise or local probabilities. An example of these visualizations is given in Fig. 1.

The reflectivity field was the default field that appears when selecting the 'ensemble' model page. For this field and other fields that are characterized by local extreme maxima such as reflectivity, precipitation, integrated graupel, SDI or wind gusts (in convective situations), some visualizations were clearly favored above others:

- 1) **Of all visualizations, the 'upscaled probabilities' were very favorably received by the participants** along with the 'all members exceeding threshold' product. In addition the 'fraction of members exceeding threshold' and the 'maximum of all members' ("sometimes a little messy") were found to be somewhat useful.
- 2) **The smoothed probabilities were not very useful for fields like reflectivity, precipitation or SDI.** The smoothed probabilities were much criticized for smoothing out the signal to below the lowest threshold. Naturally, this was partly caused by the setting of the value of that lowest threshold to 10%, which a participant commented is "*just throwing out information*". Setting the threshold to much lower values would, however, make them hard to interpret. E.g. it would be somewhat complicated to understand what a 2% or a 8% probability of rainfall means.

This is where the **benefit of the upscaled probability** comes in by enlarging the area of reference to a 40 km square. A 2% probability of occurrence of an extreme event at a particular location would then, for instance, change into a 40% occurrence of the event within a 40 x 40 km box near that location, a concept could work with quite intuitively.

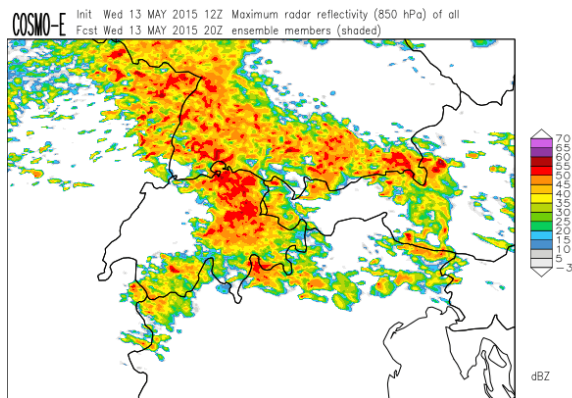
The example in Fig. 1 shows a COSMO-DE-EPS forecast that denotes a zone stretching west-to-east across Baden-Württemberg and Bavaria in which rotating updrafts (i.e. high SDI) are possible. Because of the varying locations of the rotating storms within the zone, the fraction of members that exceed the threshold is low at any given location. The smoothing in the 'fuzzy' product even reduces the map to an empty one. The signal is, however, clearly visible in the upscaled probability image (left bottom). The radar at 2000 UTC (bottom right) shows several strong thunderstorms, one that has just produced a significant F2 tornado (labelled A) and one that is about to produce an F3 tornado at 2030 UTC (labelled B).

The signals in this case, and the case of tornadoes on 5 May do not stand out as clearly as one would perhaps want, so a recalibration of the thresholds value and/or the size of the area of reference for SDI

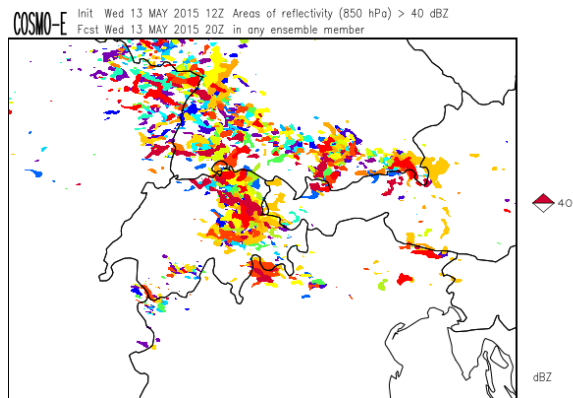
is probably desired. The supercell detection index is sometimes “triggered” by non-convective flow that is deflected by orography. One may want to explore the use of **updraft helicity**, a very well researched proxy to severe weather in the USA, which perhaps does not suffer from this effect.

3) **For smoother fields like CAPE or CIN, the upscaled product is not evidently better** than products like the ‘fraction of members exceeding the threshold’ or the smoothed variety. Smoothed probability may have the highest value.

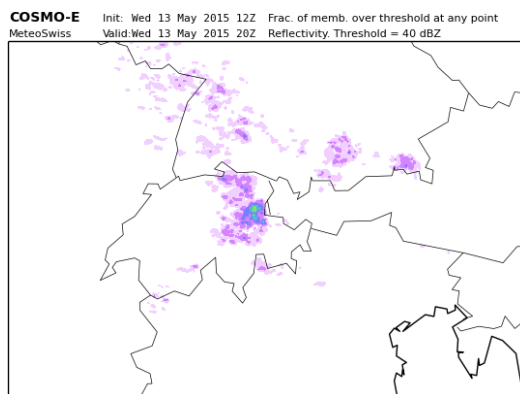
A participant remarked that “*more thresholds are needed*”, since the present implementation used only one threshold for CAPE (500 J/kg) and CIN (50 J/kg).



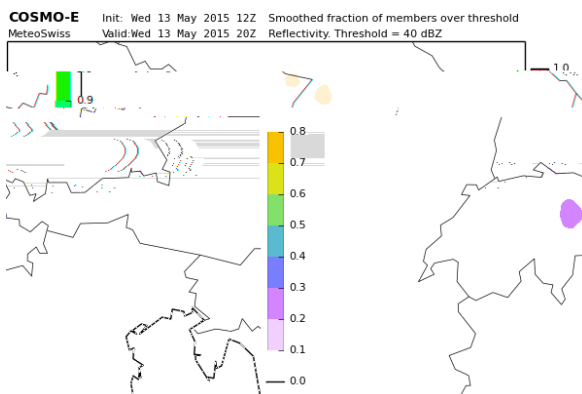
maximum of any member



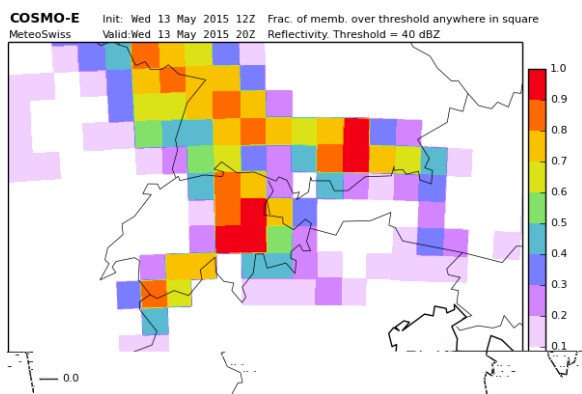
all members exceeding threshold



fraction of members exceeding threshold

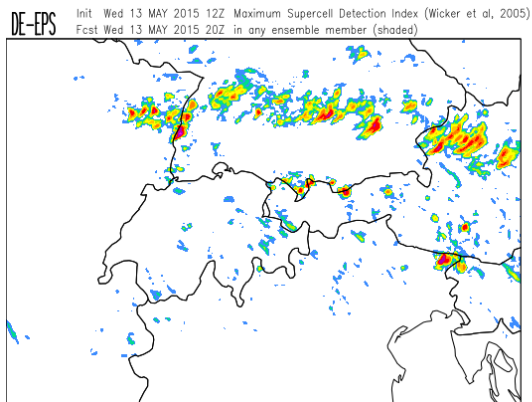


smoothed fraction exceeding threshold ('fuzzy')

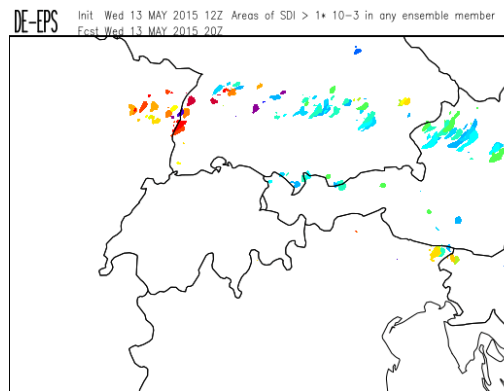


upscaled probability of exceedance ('upscaled')

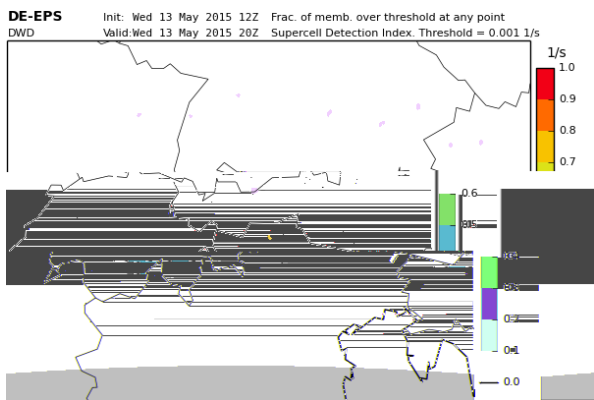
**Fig. 1. Example of the visualizations of ensemble products at the ESSL Testbed 2015. COSMO-E forecast of Reflectivity of Wed 13 May 2015 1200 UTC + 8 hours, valid 2000 UTC.**



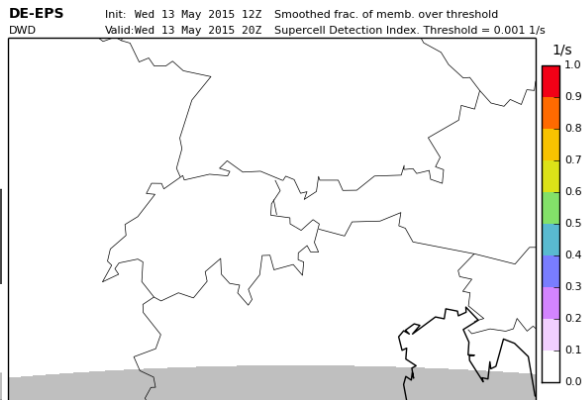
maximum of any member



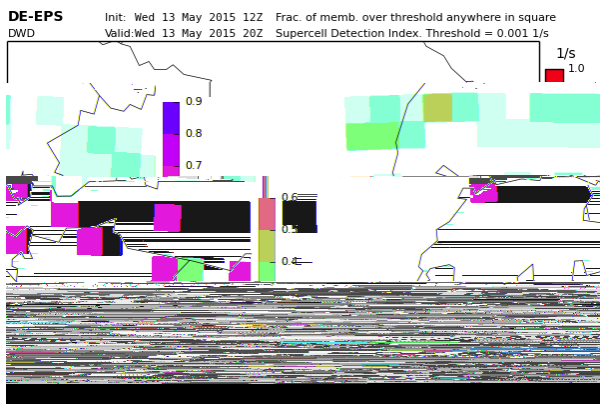
all members exceeding threshold



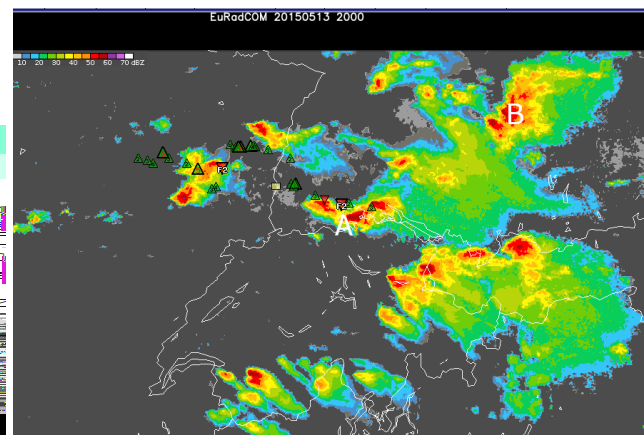
fraction of members exceeding threshold



smoothed fraction exceeding threshold ('fuzzy')



upscaled probability of exceedance ('upscaled')



Left bottom: radar reflectivity (N, W and central part of domain) and ESWD reports (17-20 UTC). "A" and "B" are tornadic storms.

**Fig. 2.** Example of the visualizations of ensemble products at the ESSL Testbed 2015. COSMO-E forecast of Supercell Detection Index of Wed 13 May 2015 1200 UTC + 8 hours, valid 2000 UTC.

### **Specific comments regarding extreme precipitation**

- The upscaled visualization was found to be useful for precipitation amounts.
- Forecasts of high precipitation was possible with long lead times (**Event 4**) for an event tied to the mountains. Forecasts in lower terrain were more difficult.

### **Specific comments regarding graupel**

- The upscaled and ‘members exceeding threshold’ visualizations were found to be very useful for graupel
- Subjectively, integrated graupel worked rather well as a predictor for large hail (see **Events 6 and 7**)
- The used threshold value of 10 kg/m<sup>2</sup> appears to work fairly well for forecasting hail of 2cm and larger. Possibly higher thresholds are required to predict even larger hail
- The success of the large hail forecasts depends strongly on the accuracy of the forecast regarding convective initiation, mode and coverage.

### **Specific comments regarding wind gusts**

- The upscaled visualization was found to be a useful visualization for wind gusts
- Occasionally, COSMO-DE-EPS members would produce too strong cold pools with extreme wind gusts (such as in **Event 7** over southern Germany), essentially predicting a “bow-echo” whereas the convective mode turned out to be more isolated.

### **Specific comments regarding SDI and vorticity; comparison with radar rotation tracks**

- The SDI showed very clear and high signals on 13 May (see **Event 2**, Chapter 4) when powerful supercells, some tornadic, affected South Germany. On other days, signals were weaker. On 29 May (see **Event 3**, Chapter 4), another day with a tornado, SDI again showed clear signals.
- The upscaled visualization was found to be particularly useful for SDI
- We visualized vorticity at 850 hPa in addition to SDI, but this did not allow one to detect rotating updrafts very well. In order to detect these vertical speed must also be provided to compute updraft helicity, which was not the case at the 2015 Testbed.
- In strong flow, SDI signals also occur of the Alps in absence of convection. One may want to look for alternatives of detecting convective updraft rotation that are not ‘triggered’ by flow around mountains.

### **Specific comments regarding CAPE and CIN**

- The visualizations of fraction of members over a threshold was used most often. No need for ‘fuzzy’ or ‘upscaled’ visualizations.
- Introduction of multiple threshold values would be welcome, or perhaps a display of the median values.

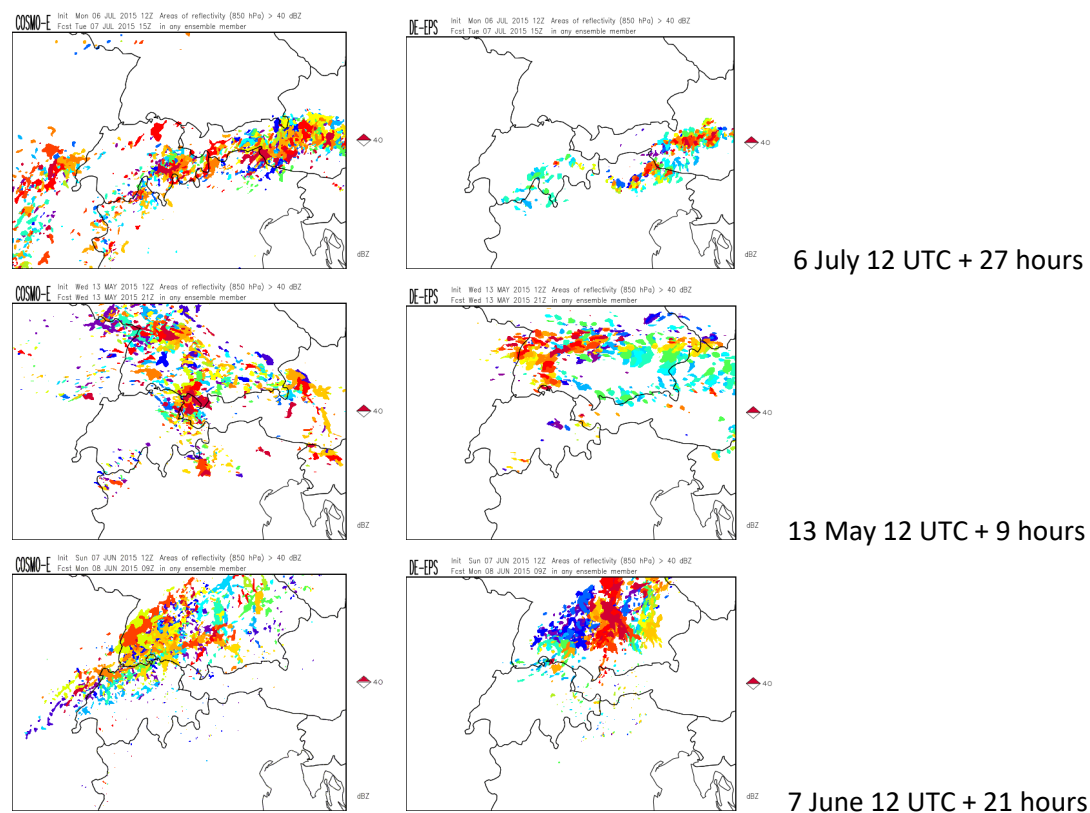


### 3. General assessment of model performance

#### Convective initiation (CI), timing, and coverage in all convection permitting COSMO models

Often, convective initiation was late in all of the convection permitting COSMO models. This effect was noted especially often over flat terrain, in absence of weather systems that produce local lift, and in warm air-masses. Between COSMO-DE or COSMO-1 or between COSMO-DE-EPS and COSMO-E no large differences could be found in this respect, although COSMO-E was perhaps slightly more eager to develop convection.

On a number of occasions, marked differences between the two ensembles were noted, see for instance these examples:



**Figure. Display of all members exceeding 40 dBZ reflectivity threshold. Left: COSMO-E. Right: COSMO-DE-EPS.**

Convective coverage, away from the mountains, was also often underestimated by all of the convection-permitting COSMO models. In cases in which the deterministic COSMO-1 and COSMO-DE would not initiate storms, COSMO-DE-EPS and COSMO-E would often, but not always, still have a few members with convection initiation.

Across the Alps, a few cases of overforecasting of CI were noted in COSMO-E, e.g. on 7 June 2015. Participants called these “false alarms”.

In one case (**Event 4**), predictability of severe rain was very good over the Alps for an event during the afternoon, but the inability of the ensemble models (COSMO-E and –DE-EPS) to simulate the movement of the convection towards lower areas (in this case the region around Basel), made the flooding that occurred there quite unpredictable.

**Differences in model philosophy (convection permitting models)**

A very interesting experience at the Testbed were the different philosophies of the Swiss and German ensembles, COSMO-E and COSMO-DE-EPS and deterministic runs. These differences are evident from the following table:

	<b>COSMO-E</b>	<b>COSMO-DE-EPS</b>
Initialization:	Every 12 hours	Every 3 hours
Forecast time:	Out to 120 hours	Out to 27 hours

	<b>COSMO-1</b>	<b>COSMO-DE</b>
Initialization:	Every 6 to 12 hours	Every 3 hours
Forecast time:	Out to 48 hours	Out to 27 hours

**Participants and Testbed staff were sometimes positively surprised by the accuracy of COSMO-1 forecasts and COSMO-E forecasts at forecast ranges 24 - 48 hours ahead.**

The impression was that for convective events, the forecast quality increases strongly when lead times become so short (i.e. 15 UTC forecast for 18 UTC) that assimilated radar data “tells” the model where convection has initiated. However, the difference between a run on the same day in the morning, or the afternoon the day before is not as large. **It seems useful to run COSMO-DE(-EPS) (occasionally) out to 48 hours.**

This would mean in practice that forecasters can use convection permitting model data to forecast the convective evolution on the next day, something not possible with DWD’s present set up.

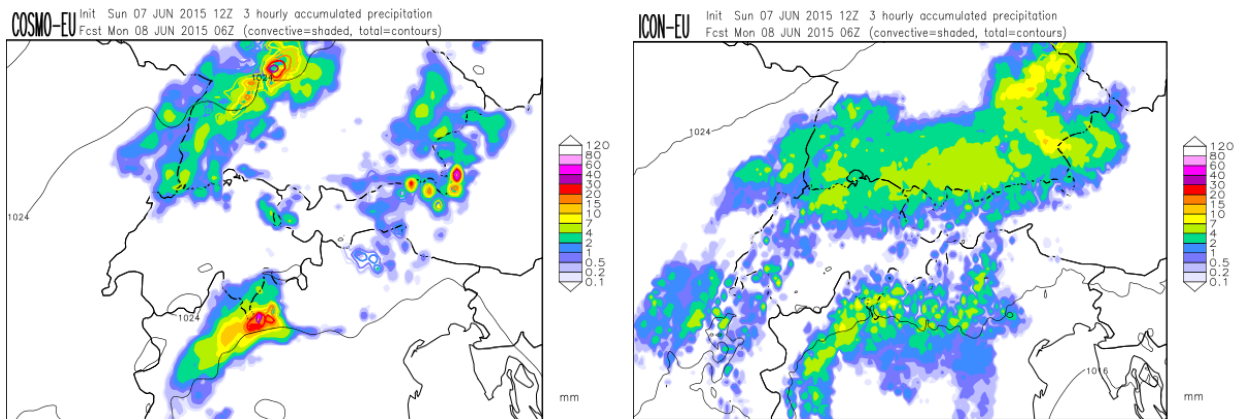
**Models with parameterized convection (COSMO-EU/ICON-EU)**

It was noted at many different occasions that the structures of precipitation in COSMO-EU and the new ICON-EU differed. The characteristics are summarized in the following table.

	<b>COSMO-EU</b>	<b>ICON-EU</b>
Coverage, i.e. extent of area affected by parameterized precipitation	Somewhat low	High
Location, compared to reality	Often biased to the close vicinity of forcing boundaries and fronts. Not typically as far	No obvious bias.

	into the warm air-mass as real storms would.	
Structure and intensity	Very intense local maxima. To some extent, resolution of structure of mesoscale convective systems.	Unrealistic weak precipitation rates over a large area. No structure.

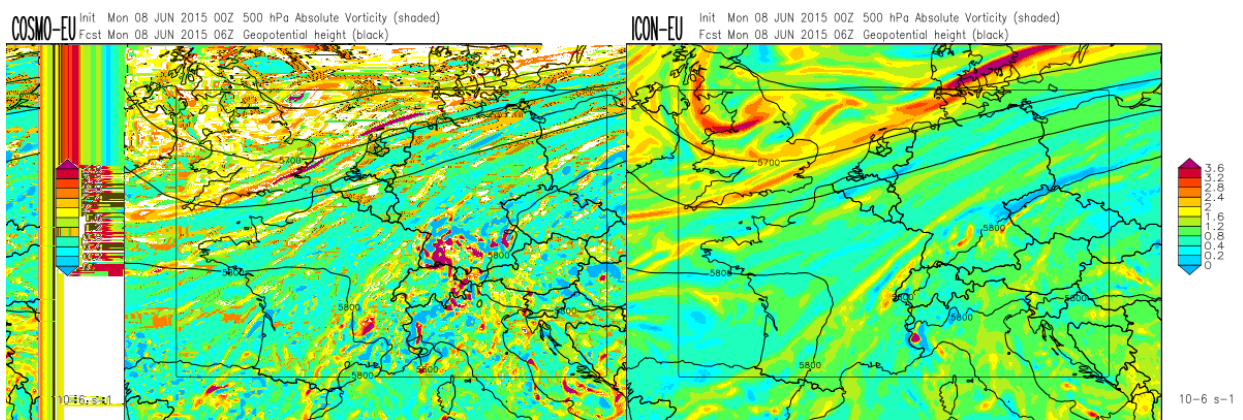
A good property of ICON is that the model seems to be less reluctant in producing convective precipitation further away from active weather systems. In the following example case on 8 June, the “warm air” is to the southeast:



COSMO-EU + 18 hour forecast (3 hour acc. precip.)

ICON-EU +18 hour forecast (3 hour acc. precip.)

ICON-EU’s smoothness seems unrealistic and little variability is created on the mesoscale. This may be a problem if the model is at some point run in ensemble mode. On 8 June, a mesoscale convective vortex occurred and was forecast by COSMO-EU, and the convection-permitting models, but not by ICON-EU. This figure shows the different amounts of variability introduced by deep convection by comparing 500 hPa vorticity between these two models:



COSMO-EU + 6 hour forecast  
500 hPa vorticity

ICON-EU +6 hour forecast

Another difference is that the larger areas of weak precipitation in ICON ( as opposed to the small areas of intense precipitation in COSMO-EU) will require users of direct model output to get used to. They will need to mentally adjust the forecast precipitation field in the opposite way as they used to with COSMO-EU. In COSMO-EU, they would need to downplay local maxima, whereas in ICON they should probably expect higher maxima than the direct model output.

### COSMO-1

COSMO-1 storms and reflectivity structures seem to be smaller in the model than in reality. It was often noted that areas of stratiform precipitation were underestimated/almost absent:

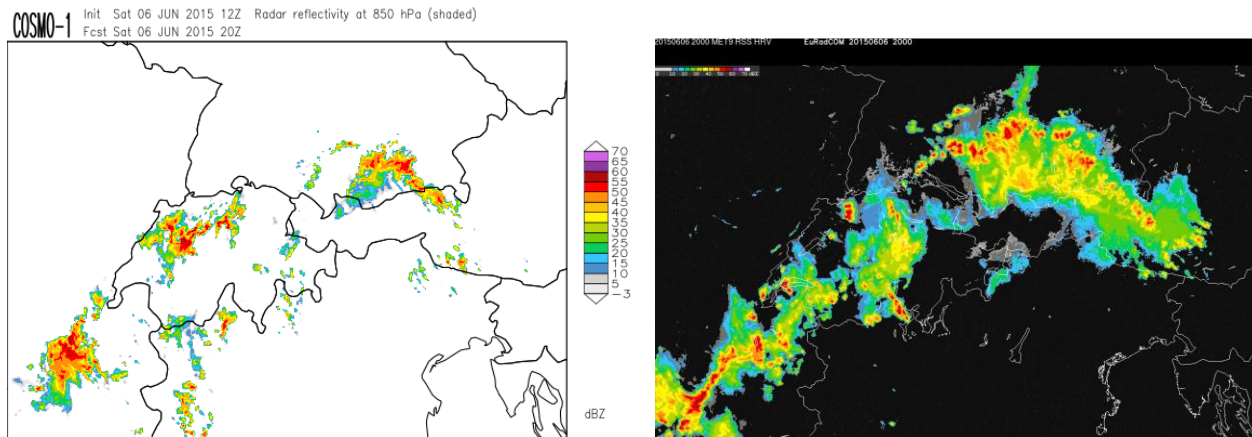


Figure. Left: COSMO-1 forecast (+8 hours) of reflectivity at 850 hPa. Right: radar (same colorscale).

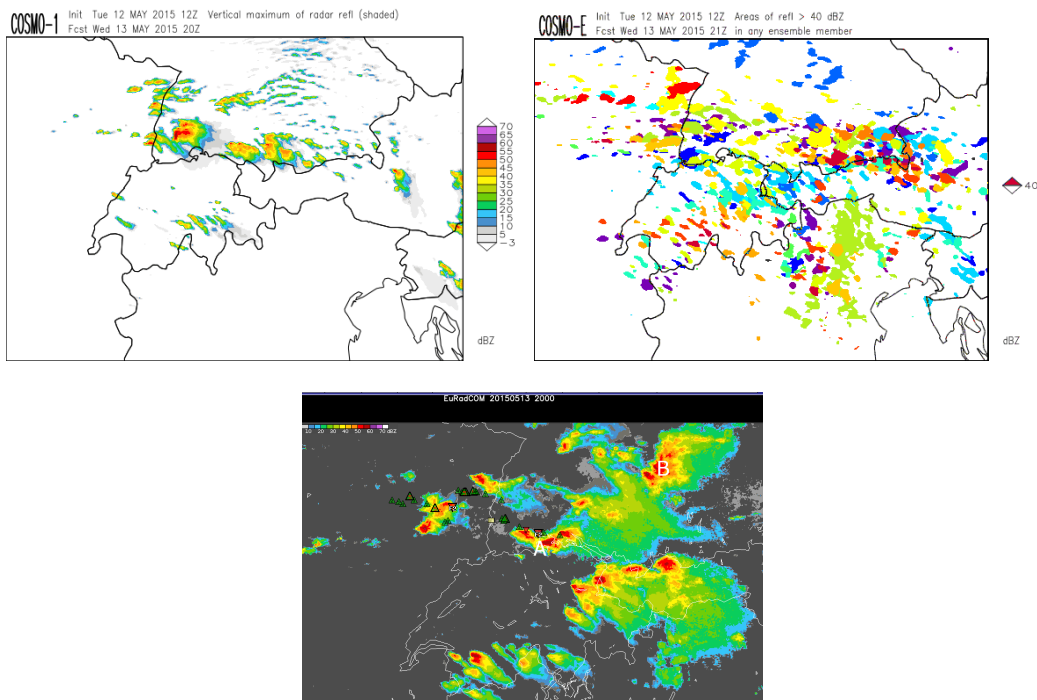


Figure. Comparison of deterministic, ensemble forecast (both 32 hour lead time) and radar observations. Which one is more useful?

## COSMO-E/-1 specific questions from MeteoSwiss

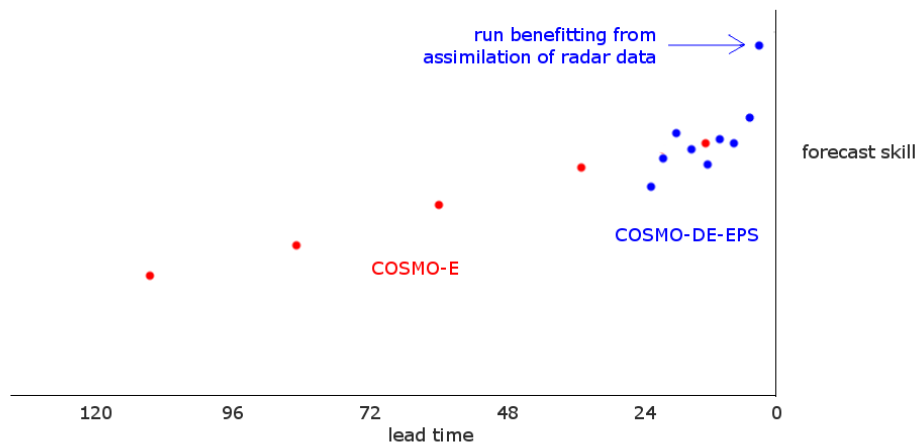
The following questions were discussed among participants

- Accuracy of forecast as a function of lead time. The question how far into the future it makes sense to run and use a deterministic model with 1 km grid-spacing. The question at which lead time ensemble forecasting clearly outperforms deterministic solutions.

In order to compare the accuracy of COSMO-E and COSMO-1 as a function of lead time, we must first establish the characteristics of forecast quality of these two systems.

### Subjectively estimated accuracy of ensemble systems

For COSMO-E and COSMO-DE-EPS, whenever the quality was fair to good 6 or 9 hours before an event, a significant part of the skill was often already present 27 hours before the event. A strong increase in skill was usually observed for the last one or two COSMO-DE-EPS forecasts before an afternoon/evening convective event, as these would benefit strongly from the assimilation of radar data after the moment of convective initiation. This is something the modelling system has great difficulty in simulating accurately on its own. This is illustrated in the following figure:



See **Event 7** (Chapter 4), for an example of this. The deterministic solution of the COSMO-1 runs are hard to compare in a formal way to the output of an ensemble system like COSMO-E, but a subjective estimate can naturally be given. Participants noticed on several occasions that particular features of the deterministic run were forecast well 24 or even 48 hours in advance.

For instance, the location of one of the two tornadic supercells on 13 May was predicted very accurately by the COSMO-1 run of 12 UTC the day before, better even than the COSMO-1 run of 12 UTC the same day. The COSMO-E ensemble was useful by adding information on the possible locations of other intense cells. That said, the real distribution was not entirely accurate in that COSMO-E underestimated the activity over Bavaria.

- The question if and, if yes, how forecasters can determine at an early stage which ensemble member is most accurate for the medium-range. Given that the lead-time of COSMO-E is 5 days and the model is run every 12 hrs, can a best member be selected, e.g. after 10 hours into the forecast?

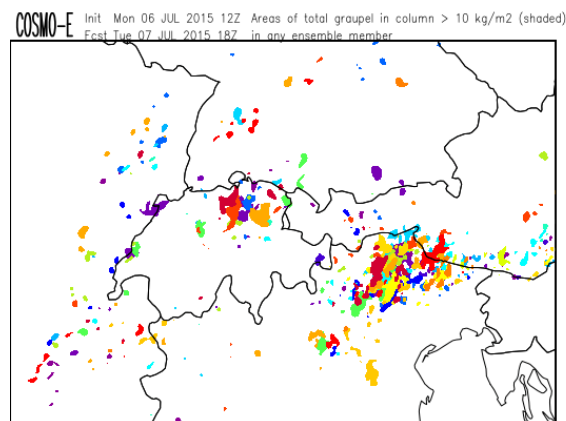
The honest answer to this question is **no**, not at the moment, for the following reasons:

First, for technical and practical reasons tracking an individual ensemble members is very difficult. At the Testbed, the only way to identify individual ensemble members is through the color code in the “members exceeding the threshold” plot. Tracking a particular ensemble member by its color code is next to impossible, since the eye cannot easily distinguish 20 color shades and pick out one particular color shade to track through time. A much more sophisticated system must be developed to make this possible.

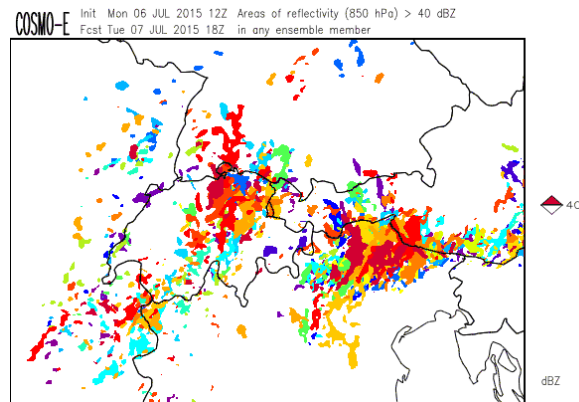
Second, it is unclear which criterion one would use to identify the best member 10 hours into the integrations. The development of convection in the 24-120 hour period is tied to large-scale weather patterns that, at that time, are located far out of the COSMO-E domain. Hence, identifying a “best member” would require an investigation of the driving ECMWF EPS members, rather than the evolution within the COSMO-E domain.

- The question if it is possible, and if so, how to identify the physical processes in particular ensemble members, in addition to using statistical quantities based on the entire ensemble.

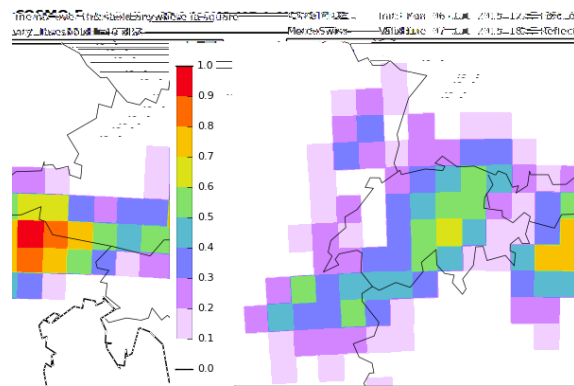
The “members exceeding the threshold” displays the outlines of the reflectivity, the SDI, or the integrated graupel. Hence, they give an idea of the rate at which precipitation is produced, the extent to which updrafts rotate or hail is formed in particular members. The individual members can be distinguished by having varying color codes. This allows one to get an impression, e.g., of the number of cells in which large amounts of graupel are simulated to have formed, see figure:



For reflectivity, the image looks like this:



It shows clearly that many ensemble members develop linear clusters of storms over Northeast Italy and some over Northeast Switzerland as well. In contrast, the statistical 'upscaled probability' display does not give this information, but summarizes well which fraction of ensemble members produce a storm in a particular area of reference:



- The question where human forecasters can contribute most to the forecast process, and where can (ensemble) model output be automatically be translated into forecast products.

This question was discussed quite in detail in several participant groups. The answers can be found in the attached questionnaire forms. A very comprehensive answer that was given, was the following:

*Forecasters can adjust NWP forecasts when they recognize particular phenomena (like showers behind a cold front that a forecaster knows are typically not represented a model). Forecasters can also adjust e.g. for biases for late initiation in NWP. A MOS will not easily be capable of doing that.*

*Also, as soon as observations diverge from NWP, NWP forecasts cannot be used any longer. Some of these adjustments to observations can be made by forecasters. They can partly be automated, but there are reasons that limit to what extent this can be automated:*

*1. NWP forecasts need to be adjusted to the needs of the customers. Needs are very different, hence automatization would be very laborious.*

*2. The forecaster has added value because of quick reactions to local and quickly changing weather that deviated from the model.*

*3. The customers would like to have the forecasts interpreted, preferable by a human person.*

*4. Forecasters are often - perhaps this is not right, but it is a fact - incorporated in decision-making based on the forecast and is able to incorporate knowledge of what is a stake. The forecast can be adjusted by the forecaster to trigger the optimal decision.*



## 4. Events

### Event 1

(omitted in this version)

### Event 2: 13 May 2015

At the Testbed, we asked participants to look into the severe weather situation on this day in detail and evaluate model performance.

Description of the event:

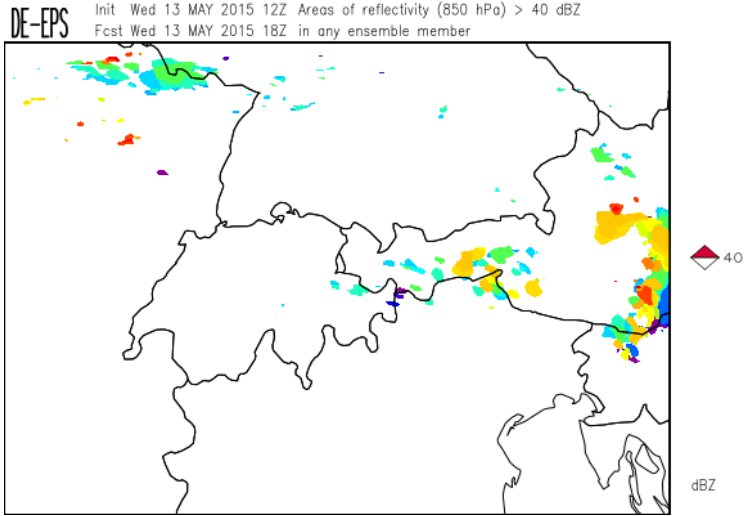
South of a west-east oriented front, a somewhat moist and rather unstable air-mass was present. Very strong wind shear (30 m/s 0-6 km bulk shear and 15 m/s 0-1 km bulk shear) was in place, as well as substantial storm-relative helicity. In this environment, intense supercells developed over eastern France in the late afternoon and early evening. They produced very large hail and a number of strong tornadoes as they tracked eastward across the Vosges and the southern Black Forest (between Feldberg and Lake Constance, just before 20 UTC). Further storms developed over central and eastern Switzerland and across northern Baden-Württemberg. One of the latter storms produced another significant tornado (F3) slightly north of Augsburg (Gersthofen, Affing in Bavaria just after 20 UTC).

Some observations regarding model performance:

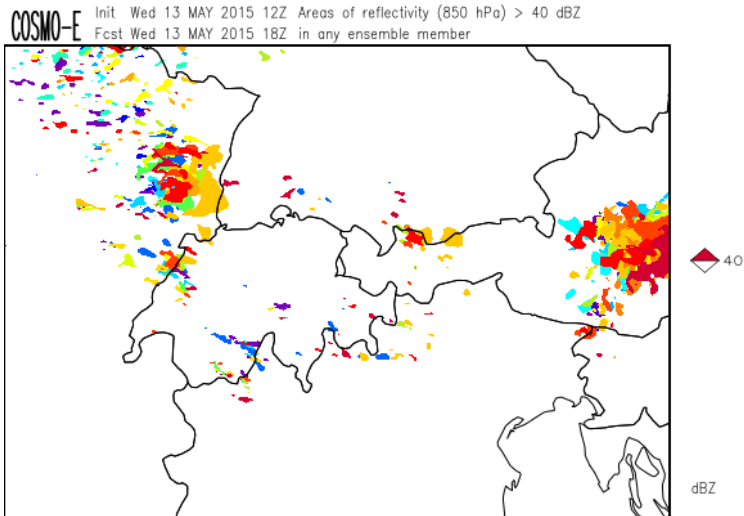
- 1) The COSMO-DE-EPS runs initialized at 00/03/06/09/12/15 UTC simulated storms with very high supercell detection index, SDI, (rarely if ever seen at the testbed before) moving from west to east across southern Germany. The COSMO-E run of 12 UTC agreed. In fact, the COSMO-E runs of 12 UTC the day before (12 May) already produced such high SDI storms.
- 2) The COSMO-E of 12 UTC 13 May was able to capture the actual location of the convective initiation over eastern France better than any of the COSMO-DE-EPS runs until to the point where the storms were assimilated at 18 UTC.
- 3) Especially COSMO-DE and COSMO-DE-EPS made the northern storms too active and underestimated the storms "in the warm air" across the southern Vosges and Baden-Württemberg. COSMO-E and COSMO-1 did not have this problem as much. All models underestimated the activity even further south over central and eastern Switzerland, although a few members had signals.
- 4) COSMO-1 was surprisingly accurate with the placement of the most severe cell in southwestern Germany 32 hours in advance, i.e. the 12 UTC run of the day before. Likewise, early morning runs of COSMO-DE (i.e. +20, +17 +14 hour forecasts) already showed potential for strong rotation along the northern line, very close to the location of the tornado near Affing, Bavaria, during the evening.
- 5) COSMO-1 showed clear supercell hook-echo structures. The storm structures tended to be finer and the storms smaller in the model than on radar.

- 6) The ICON model showed precipitation more or less in the right areas, possibly a bit too much. The precipitation field in the ICON model shows very little structure. COSMO-EU generally underestimated the convective precipitation and simulated not enough over France and Baden-Württemberg.

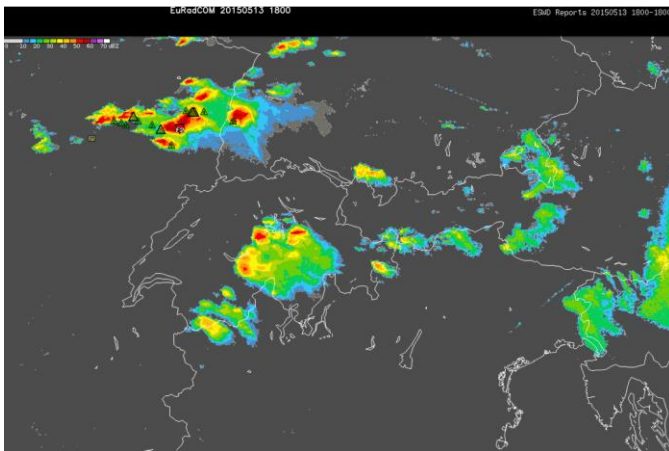
1800 UTC. 1200 UTC + 6 hour forecasts



DE-EPS

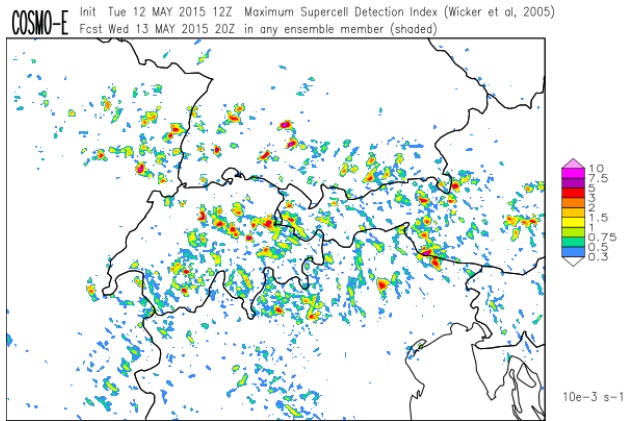


COSMO-E



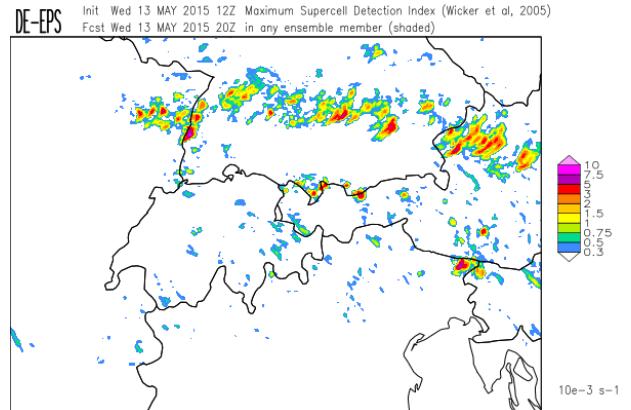
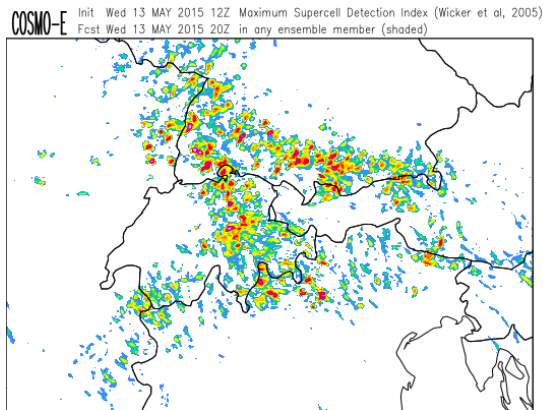
Radar and ESWD (past 3 hours)

2000 UTC. COSMO-E (left), COSMO-DE-EPS (right). Supercell-detection index.

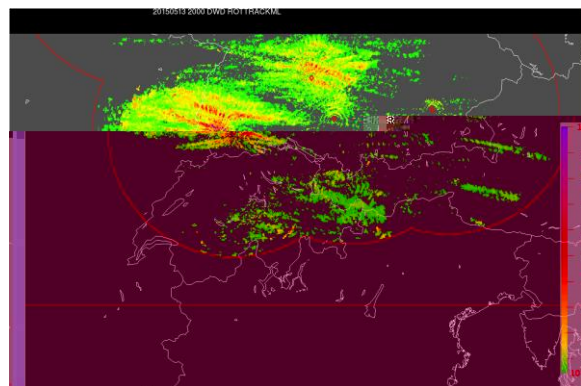
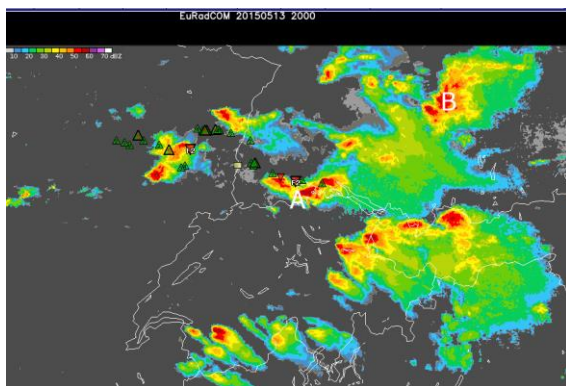


COSMO-DE-EPS not available

COSMO-E. Supercell detection index (maximum of all members). 32 - hour forecast.

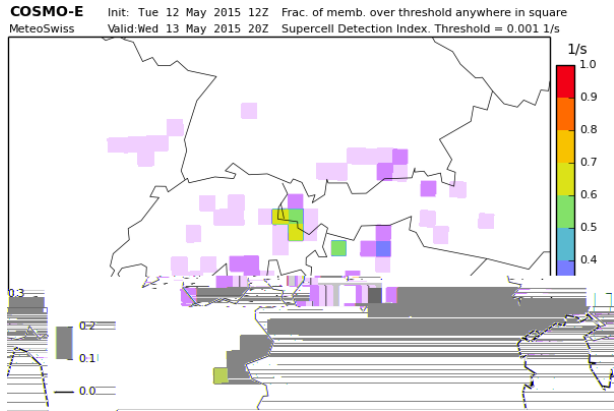


COSMO-E. Supercell detection index (maximum of all members). 8 - hour forecast.



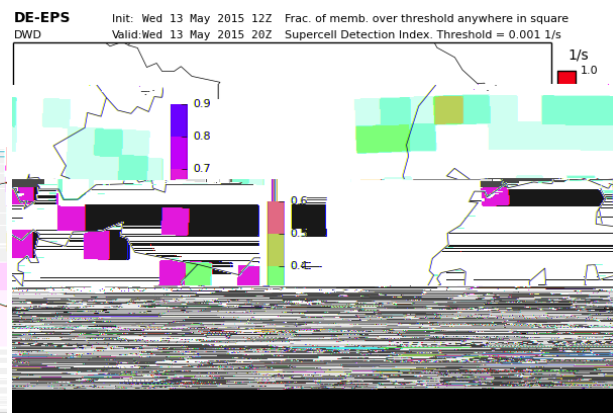
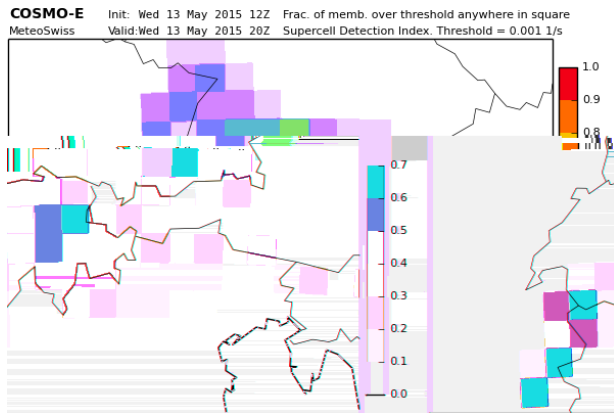
Left: Radar + ESWD reports 17-20 UTC (A and B the significant tornado producing storms)  
 Right: Radar-based mid-level rotation tracks between 17 and 20 UTC.

2000 UTC. COSMO-E (left), COSMO-DE-EPS (right).  
 Supercell-detection index, now displayed using the **upscaled** product.

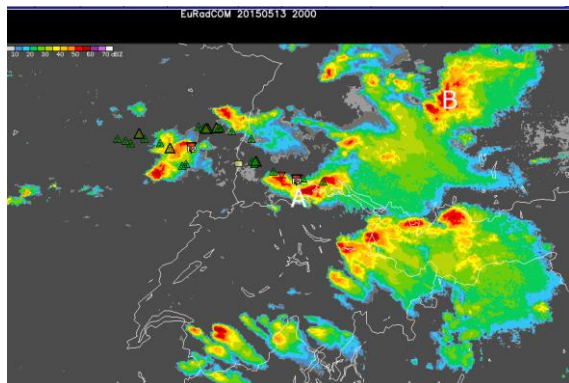


COSMO-DE-EPS not available

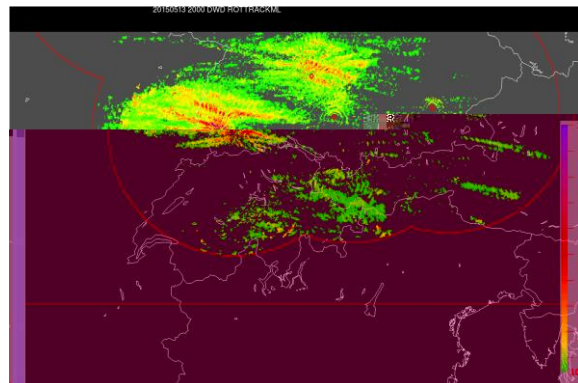
COSMO-E. Supercell detection index (maximum of all members). 32 - hour forecast.



COSMO-E. Supercell detection index (maximum of all members). 8 - hour forecast.

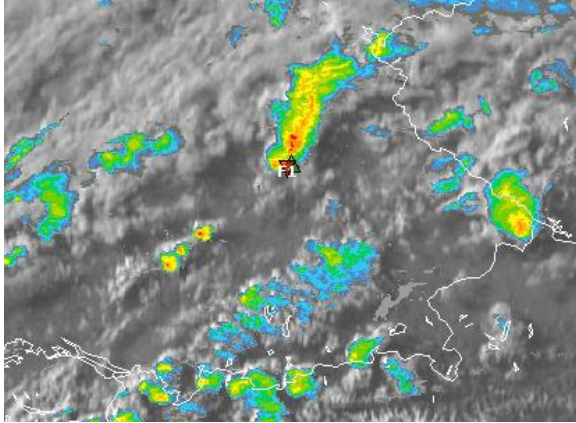


Radar + ESWD reports 17-20 UTC

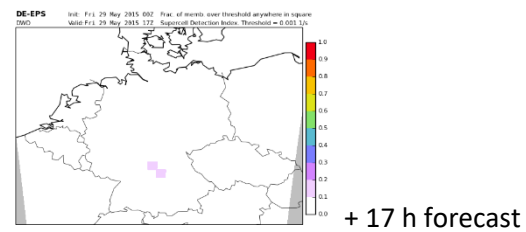
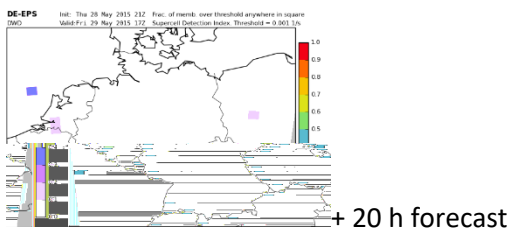
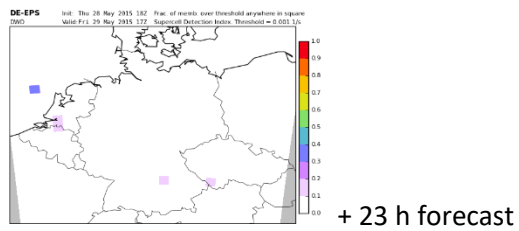
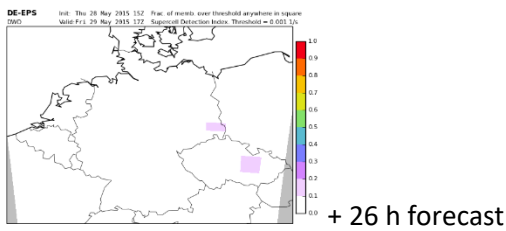


Radar-based mid-level rotation tracks.

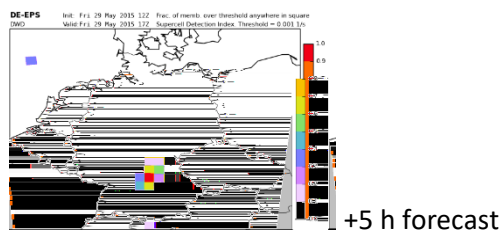
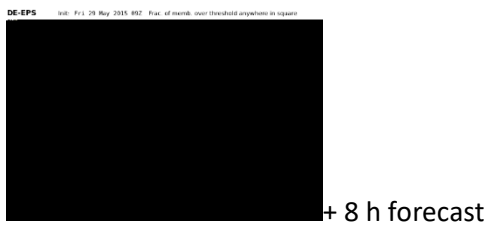
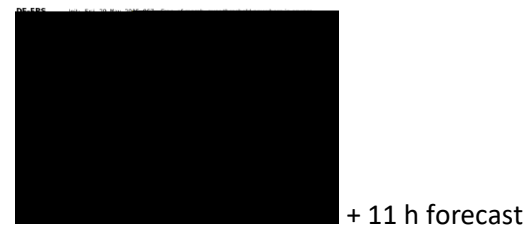
### Event 3: 29 May 2015



Radar and satellite at 1735 UTC. ESWD reports 1435-1735 UTC.

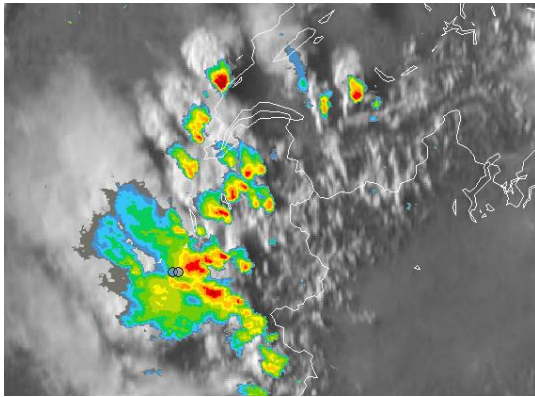


Not available + 14 h forecast



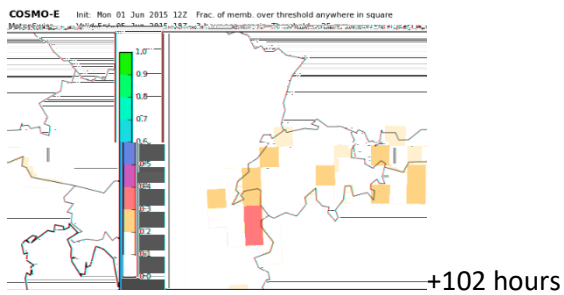
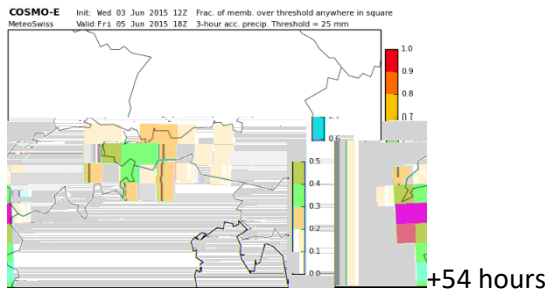
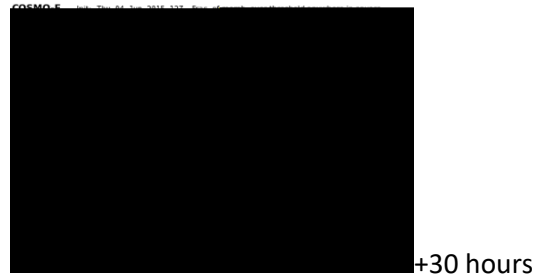
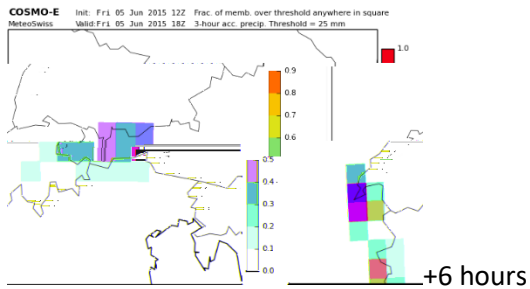
COSMO-DE-EPS Supercell Detection Index forecasts, “upscaled” visualization.

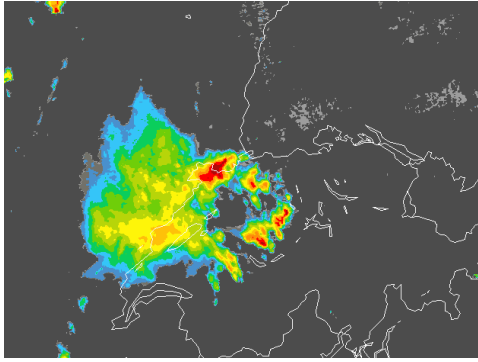
### Event 4: 5 June 2015



Radar and ESWD flood reports (two blue circles) at 1700 UTC 5 June 2015.

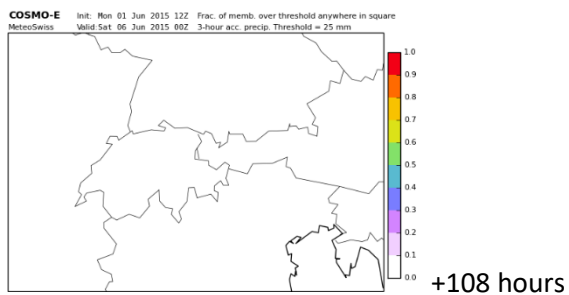
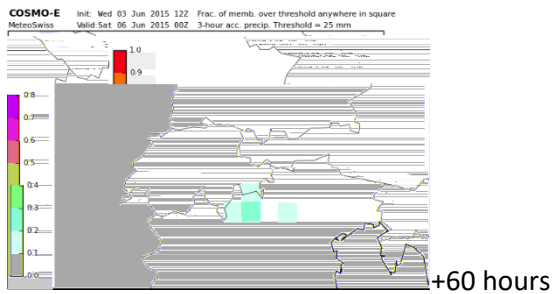
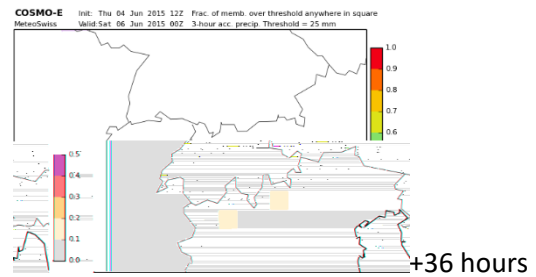
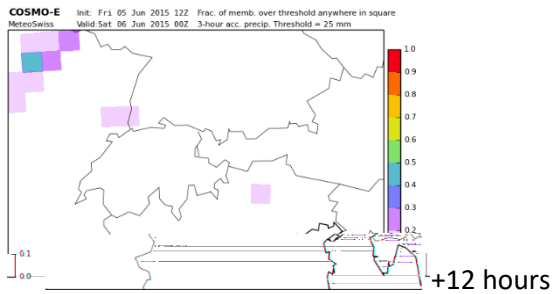
**COSMO-E** forecasts of 3 hourly precipitation (fraction of members exceeding 25 mm / 3 hours) between 1500 and 1800 UTC:



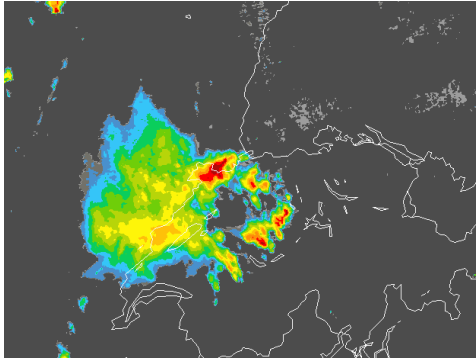


Radar image at 2200 UTC, showing a storm producing severe wind gusts and floods near Basel.

**COSMO-E** forecasts of 3 hourly precipitation (fraction of members exceeding 25 mm / 3 hours) between 2100 and 0000 UTC:

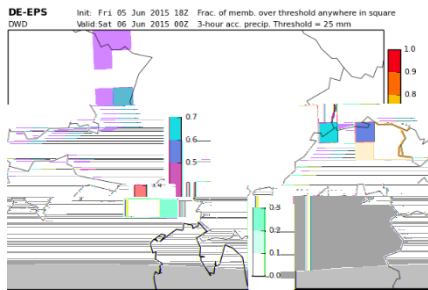






Radar image at 2200 UTC, showing a storm producing severe wind gusts and floods near Basel.

**COSMO-DE-EPS** forecasts of 3 hourly precipitation (fraction of members exceeding 25 mm / 3 hours) between 2100 and 0000 UTC:



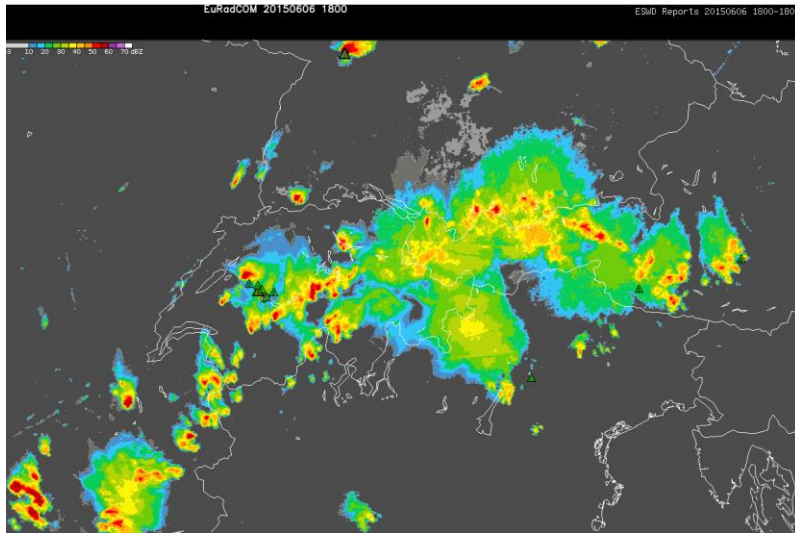
+6 hours



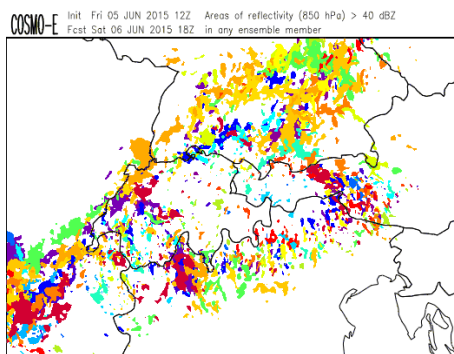
+9 hours

Earlier runs show < 0.1 probability of exceedance in the Basel area.

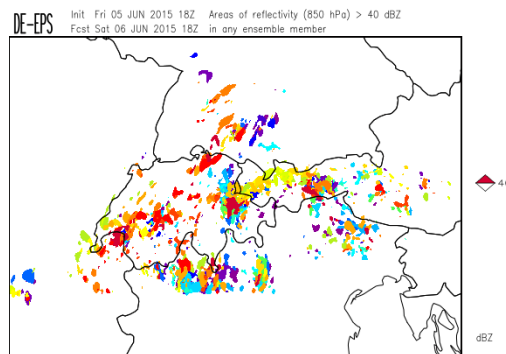
**Event 5: 6 June 2015**



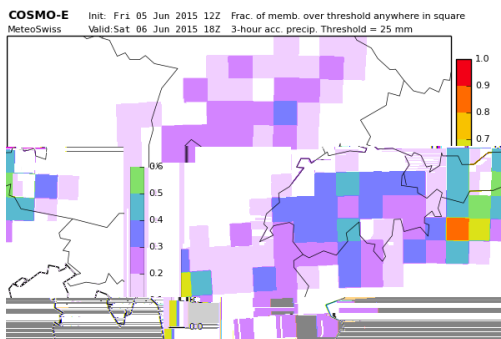
Radar 6 June 1800 UTC. ESWD reports 1500 – 1800 UTC (large hail over Switzerland and near Karlsruhe).



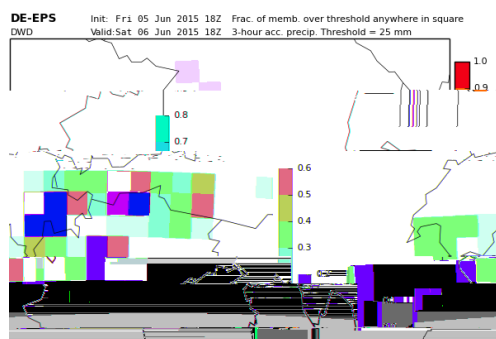
+ 30 hour forecast COSMO-E  
Reflectivity > 40 dBZ



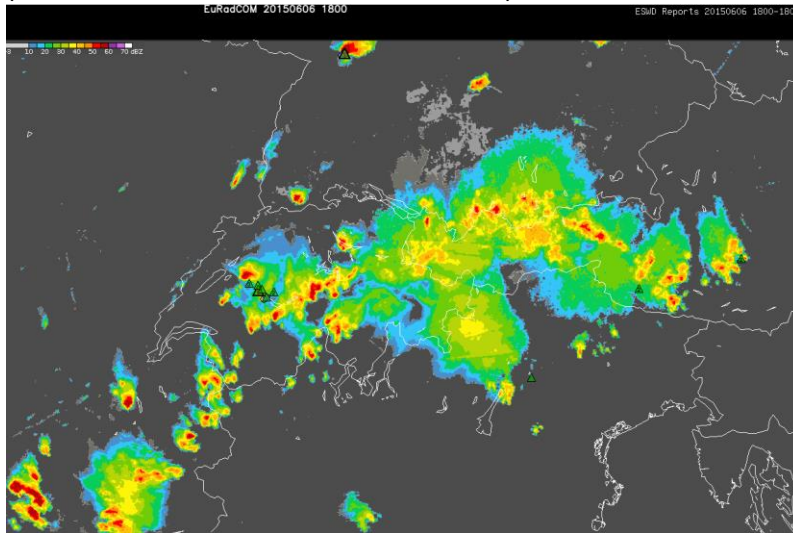
+ 24 hour forecast COSMO-DE-EPS



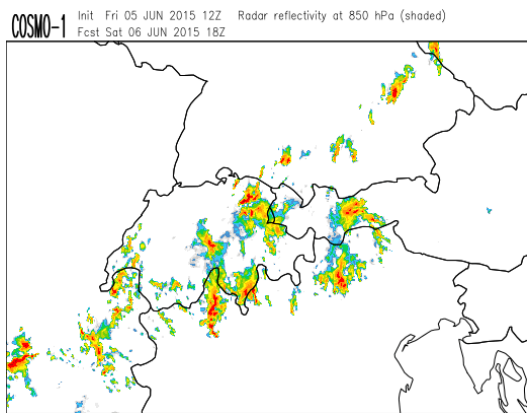
Upscaled probability of 25 mm/3h



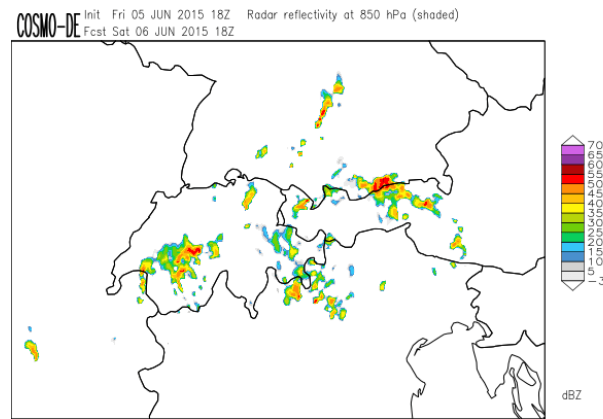
Radar observation at 1800 UTC (EuRadCom) showing much larger areas of reflectivity than models (model and radar colorscales are identical):



Deterministic forecasts for 1800 UTC:



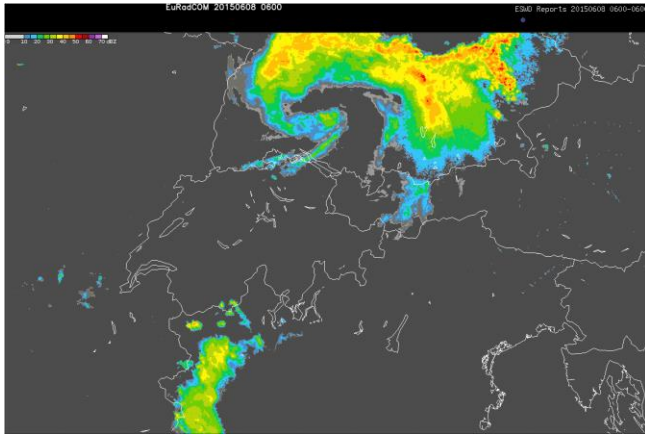
COSMO-1 forecast (reflectivity at 850 hPa)  
+30 hours



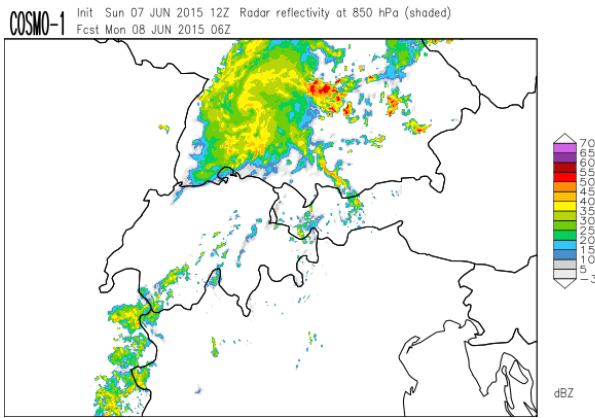
COSMO-DE forecast (reflectivity at 850 hPa)  
+24 hours

**Event 6: 7-8 June 2015**

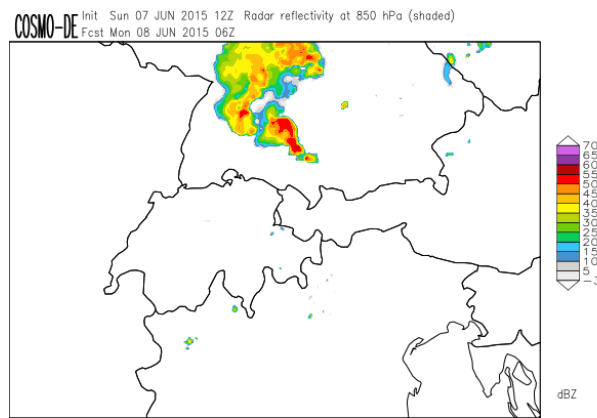
Radar 7 June 1800 UTC



Radar reflectivity at 0600 UTC 8 June 2015.

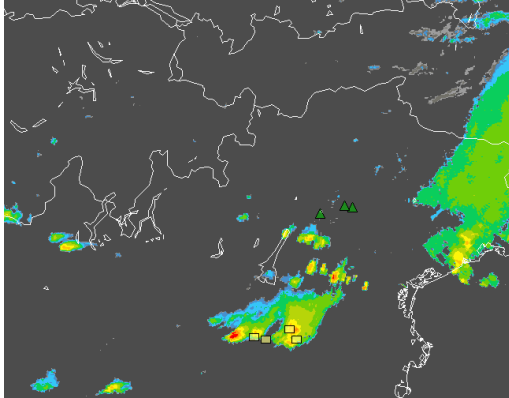


COSMO-1 + 18 hour forecast (reflectivity)

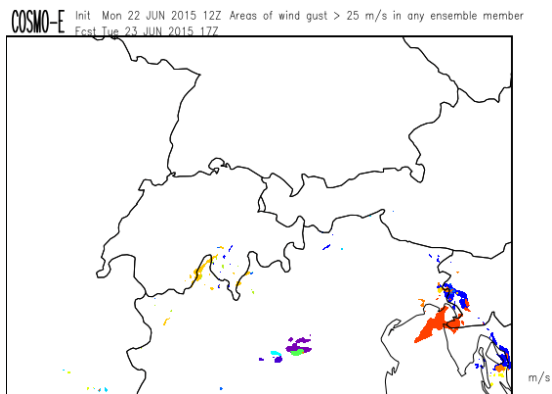


COSMO-DE +18 hour forecast (reflectivity)

**Event 5: 23 June 2015**

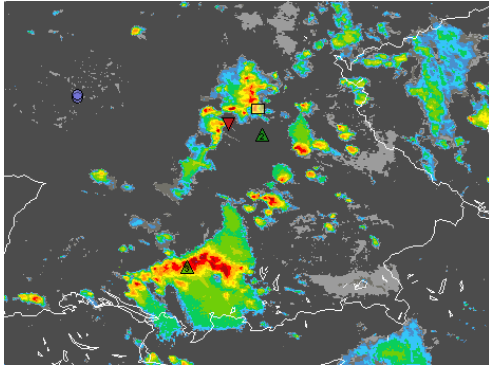


Radar image at 1745 UTC 23 June 2015 and ESWD reports (yellow rectangles: severe wind gusts; green triangles: hail reports) 1445-1745 UTC.

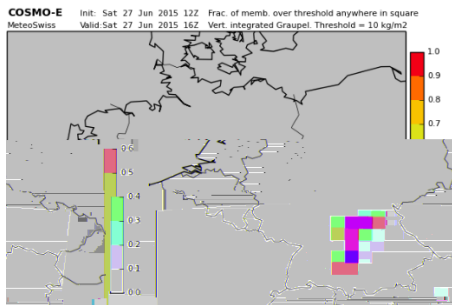


COSMO-E forecast of day before 12 UTC.

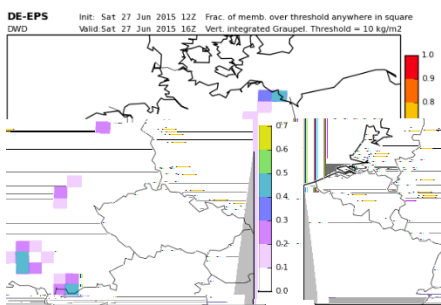
**Event 6: 27 June 2015**



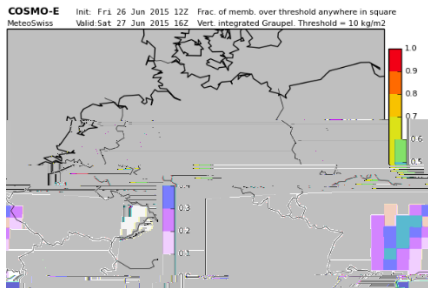
Radar at 1600 UTC 27 June 2015 and ESWD reports 1300-1600 UTC (blue circle = heavy rain, yellow rectangle = wind, green triangle = hail, red triangle = tornado).



COSMO-E (4 hour fcst)



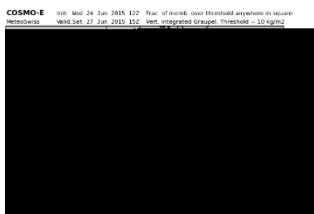
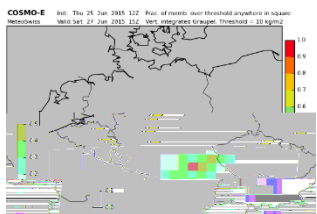
DE-EPS (4 hour fcst)



COSMO-E (28 hour fcst)

**not available**

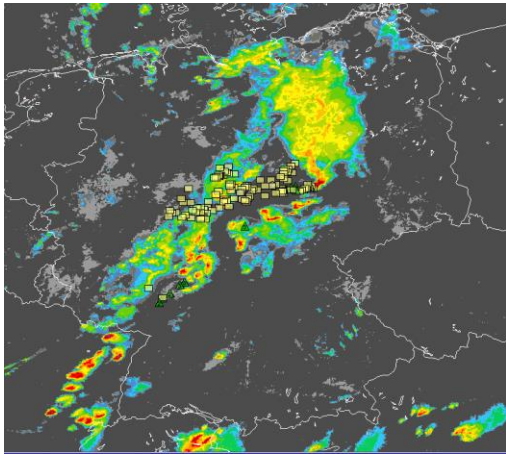
DE-EPS (28 hour fcst)



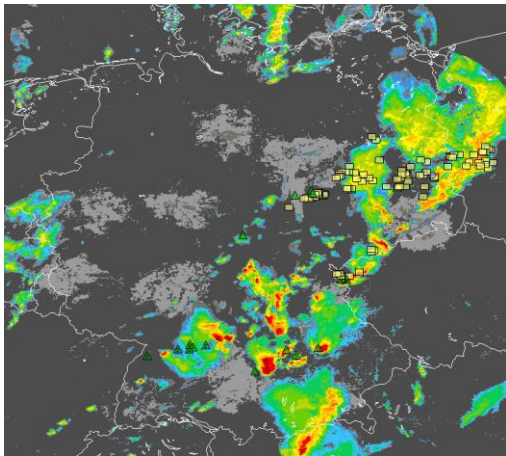
COSMO-E + 51 hours, + 75 hours, +99 hour forecasts

**Forecasts of integrated graupel > 10 kg/m2**

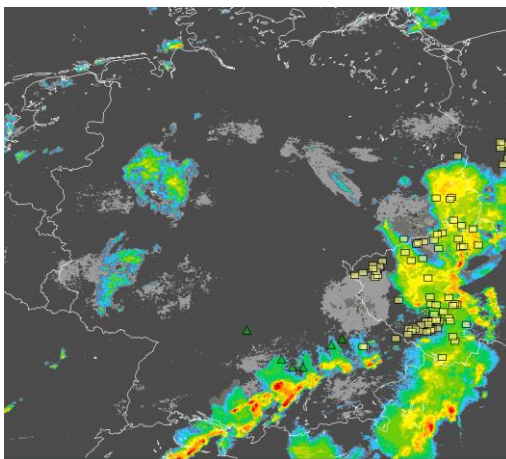
**Event 7: 7 July 2015**



1800 UTC (reports 1500-1800)



2100 UTC (reports 1800-2100)



0000 UTC (reports 21-00 UTC)

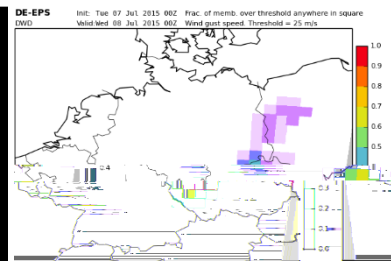
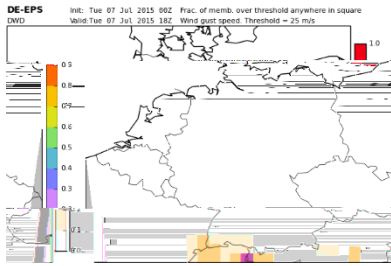
**Radar and ESWD reports (wind: yellow rectangles; green: hail reports)**

For 1800 UTC

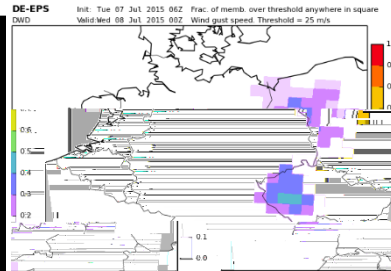
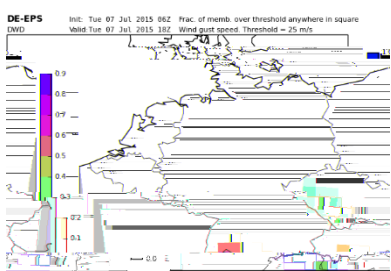
2100 UTC

0000 UTC

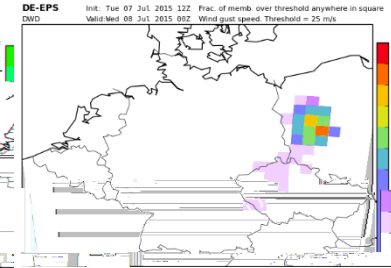
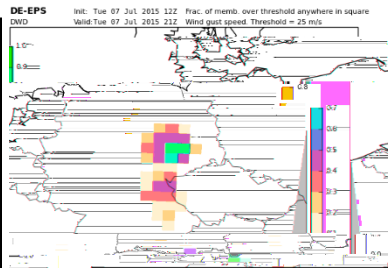
Initialized: 00 UTC



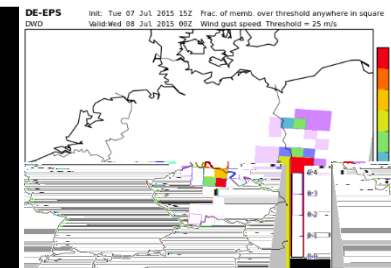
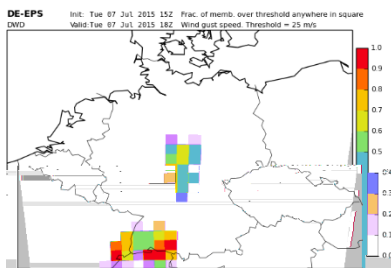
Initialized: 06 UTC



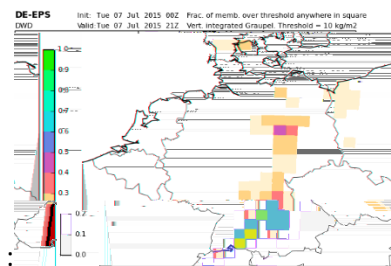
Initialized: 12 UTC



Initialized: 18 UTC



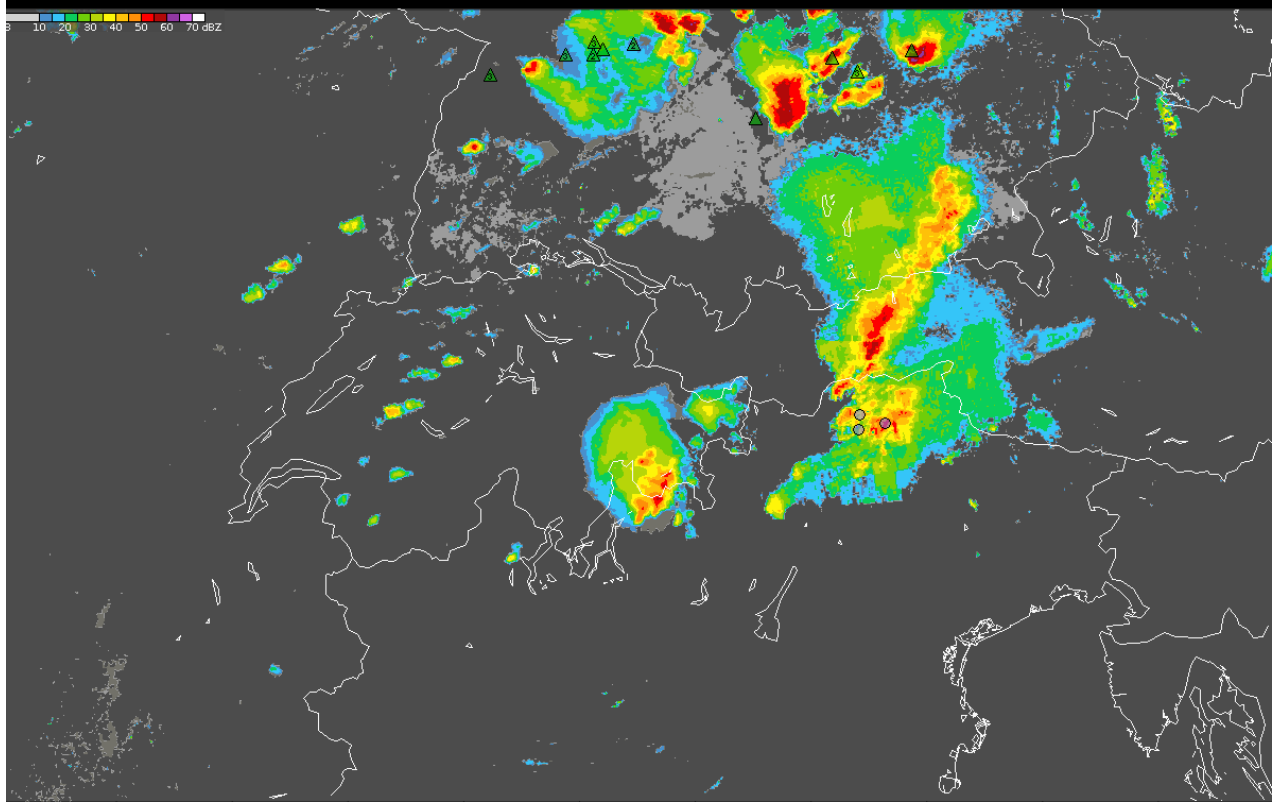
### Forecasts (COSMO-DE-EPS) of exceedance of 25 m/s wind gust



### DE-EPS forecast of vertically integrated graupel



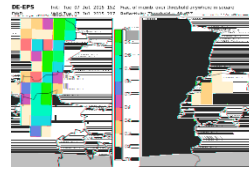
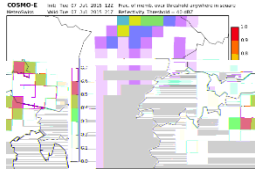
0 10 20 30 40 50 60 70 dBZ



COSMO-E

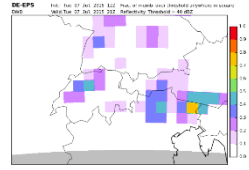
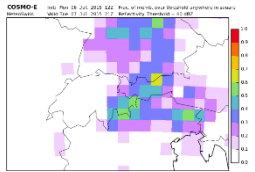
COSMO-DE

n/a

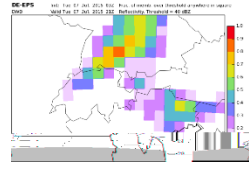
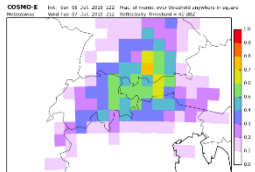


+ 6 hours

n/a



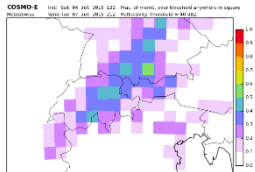
+ 9 hours



+ 18 hour

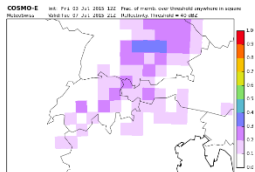
n/a

+ 33 hours



n/a

+ 57 hours



n/a

+81 hours

n/a

+105 hours

Probability of exceedance of 40 dBZ in COSMO-E and COSMO-DE-EP5

## 5.

### June and July 2015

5 May 2015	<ul style="list-style-type: none"> <li>- A strongly-forced convective line in an environment of high shear was well-forecast by COSMO-DE and (-EPS) at 06 UTC to pass NE Germany at 12-18 UTC. The line produced a significant F3 tornado in Bützow and several others. Modest signals were visible in the “upscaled” version of the SDI. A refinement of thresholds/areas of reference is probably necessary.</li> </ul>
12 May 2015	<ul style="list-style-type: none"> <li>- A cold front moves into Germany from the northwest. In an environment of strong, mostly unidirectional shear, a number of supercells develop. Convective initiation took place further to the southwest (away from the forcing) and 2-3 hours earlier than most COSMO-DE-EPS members of the 15/18/21/00/03/06/09 UTC runs. The location and timing of the storms was no better in the later runs than in the earlier runs. Potential for severe wind gusts and hail was rather well forecast 26-27 hours in advance.</li> </ul>
13 May 2015	<ul style="list-style-type: none"> <li>- <b><i>See chapter on this event.</i></b></li> </ul>
29 May 2015	<ul style="list-style-type: none"> <li>- A small number of storms developed over central Bavaria and moved eastward. One storm produced a tornado near Nürnberg. COSMO-DE-EPS produced increasingly strong signals in SDI in subsequent runs.</li> </ul>
5 June 2015	<ul style="list-style-type: none"> <li>- Storms developed over Alps, especially the French Alps, where some local flooding occurred, and Tyrol. The local flooding events were well forecast by COSMO-E with long lead times. During the late evening, floods occurred in the Basel region as storms moved into the lowlands. This evolution was not captured by COSMO-E or COSMO-DE-EPS until the 12 UTC COSMO-DE-EPS run.</li> </ul>
6 June 2015	<ul style="list-style-type: none"> <li>- Widespread storms formed over the Alps, and isolated intense cells over lower terrain across South Germany and East France, Switzerland and France. The Alps storms clustered and moved into southern Germany. The evolution of the Alpine storm system was well forecast by COSMO-DE-EPS and COSMO-E. The possibility of convective initiation across lower terrain across South Germany was also indicated by the ensemble systems as well as COSMO-DE. COSMO-1 was too reluctant.</li> </ul>
7-8 June 2015	<ul style="list-style-type: none"> <li>- Widespread storms formed once again over the Alps, but also over lower terrain in South Germany. During the evening, a line of storms moved out of Switzerland to the northeast, developing into a MCS with a well-defined Mesoscale Convective Vortex. Especially COSMO-1 produced a fairly good forecast of the vortex 18 hours ahead, COSMO-DE was not too bad either. Both models had a bias away from the warm air. Participants found that the ensembles did not add much information in this particular case.</li> </ul>

	<ul style="list-style-type: none"> <li>- ICON did not resolve the vortex. COSMO-EU was better at that, but it had a displacement of precipitation “away from the warm air” that was not present in ICON.</li> </ul>
19 June 2015	<ul style="list-style-type: none"> <li>- A few of strong (super) cells moved southward across the Po-Valley during the evening. COSMO-E provided almost no hints that these storms would form, but COSMO-1 did (12 UTC run, 00 UTC run).</li> </ul>
23 June 2015	<ul style="list-style-type: none"> <li>- Storm with severe wind gusts moved eastward across Po-Valley. COSMO-E correctly hinted at the gust threat over 24 hours in advance (few ensemble members had this scenario). COSMO-1 was not as clear-cut.</li> </ul>
27 June 2015	<ul style="list-style-type: none"> <li>- A number of fast-moving storms developed across central and southern Germany, producing large hail, local very heavy rain, severe wind gusts and a tornado. COSMO-DE-EPS and COSMO-E data show the large hail risk clearly, less so the wind gust risk.</li> </ul>
7 July 2015	<ul style="list-style-type: none"> <li>- An MCS tracked eastward over northern Germany with strong wind gusts. Isolated supercells developed over southern Germany, producing large hail, transforming into another severe wind producing MCS before crossing the Czech border. COSMO-DE-EPS had trouble with predicting the wind gusts with the northern system. The southern system was simulated as a bow echo (instead of isolated supercells) by many ensemble members hence overforecasting the wind gust threat in the south.</li> </ul>

# Feedback des DWD zum ESSL Testbed 2015 und Verbesserungsvorschläge

---

## Teilnahme des DWD am Testbed 2015

Produkte/Modelle: COSMO-DE, COSMO-DE-EPS (verschiedene Parameter und unterschiedliche Visualisierungsvarianten), COSMO-EU, ICON-EU,

NowCastMIX, Mesocyclone Detection, Mesoanticyclone Detection, Rotation-Low-Level, Rotation-Mid-Level, VIL, VII, and corresponding track products.

Vorhersager: Robert Herrman, Katrin Hohmann, Martin Jonas, Thomas Schumann, Helge Tuschy, Bernd Zeuschner

Entwickler: Thomas Hengstebeck, Paul James, Susanne Theis, Kathrin Wapler

Expert lectures (onsite und remote): Thomas Hengstebeck, Paul James, Susanne Theis, Helge Tuschy, Kathrin Wapler

## Erwartungen / Anforderungen

- Visualisierung der Produkte zur effizienten Evaluierung
  - Absprache dazu im Vorfeld des Testbeds mit den Entwicklern, Rücksprache mit Entwicklern falls Änderungen zu den Absprachen eintreten; Dokumentation der ggf. daraufhin vereinbarten Änderungen
  - Darstellung der Grenzen / des Rahmens der Vorhersageprodukte (damit klar wird, wo das Modellgebiet aufhört und Fehlinterpretationen vermieden werden können (z.B. kein Niederschlag an einem bestimmten Ort da nicht vom Modell vorhergesagt, oder da außerhalb des Modellgebietes))
- Neutrale Präsentation der zu evaluierenden Verfahren und Produkte sowie der vorbereiteten Visualisierungsvarianten (ohne Wertung) für die Teilnehmer des Testbeds
- 
- Dokumentation der Evaluierungsergebnisse für die bereitgestellten Verfahren im anschließendem Evaluationsbericht

## Wünsche

- Aktive(re) zielgerichtete Förderung des Austausches zwischen Vorhersagern und Entwicklern
- Berücksichtigung der jeweiligen speziellen Expertisen und Erfahrungen der Teilnehmer

## Generelles Feedback der DWD-Teilnehmer (Testbed 2015)

- Testbed wurde als anregend und bereichernd wahrgenommen
- Austausch mit Entwicklern und Vorhersagern von unterschiedlichen Wetterdiensten wurde als positive Erfahrung beurteilt

- Diskussion über den „State of the art“ der Vorhersage von schweren konvektiven Ereignissen interessant und hilfreich

## Verbesserungsvorschläge

- Im Operations Plan und in einer Einladungsmail klar deutlich machen, dass eine Vorbereitung der Teilnehmer erforderlich ist. Dies kann z.B. enthalten, dass den Teilnehmern ein Arbeitstag an dem sie keine Verpflichtungen im operationellen Dienst haben zur Verfügung gestellt werden muss, um:
  - (a) den Operations Plan detailliert zu lesen (zu erstellende Vorhersagen und Warnkriterien einprägen, Kennenlernen der zu evaluierenden Produkte, etc.),
  - (b) sich mit dem Testbed Data Interface vertraut zu machen (z.B. an vorgegebenen Beispielfällen),
  - (c) die aufgezeichnete Vorlesung zum ingredients-based forecasting anzusehen und zu verstehen,
  - (d) ggf. weitere ausgewählte Vorträge aus den letzten Jahren anzusehen.
- Klarere Beschreibung der Aktivitäten im Testbed insbesondere mit Hinblick auf die Erwartungen an die Teilnehmer (Vorhersager wie Entwickler, ggf. unterschiedliche Erwartungen), z.B.
  - adäquate Vorbereitung ,
  - Teilnahme an den Diskussionen,
  - Evaluation der Produkte,
  - Präsentation der Vorhersage/Verifikation in der Gruppe,
  - Expertenvorlesung (Erwartungen an den Inhalt)
- Karten der verschiedenen Vorhersagegebiete (W, SW, S, SE, E, NE, D, Alps) mit Orographie, Ländernamen, und einigen geographischen Namen wären hilfreich, da nicht alle Teilnehmer im Detail mit allen Regionen vertraut sind. Die Karten könnten im Operations Plan enthalten sein, und/oder über das Testbed Data Interface verfügbar und/oder ausgedruckt und an den Vorhersage-Arbeitsplätzen aufgehängt werden.
- Für die Fern-Vorträge: kurze technische Anleitung und verbindliche Verabredung zum technischen Test (mit fester Uhrzeit und Angabe der Test-Session)
- Klarere Struktur in der Gruppenarbeit sowohl bei den Vorhersagen als auch bei den Evaluationsaufgaben: Zuständigkeiten und Aufgaben definieren, unsachliche Unterbrechungen ausschalten, unterschiedliche Erfahrungen/Expertisen der Vorhersager und Entwickler zielorientiert berücksichtigen
- Alle Daten die im Testbed Data Interface dargestellt sind, sollten eine Beschriftung / Legende / Farbtabelle haben; für nicht so gängige Parameter wäre eine kurze Erklärung hilfreich
- Vorhandene NinJo-Workstation läuft sehr langsam, Es ist zu klären woran dies liegt (Datenverbindung zum Server, geringer Rechner-Power?), Sollten die Performanzprobleme bestehen bleiben ist es besser, den NinJo-Rechner während des Testbeds nicht einzusetzen.)
- Für die Bodenanalyse wäre das vollständige Bodeneintragungsschema hilfreich; außerdem müssten die Stationseintragungen größer (Schriftgröße) dargestellt werden, um eine bessere Auswertung zu ermöglichen