ESSL Testbed

Report on the evaluation of DWD nowcast and warning products at the ESSL Testbed 2023

Combined final report of:

DWD-Auftrag 3069293/23-TRA

„Werkvertrag über ein Testbed der DWD Vorhersage- und Warnprodukte “

and

DWD-Auftrag 3070927/23-MEY

“Erstellung einer detaillierten Evaluierung des SUNFONY Rapid Update Cycles und seinem mit dem KNORAD3D-EPS kombinierten Zell-Objekt-Produkt im Rahmen des ESSL Testbeds 2023“

ESSL Report 2023/02

Authors: Tomáš Púčik, Christoph Gatzen and Pieter Groenemeijer

European Severe Storms Laboratory – Science & Training
Bräunlichgasse 6a/6
2700 Wiener Neustadt
AUSTRIA

Contents

1 Introduction

1.1 This report

This report summarizes the findings of the evaluation of DWD nowcast and warning products at the ESSL Testbed 2023. It condenses the feedback collected from discussions at the Testbed, the feedback collected through questionnaires, and of other input by the ESSL Testbed team. The direct written feedback of Testbed participants is included as an annex to this report.

1.2 The ESSL Testbed 2023

The summer ESSL Testbed 2023 took place during the weeks of 5 – 9 June, 12 June – 16 June, 26 – 30 June, and 3 – 7 July. This was the 12th edition of the ESSL Testbed, a series of testbeds that began in 2012. The third week was reserved for more expert participants and was done in a hybrid mode, while all other weeks were done in the purely onsite format. The first and the second week were co-organized together with EUMETSAT, who sponsored the participation of forecasters at the Testbed. Participants ranked the Testbed with an average grade of 9.45 out of 10. The 2023 version of the ESSL data displayer brought an upgrade. To better facilitate the evaluation process, three different modes of data visualization, “verification”, “analysis” and “forecast” mode were introduced. For example, verification allows to combine the forecast/nowcast product issued at a particular time in history with future radar/satellite/ESWD data.

Altogether 48 participants took part in the Testbed in addition to 5 ESSL staff members. The participants came from 23 different countries: Austria, Croatia, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Latvia, Lithuania, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Switzerland, Turkey, United Kingdom. Some of the participants were not forecasters but developers of the products evaluated at the Testbed. Overall, 8 DWD personnel took part: Matthias Gäßl, Sebastian Balders, Franka Nawrath, Sophie Löbel, Nikolaos Antonoglou, Matthias Gottschlack, Lukas Josipovic, Reinhold Hess.

1.3 Testbed Resources

The following online resources contain further information about the Testbed:

The Testbed Data Interface showing all products and all data, is available online after the end of the Testbed at: https://weather.essl.org/wx/displayer.php

Username: testbed
Password*: hailstorm2020

These credentials may expire. Please contact pieter.groenemeijer@essl.org to request an updated password in case the credential do not work.

1.4 Feedback

Feedback on the products was collected throughout the Testbed period, partly **i) in direct discussions with the on-site participants**, and in part **ii) through the documentation of answers to questionnaires** that were filled out jointly by participants, who typically worked in groups of 2-4 persons in dedicated online sessions during the afternoons. The direct feedback from participants collected through questionnaires has been attached to this report.

The evaluation concerned five different products: ICON-RUC, Radar Maxima, KONRAD3D-EPS, KONRAD3D-SINFONY and NowCastMIX. NowCastMIX is evaluated as an operational product in comparison with the KONRAD3D-related products. Attention has been given to the comparison with verification data from various sources, such as surface observations, ESWD reports, as well as satellite imagery and radar data.

2 ICON-RUC

ICON-RUC (RUC hereafter) and its ensemble is being developed as part of the effort in the Seamless Integrated Forecasting System (SINFONY) project. The model is a vital part of the seamless nowcasting system up to 12 hours ahead. The goal of RUC is to provide accurate forecasts of precipitation fields, including convective storms, at times when nowcasting systems using observation systems quickly decreases in accuracy. RUC has the same resolution and vertical levels as the ICON-D2 model setup. RUC is run every hour with faster computation compared to the ICON-D2 to allow forecasters to access the data as early as possible. The three main differences to the ICON-D2 are:

1. Aggressive hourly data assimilation of observations, including radar and satellite data.
2. Usage of a 2-moment microphysical scheme that includes a separate hail hydrometer class and a radar reflectivity output based on Mie-Scattering.
3. More frequent output, including 5-minute intervals for the simulated radar reflectivity variable[1].

2.1 Visualisation

RUC data visualisation was similar to that of the D2 data. Ensemble data were displayed using "paintball" plots, the median and the maximum value per grid point, and the fraction of members exceeding a pre-set threshold per grid point and for an upscaled area. Compared to the D2, more variables were available as direct model output of the RUC, including updraft helicity in the 0-3 km layer, the 18 dBZ echo top height, and total column hail.

[1] In 2023, only 15-minute intervals were shown in the Displayer.

2.2 Conclusions

The evaluation of ICON-RUC led to the following main findings:

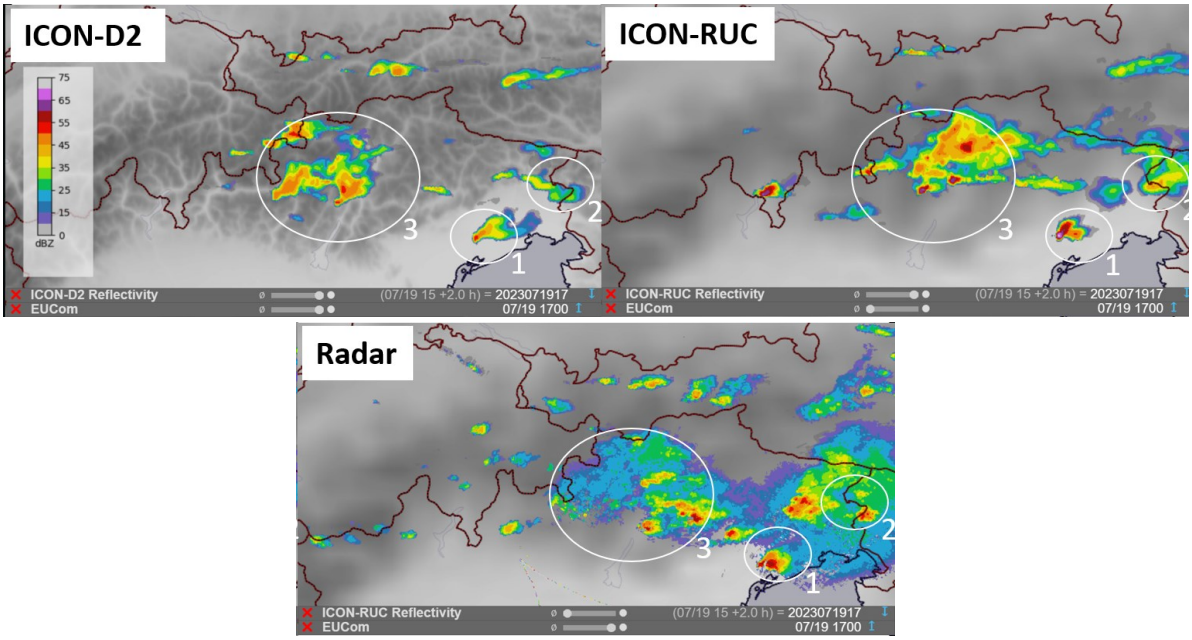
- ICON-RUC was well-received by the forecasters at the ESSL Testbed 2023. Participants appreciated the hourly updates of the model and considered the forecasts significantly more skilful than the storms simulated by ICON-D2.
- The distribution of reflectivity within convective systems deviated from the radar, as the highest reflectivity within storms was often higher and included in relatively small intense cores, i.e. smaller and more intense than on radar. This had been similar in 2022.
- In some cases, RUC was able to reproduce the life cycle of supercells and convective systems well. In a few cases supercells that were assimilated decayed too fast. It may be that the fact that intense cores of isolated storms were often smaller than observed by radar contributed to their quick decay.
- Compared to ICON-D2, RUC is not as reluctant in developing convection. The initiation of storms was either on time or only slightly delayed compared to reality. Only a few times, it was too early. When comparing many cases, the probability of storms occurring in the ensemble was usually about right or possibly slightly too low. In summary, ICON-RUC EPS performed better than ICON-D2 EPS with respect to storm initiation, as it did not suffer (as much) from late or absent initiation.
- Areas within the cold pools of both supercells and convective systems were often somewhat colder than surface observations in the RUC, whereas ICON-D2 forecasts tended to be too warm.
- The life cycle and the occurrence of larger convective systems and bow-echoes was typically captured well when the model had already assimilated the existing storm system.
- The performance of updraft helicity and reflectivity corresponded better with supercell occurrence in ICON-RUC(-EPS) than in ICON-D2(-EPS).
- Wind gust speed was higher in ICON-RUC than in ICON-D2. Both models often produced too strong wind gusts, but this was worst for ICON-RUC.
- Stratiform precipitation areas in ICON-RUC are more accurate, i.e. larger, in extent than in ICON-D2 which strongly underestimates their extent.

2.3 Discussion illustrated with cases

2.3.1 Does RUC simulate more realistic reflectivity patterns of convective storms compared to the ICON-D2? Think about squall lines, supercells, updraft cores, stratiform regions, and their shapes.

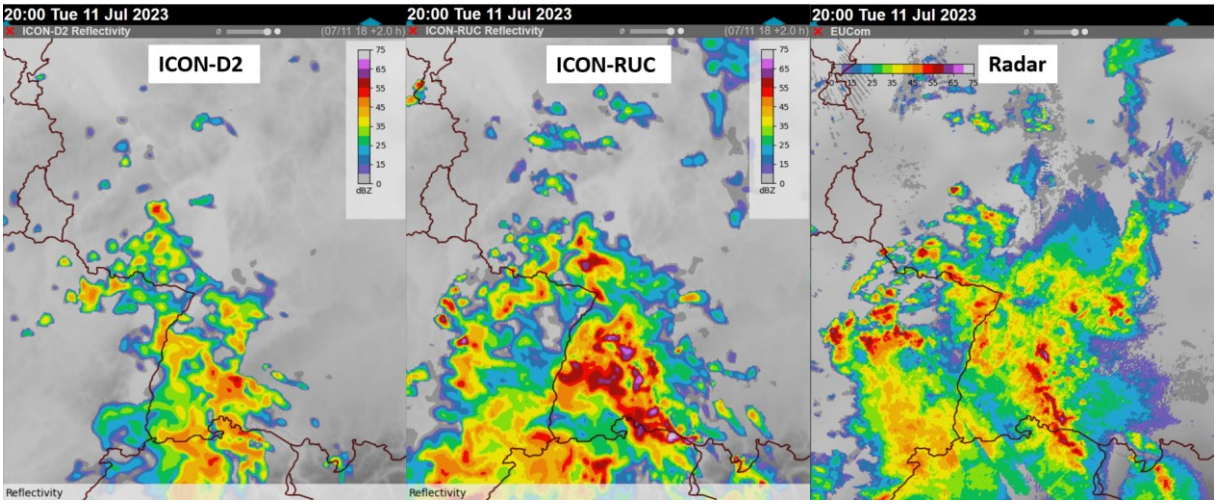
Participants rated this question using three possible answers: “no”, “a bit more realistic” and “much more realistic”. Eight groups answered this question, five of them answering “much more realistic” and 3 “a bit more realistic”, yielding an average score of 2.6 out of 3. This result is even slightly above that of last year’s Testbed. Some groups have noted that the reflectivity areas were more intense than in ICON-D2, and than radar observations. Additionally, RUC was better than ICON-D2 in showing the correct storm mode in some cases. We illustrate this with two examples, one of isolated supercells, one with a larger convective system.

On 19 July 2023, isolated supercells moved across northern Italy, producing very large hail. A comparison of the simulated reflectivity of ICON-RUC and ICON-D2 with the radar data indicates that ICON-RUC is better in capturing the overall structure of the cells. In particular, the cores of the precipitation areas look better since they are rather small with tight reflectivity gradients. However, stratiform areas also looked better on average: ICON-D2 produced rather large and intense forward-flank downdrafts for the cell in area 1 and the two most prominent southern cells in area 3 compared to ICON-RUC and the radar display. In area 2, cells are embedded in stratiform precip that is better reproduced by ICON-RUC as well.



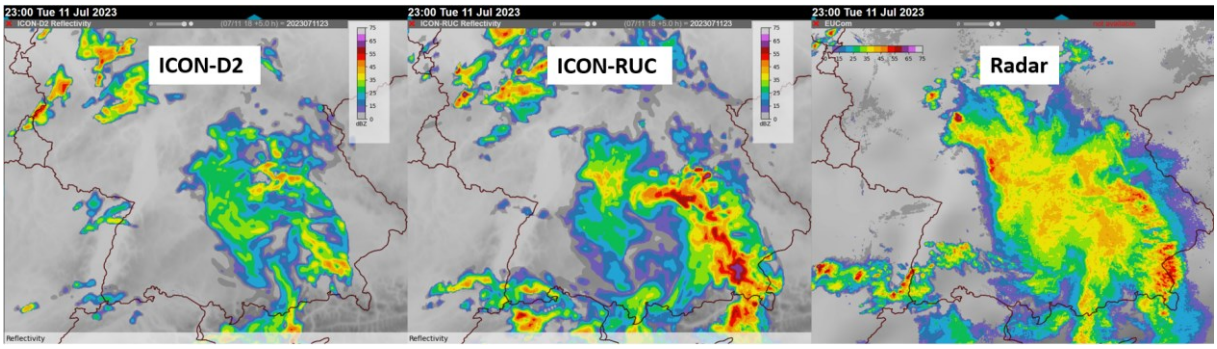
19 July 2023 15 UTC ICON-D2 and ICON-RUC simulation of reflectivity at 17 UTC compared to the radar observations.

In the case of 11 July 2023 across southern Germany, both ICON-D2 and ICON-RUC simulate the transition from isolated supercells to a large convective system. The system looks considerably weaker in ICON-D2 compared to the radar display, especially for the the storm cores of embedded (super)cells and the leading squall line, but additionally parts of the stratiform precipitation region, e.g., across eastern France. ICON-D2 has a significantly too weak, small, and unorganized reflectivity structure for this case. It would be difficult to anticipate a severe squall line based on the simulated radar display of the ICON-D2.



11 July 2023 18 UTC ICON-D2 and ICON-RUC simulation of reflectivity at 20 UTC compared to the radar observations.

Later in the development of the mesoscale convective system into a severe bow echo, the forecast of the structure and movement speed, intensity of the stratiform precipitation, and trailing convective cells across eastern France and south-western Germany is significantly better for the ICON-RUC compared to ICON-D2 as indicated by a 5-hour-forecast shown below. In this case, ICON-RUC overestimates the reflectivity intensity of the precipitation cores while ICON-D2 massively underestimates it.

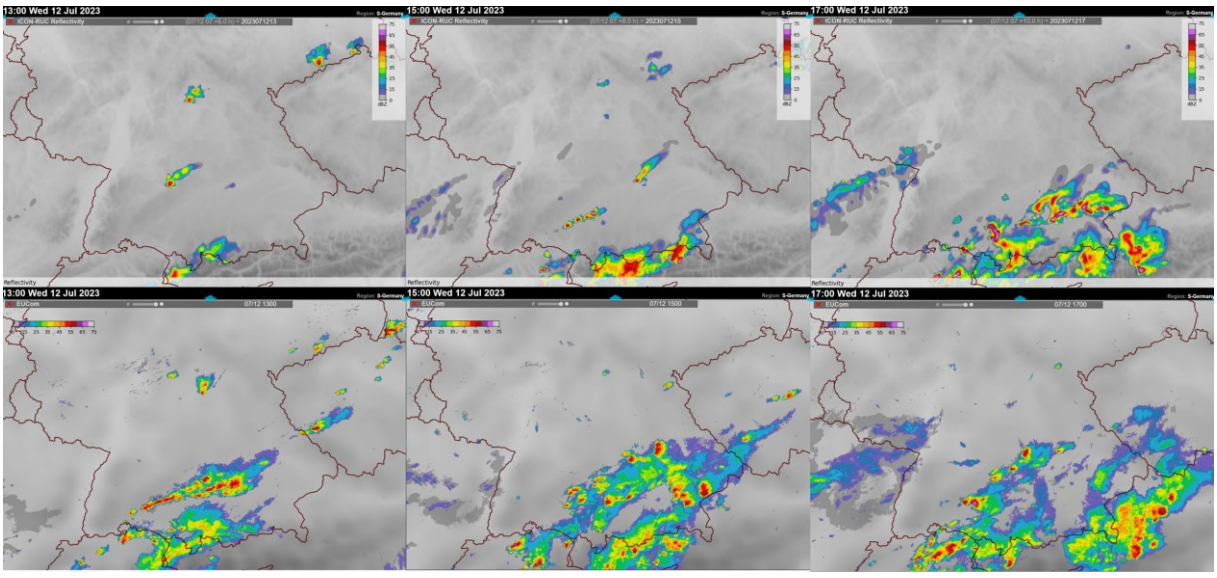


11 July 2023 18 UTC ICON-D2 and ICON-RUC simulation of reflectivity at 23:00 UTC compared to the radar observations.

2.3.2 Compare the timing of the storm initiation in the RUC to reality.

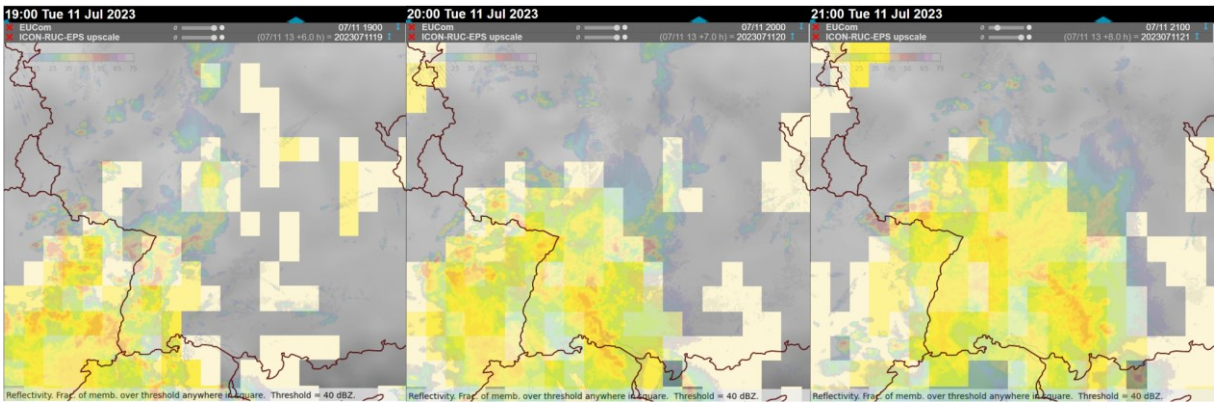
Twelve events were rated by the groups that decided between the possible answers “2+ hours too early”, “1 hour too early”, “On point”, ”1 hour too late”, and “2+ hours too late”. In eight cases the initiation was forecast exactly “On point,” in two cases it was “1 hour too late” and in one case the delay was “2+ hours too late”. In one case, initiation was “1 hour too early”. The answers show that the ICON-RUC initiation is almost on time. Some groups mentioned that ICON-RUC failed to produce storms in a situation with Föhn winds close to the Alps, another group found that ICON-RUC initiated storms that did not develop, and there were both observations of storms forming a little too early, and of them forming a little too late. The overall impression is that ICON-RUC has a rather balanced handling of storms initiation and no strong bias regarding too early or too late initiation. Some groups also mentioned that ICON-RUC performed better compared to ICON-D2.

An example when ICON-RUC was too late with initiation is given below. On 12 July 2023, 6-, 8-, and 10-hour forecasts of the ICON-RUC is compared to the radar image. In the situation, storms developed along a linearly oriented convergence zone. Looking at the simulated radar display of the ICON-RUC forecast, the general nature of developing convection is reproduced quite accurately. In particular, the linear nature of the convection is evident. However, it looks like the forecast lags behind the real development by one to two hours: At 15 UTC, ICON-RUC predicts the initiation of a line of storms across southern Germany. This line had already developed at 13 UTC. The same time lag is visible 2 hours later.



12 July 2023 07 UTC ICON-RUC forecasts of simulated radar reflectivity at 13, 15, and 17 UTC (upper figures) compared to the radar observations at the respective times (bottom figures).

Another example illustrates the timing of convection initiation along the leading gust front of a cold pool in association with a severe bow echo. On 11 July 2023, ICON-RUC EPS captured the timing of the initiation of the bow echo across southern Germany. Additionally, the further development and movement speed of the bow echo is highlighted by the upscaled probability of the radar reflectivity pretty well.

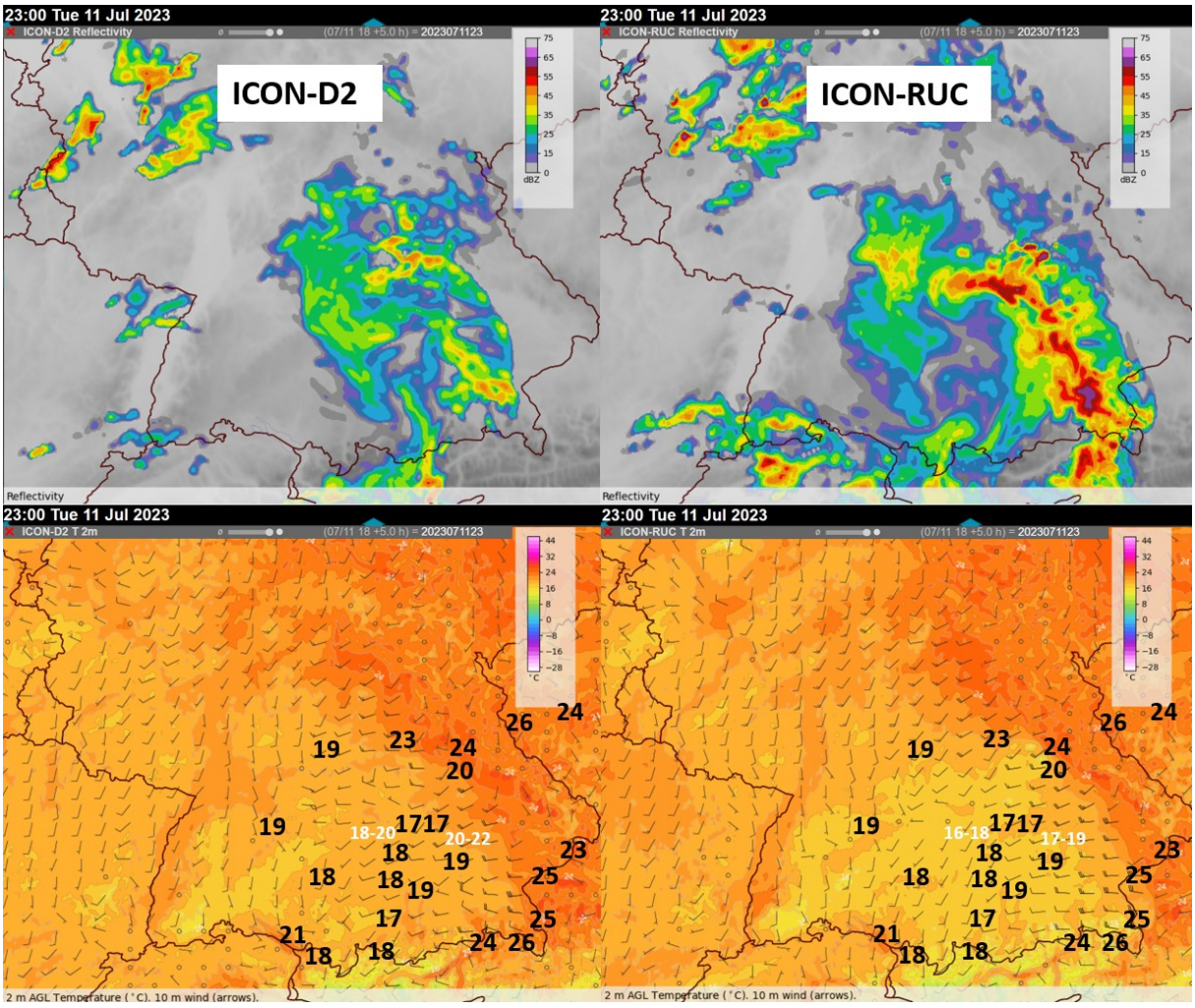


11 July 2023 13 UTC ICON-RUC EPS probability of reflectivity exceeding a threshold (40 dBZ) at 19, 20, and 21 UTC overlayed to observed radar reflectivity at the respective times (second layer).

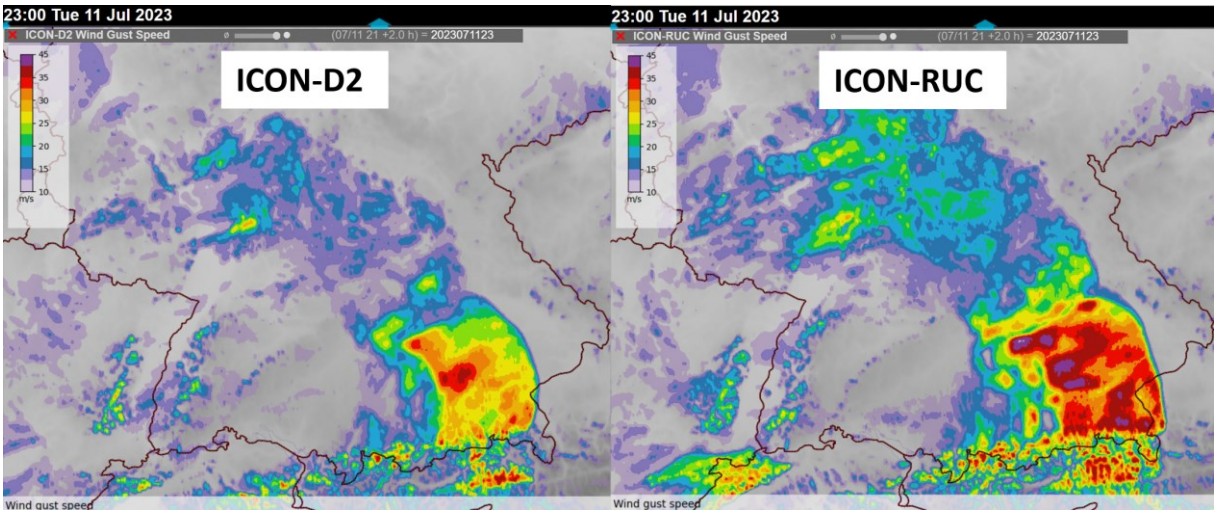
2.3.3 For cold-pool-driven storms, can you say something about the simulated cold pools in comparison to observations)?

Participants were asked to compare the cold pools of cold-pool driven storms of the ICON-RUC forecasts to reality. They rated the cold pool intensity using three possible answers: “Too weak (in terms of temperature drop and wind gust speed)”, “About right”, “Too strong”. Six groups answered this question, three of them answering “Too strong” and three “About right”. A second question referred to the speed of the cold pool in the simulation. Comparing it to the observations, participants addressed ICON-RUC using the possible answers “Too fast moving”, “Moving at correct speed”, “Too slow moving”. Seven groups answered the question, with six choosing “Moving at correct speed” and one “Too slow moving”. In general, it seems that ICON-RUC simulated cold pools tend to be a bit colder than observed while the cold pool speed was not found to differ much from reality. In the case of 11 July 2023, a large cold pool driven system evolved in the evening across southern Germany. It formed a fast-moving, severe bow echo during the night. The speed of the gust front was about right in the ICON-RUC forecast, whereas it was too slow for the ICON-D2. Additionally, the ICON-RUC’s cold pool temperature was about right or about 1 K colder compared to observations. Instead, ICON-D2 was about 2 K too warm.

Regarding the wind gust speed, we investigate this situation a few hours later, at 23 UTC. The ICON-RUC simulation of the forecast at 21 UTC was too aggressive as it produced a large area of wind gusts exceeding 35 m/s. The maxima were even higher with 45+ m/s. Observations during that time period were about 27 to 33 m/s, with the highest wind gust measured at 38 m/s. With respect to wind speed thus, the ICON-RUC cold pool was too strong. ICON-D2 has a better wind gust estimate which is close to the observed values.

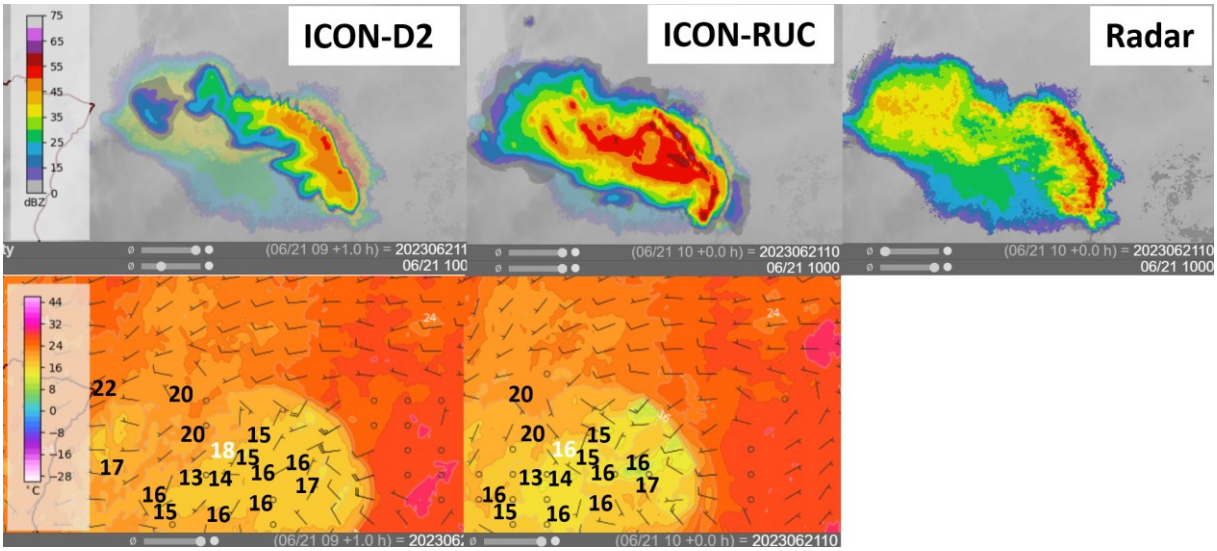


11 July 2023 18 UTC forecasts of simulated reflectivity (upper figures) and 2 m AGL temperature (bottom figures; colour shading and white labels) at 23 UTC from the ICON-D2 (left) and ICON-RUC (right) runs combined with surface observations of 2 m AGL temperature (black numbers at surface station locations).



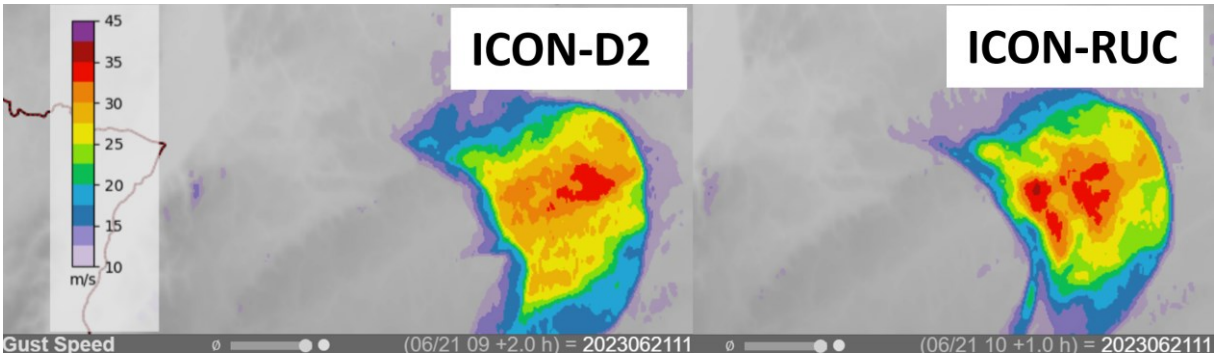
11 July 2023 21 UTC forecasts of wind gust speed at 23 UTC from the ICON-D2 (left) and ICON-RUC (right).

Another example is the severe bow echo that formed in the morning of 21 June 2023. The better performance of ICON-RUC is obvious when comparing the size, reflectivity intensity, and position of the bow echo and trailing stratiform precipitation of the simulation to the radar image. ICON-D2 tends to produce a system that is weaker, smaller, and slower compared to the observed event. Looking at the temperature at 2 m AGL, ICON-RUC indicates a broad area with a temperature between 14 and 16°C. Observations on that day indeed confirmed these values, and at numerous places, the temperature dropped to below 16°C within the cold pool. It must be noted that inside the smaller precipitation cores, ICON-RUC is even colder with a temperature down to 10°C. That was not observed by the station network. On the other hand, ICON-D2 is too warm with a cold pool temperature between 16 and 18°C. This was significant within the trailing stratiform precipitation region, where the observed temperature was 13 to 16°C. Here, ICON-D2 is just below 18°C (2 to 5 K too warm), ICON-RUC is below 16 to less than 14°C (close to observed values).



21 June 2023 09 UTC forecasts of simulated reflectivity (upper figures) and 2 m AGL temperature (bottom figures; colour shading and white labels) at 10 UTC from the ICON-D2 (left) and ICON-RUC (center) runs combined with surface observations of 2 m AGL temperature (black numbers at surface station locations). The radar image at 10 UTC is displayed in the upper right corner of the figure and as a second layer in the radar reflectivity simulations of the forecasts for reference.

Wind gusts of this case were analysed between 10 and 11 UTC. Here, both model forecasts result in a cold pool that is too strong: ICON-D2 09 UTC run simulates maximum gusts between 30 and 35 m/s, ICON-RUC 10 UTC run is a few m/s stronger. However, observations were only around 25 m/s during that time period.

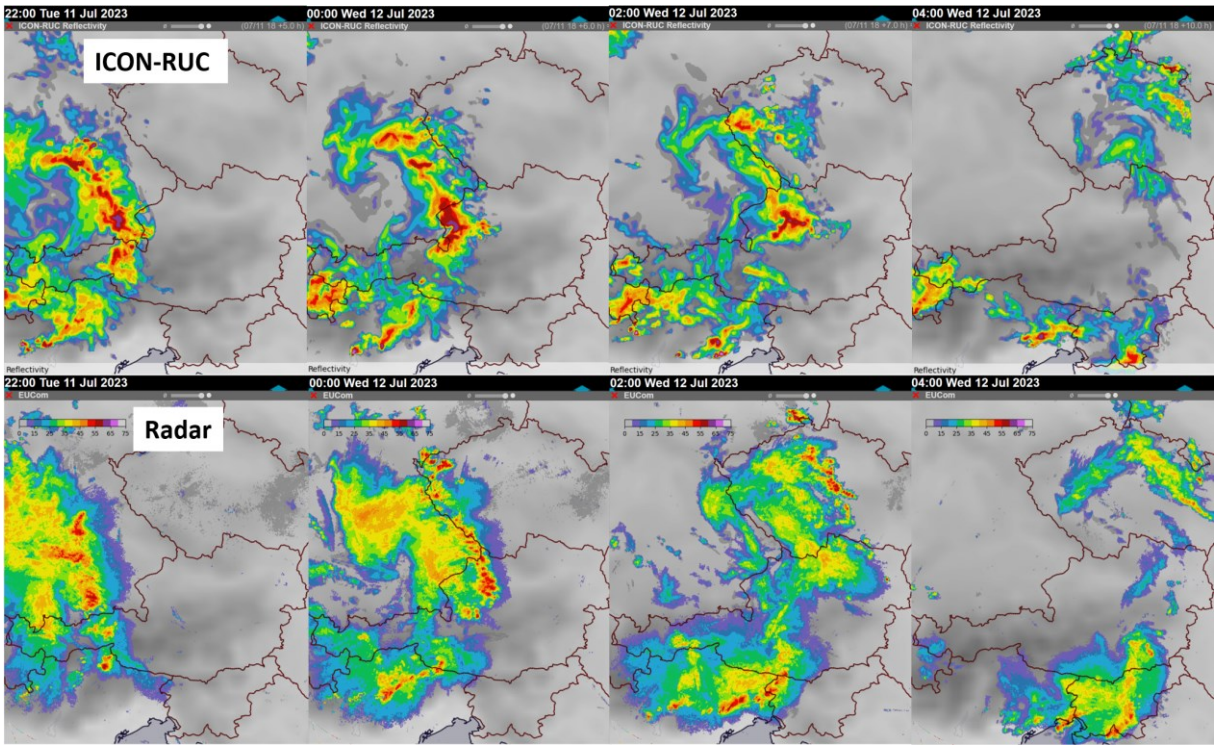


21 June 2023 forecasts of wind gust speed between 10 and 11 UTC from the ICON-D2 (09 UTC; left) and ICON-RUC (10 UTC; right).

2.3.4 Does RUC correctly forecast the duration (life cycle) of storms?

This question was answered by eight groups, using the possibilities “Single cells decay too fast”, “Single cell life cycle simulated correctly”, “Single cells persist too long”, “Squall lines decay too fast”, “Squall lines life cycle simulated correctly”, “Squall lines persist too long”, “Supercells decay too fast”, “Supercells life cycle simulated correctly”, “Supercells persist too long”. For squall lines, four groups indicated that a squall line life cycle was simulated correctly, and one group found a squall line that persisted too long. With respect to supercells, four groups indicated that the supercell life cycle was simulated correctly, and one team noted that a supercell persisted too long. Finally, the life cycle of all analysed single cells (six groups) was found to be correct. This outcome indicates a good performance of ICON-RUC with respect to the life cycle of single cells, supercells, and squall lines.

As an example, a severe squall line that developed on late 11 July 2023 across southern Germany was predicted with high performance. The mature stage of the squall line at 00 UTC on 12 July was well captured by the 6-hour forecast by ICON-RUC. Additionally, the time of decay between 02 and 04 UTC is in good agreement with the forecast of the ICON-RUC.



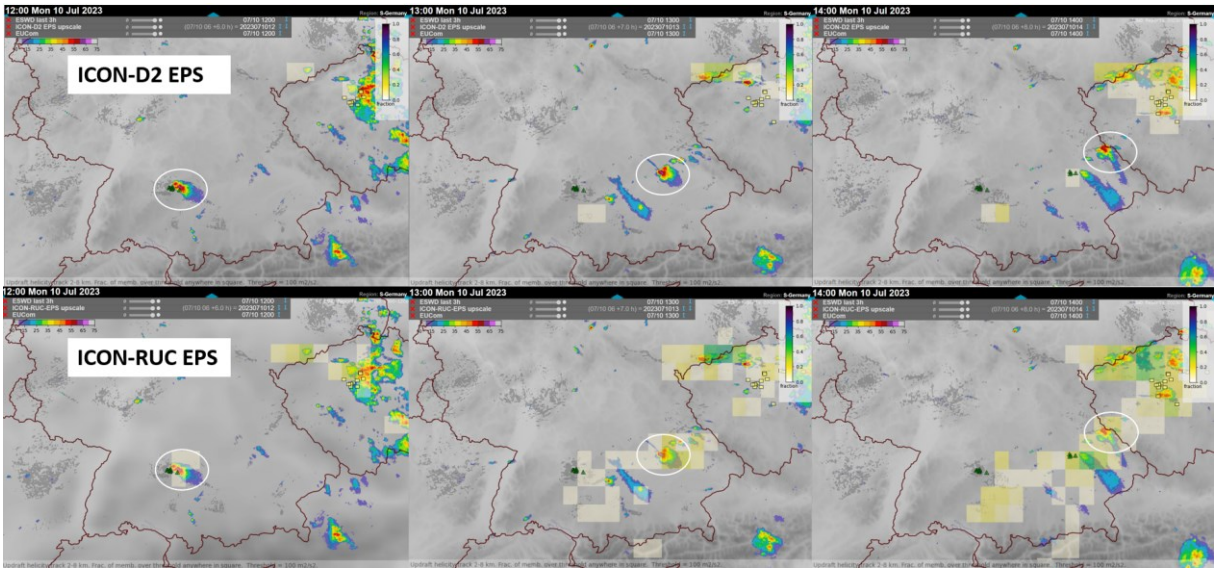
11 July 2023 ICON-RUC simulations of reflectivity at 23, 00, 01, and 04 UTC (upper panel) and the observed reflectivity at 22, 00, 02, and 04 UTC (lower panel).

2.3.5 Are the ensemble forecasts of the RUC reliable (reflectivity, updraft helicity)? In the case of intense and small-scale convection, we have to expect low exceedance probabilities for the higher thresholds over a wider region, especially for longer lead times. Do highlighted regions cover the observed cells?

Eight groups answered this question, with the options "Too many misses (convection outside of region where ensemble simulated some probability of storms)", "Sensible balance of misses, hits and false alarms", and "Too many false alarms (no storms within the region where ensemble simulated some probability of storms)". All groups answered this question with "Sensible balance of misses, hits and false alarms". This result indicates the good performance of ICON-RUC EPS with respect to the development of storms during this year's Testbed.

One group mentioned that the EPS forecast was more reliable compared to the deterministic forecast and another group noticed that the probabilities were higher for shorter lead times. An example for the updraft helicity track is given below for a case across southern Germany on 10 July 2023. Here, ICON-D2 EPS probabilities are confronted with ICON-RUC EPS probabilities. As another layer, the radar reflectivity and severe weather reports of the ESWD are plotted. It can be seen that ICON-RUC EPS highlights the potential of supercells in the 6-, 7-, and 8-hour forecast in the right locations, although there is a tendency to an overestimation in the 7- and 8-hour forecast. Instead,

ICON-D2 EPS failed to produce intense storms, what is also reflected in the probability of the updraft helicity track. The group that analysed this case remarked that it was quite tricky to forecast.



10 July 2023 06 UTC ICON-D2 EPS and ICON-RUC EPS forecasts of updraft helicity tracks at 12, 13, and 14 UTC compared to the radar observations and ESWD data (last three hours until time of radar image).

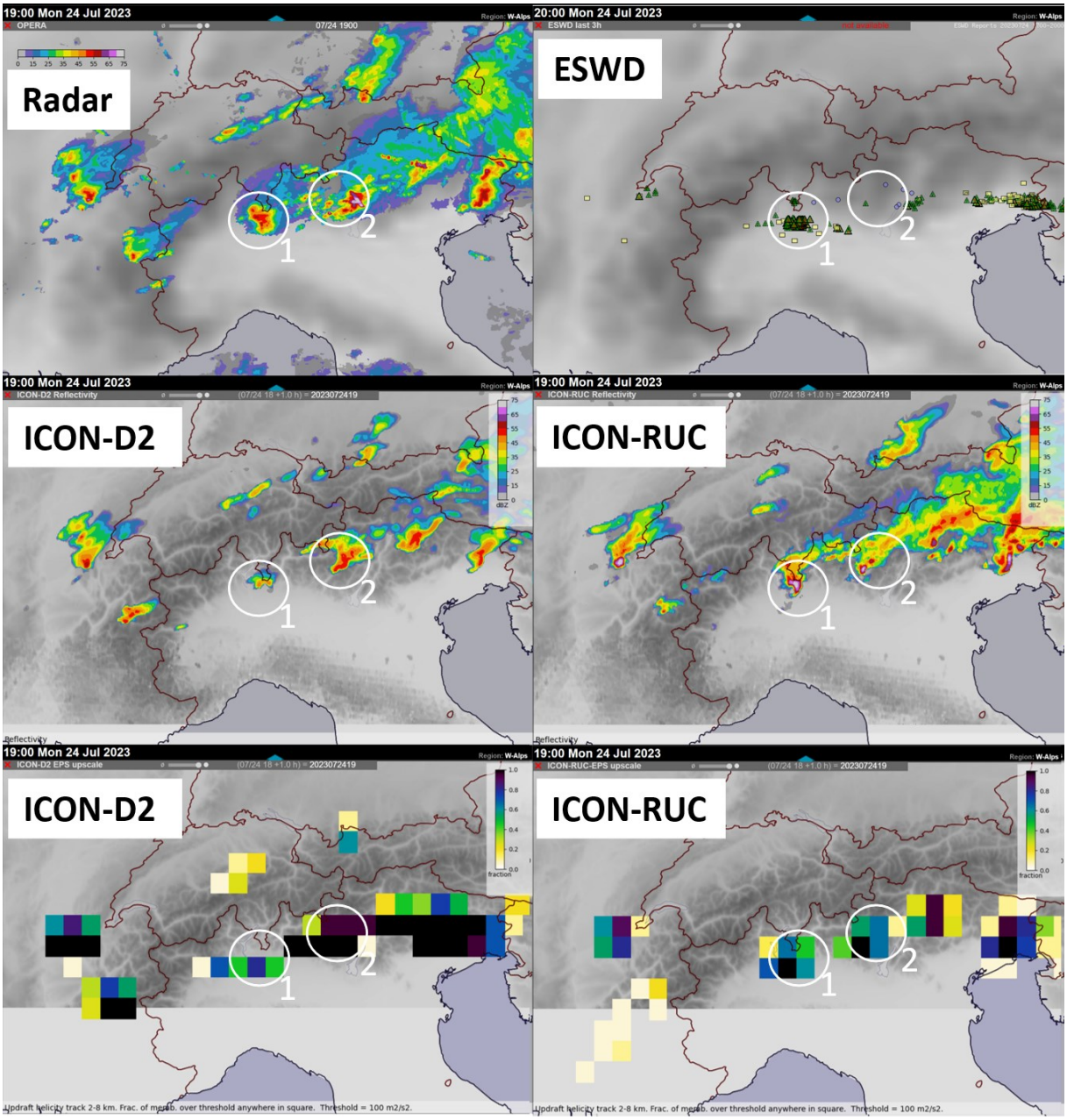
2.3.6 Have you noticed a large difference in a specific parameter (such as updraft helicity or reflectivity) in RUC that is clearly better or worse than ICON-D2?

Eight groups answered this question. Regarding the simulated radar reflectivity, there was a very clear preference to the ICON-RUC forecasts that were rated to be much better than ICON-D2. It was mentioned that convection initiation occurs more frequently in ICON-RUC. Altogether, five groups mentioned this parameter to be better in ICON-RUC forecasts.

The updraft helicity was rated by three groups. Two of them mentioned that ICON-RUC forecasts looked better compared to ICON-D2. One group saw no significant differences and considered both forecasts as “good”. One group additionally mentioned that the ICON-RUC had higher updraft helicity compared to the ICON-D2.

A case is illustrated in the following figure. On 24 July 2023, gargantuan (i.e. ≥ 15 cm diameter) hail was observed in northern Italy. Supercells that moved over the area in the evening were forecast by ICON-D2 and ICON-RUC using the 2-8 km updraft helicity tracks. While both ensembles clearly highlight the potential of the eastern most cell, the supercells in areas 1 and 2 indicate some differences. The cell in area 1 was more intense according the ESWD data that the cell in area 2. ICON-D2 instead indicates a higher potential in area 2 compared to area 1. ICON-RUC performs better here, clearly indicating

the higher potential in area 1. That said, ESWD data may not reflect the hail occurrence of the cell in area 2, though.



24 July 2023 18 UTC ICON-D2 EPS and ICON-RUC EPS upscaled probability forecasts of updraft helicity tracks at 19 UTC compared to the radar observations and ESWD data (last three hours until time of radar image).

2.3.7 In the actual weather situation that you look at, does the rapid hourly forecast update help in any way to give more precise warnings?

Seven groups answered the question, rating the value of ICON-RUC forecasts according to the possible answers “Yes, but only for shorter lead times (up to 3h)”, “Yes, but only for

longer lead times (beyond 3h)", "Yes, both for shorter and longer lead times", "No". Five groups answered the question with "Yes, both for shorter and longer lead times". This reflects a great value of the rapid update cycle. In particular, the groups motivated their answer with the argument that ICON-RUC was able to correctly predict the intensity of expected convective events quite soon, in this case a weaker development was correctly forecasted. A group mentioned a clear positive effect of ICON-RUC on their forecast. Another group noted that ICON-RUC correctly confirmed the region of highest severe weather risk. Altogether, these answers reflect very positive feedback of ICON-RUCs performance.

Two groups answered the question with "Yes, but only for shorter lead times (up to 3h)". One of these groups argued that the best results occurred in the first two hours that followed the model update.

3 Radar Maxima

The product Radar Maxima was subject to evaluation at ESSL Testbed 2023. The aim of this innovative product is to provide a good overview of the current precipitation especially in situations where convection may occur. The Radar Maxima product tries to summarize the individual forecasts of the ensemble members in a well-arranged way in order to estimate most probable locations of precipitation and their intensities rather quickly. In general, the idea of the product is to use maximum precipitation in the surrounding of a given location as a forecast variable. The Radar Maxima are derived as 95%-quantiles of the precipitation intensities within a surrounding of 40 km around each grid point. Current thresholds intensities are 1 mm/h, 2 mm/h, 5 mm/h, 10 mm/h, 15 mm/h, 25 mm/h and 40 mm/h. Statistical calibration of the probabilities is performed using a model output statistics (MOS) approach with logistic regression that has been tailored to probabilistic ensemble forecasting.

3.1 Conclusions

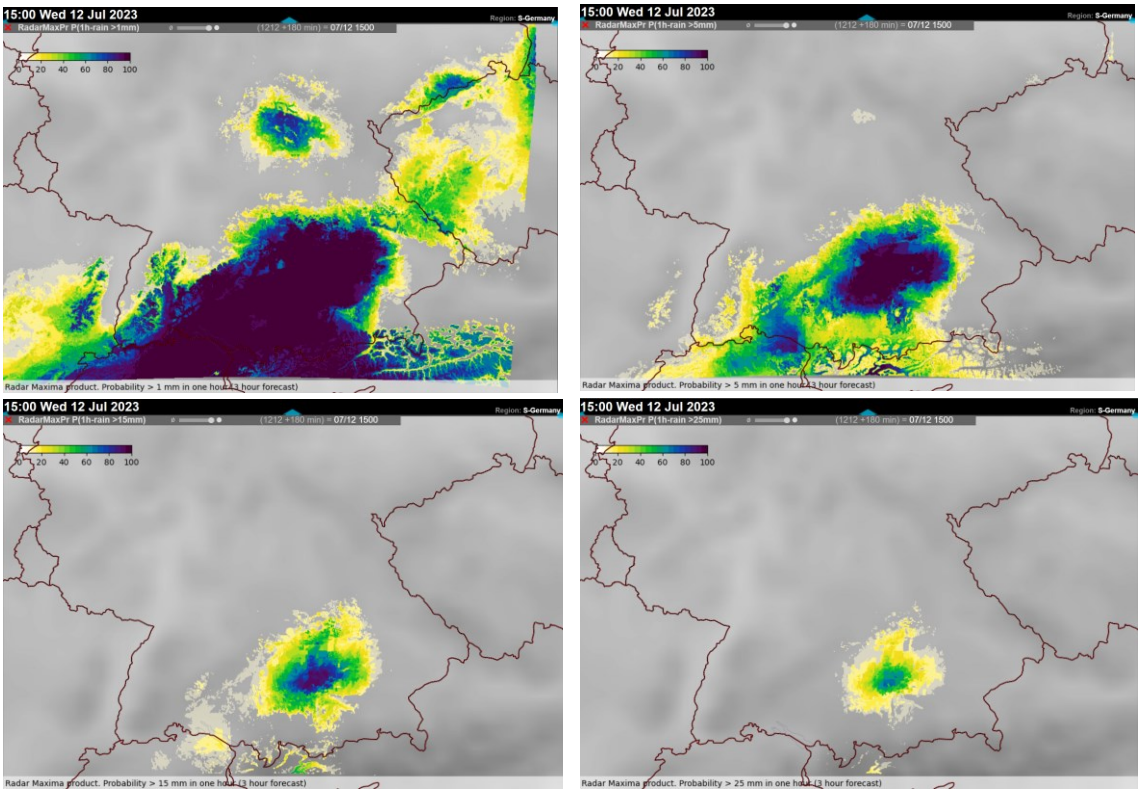
- Radar Maxima was rated to be useful in some cases, but its performance varies. In some of the analysed convective events its dependence on an ICON-D2 forecast that failed to initiate storms degraded the quality of the product.
- On average, Testbed participants answered that Radar Maxima underestimated precipitation, although the sizes of the affected area were rated to be OK. Compared to radar observations, the Radar Maxima precipitation areas were in some cases too large.
- Testbed participants agreed that there is some value using the Radar Maxima forecasts. The ICON-RUC forecast outperformed the Radar Maxima forecast when ICON-D2 EPS was too reluctant initiating storms.
- Finally, the influence of the topography on the Radar Maxima product looks too large: a pattern reflecting the topography was often recognizable in the product. A direct evaluation to rain measurements has not been done.

3.2 Visualisation

The Radar Maxima product is visualized using the probability of expected precipitation intensity exceeding various threshold in mm/h.

- Probability of the 95-quantile intensity of precipitation (1, 2, 5, 10, 15, 25, 40 mm/h)

At the ESSL Testbed, predictions were available every hour for a range of up to 24 hours at 1-hour intervals.



An example of the forecast of Radar Maxima for the probability of exceeding 1, 5, 15, and 25 mm in 1h.

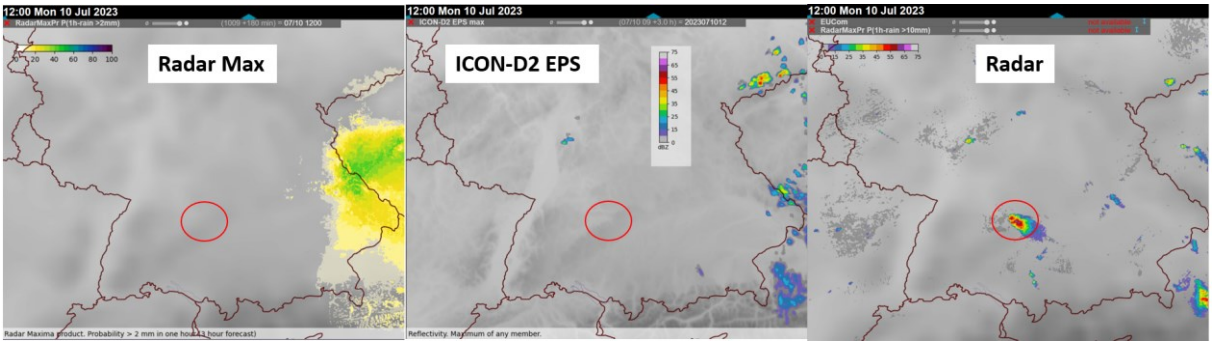
3.3 Discussion illustrated with cases

Participants were asked if the product provides a useful overview of the precipitation forecast. They were also asked if it is more useful in convective or stratiform situations. Six groups filled in the evaluation forms and decided between the answers “not at all / in some cases / in most or all cases”. All groups answered with “in some cases”. This answer is also reflected in the comments regarding the question, if the product is more useful in convective or stratiform situations: Most groups noted that they recognized varying forecast performance in the same situation for different forecast areas. Three groups found that Radar Maxima indeed worked good in convective situations, whereas one group did not find a significant difference to the performance in stratiform precipitation.

One group explicitly recognized that in situations with weak forecast performance by ICON-D2 EPS, the Radar Maxima product does not provide good forecasts either. In the case investigated, ICON-D2 EPS failed to forecast convection initiation what had an influence on the forecast of the Radar Maxima.

An example is given below. The forecast of 10 July 2023, 09 UTC for 12 UTC was considered to underestimate the precipitation intensity across southern Germany. Indeed, in the

highlighted region, rain measurements indicated intensities of more than 10 mm per hour locally.



10 July 2023, 09 UTC Radar Maxima forecast of the probability of precipitation intensity exceeding 2 mm in one hour at 12 UTC across southern Germany (left), 09 UTC ICON-D2 EPS maximum reflectivity forecast of any member for this time, and radar display at 12 UTC.

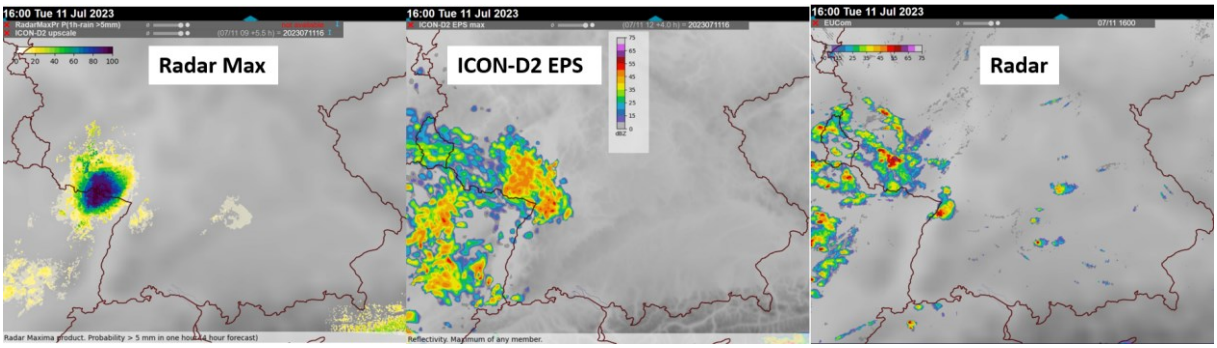
As an example of a forecast that was rated to be useful the case of 11 July 2023 is illustrated below. A large mesoscale convective system was moving eastward that was captured quite well by ICON-D2. Even though the radar reflectivity simulation by ICON-D2 looks less intense compared to the radar display, the Radar Maxima adds some value as it correctly highlights the area where more than 10 mm rain was measured in one hour.

11 July 2023, 18 UTC Radar Maxima forecast of the probability of precipitation intensity exceeding 10 mm in one hour at 23 UTC across southern Germany (left), 18 UTC ICON-D2 forecast of the reflectivity for this time, and radar display at 23 UTC.

In the next question, participants were asked if the coverage of the precipitation forecasts was meaningful. Six groups answered the question with the three option “Coverage far too small”, “Coverage OK”, “Coverage far too large”, with four groups deciding for “coverage OK”, one went for “coverage far too large”, and another for “coverage far too small”. On average, the forecast coverage is balanced according to the questionnaire.

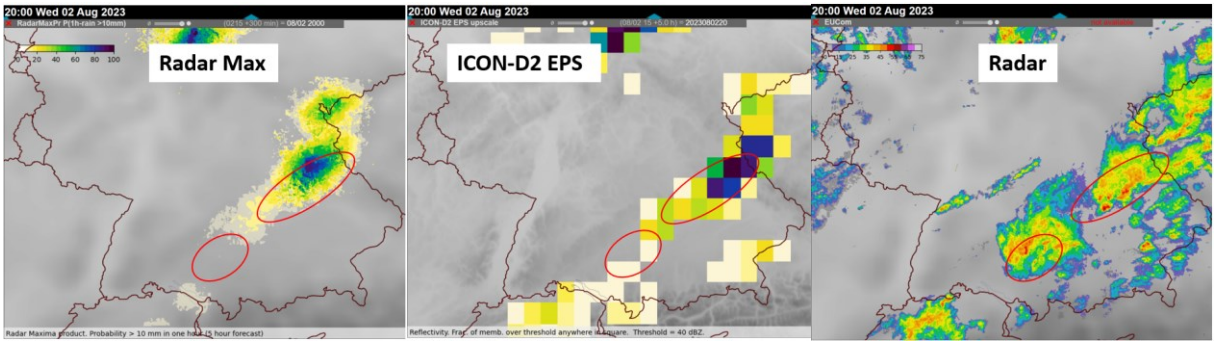
An example of a forecast of convective cells is presented below. A cluster of storms has developed across south-western Germany. It is represented by the maximum reflectivity of any member of the ICON-D2 EPS. Additrionally, the Radar Maxima highlights this area

by a increased probability of a precipitation intensity of 5 mm per hour. However, comparing the radar display to the Radar Maxima product, the convective nature of small precipitation areas is nor reproduced. Additionally, the highest reflectively did not occur in the area where the Radar Maxima expects it (based on radar data of the one-hour period prior too the time).



11 July 2023, 09 UTC Radar Maxima forecast of the probability of precipitation intensity exceeding 5 mm in one hour at 16 UTC across southern Germany (left), 12 UTC ICON-D2 EPS maximum reflectivity forecast of any member for this time, and radar display at 16 UTC.

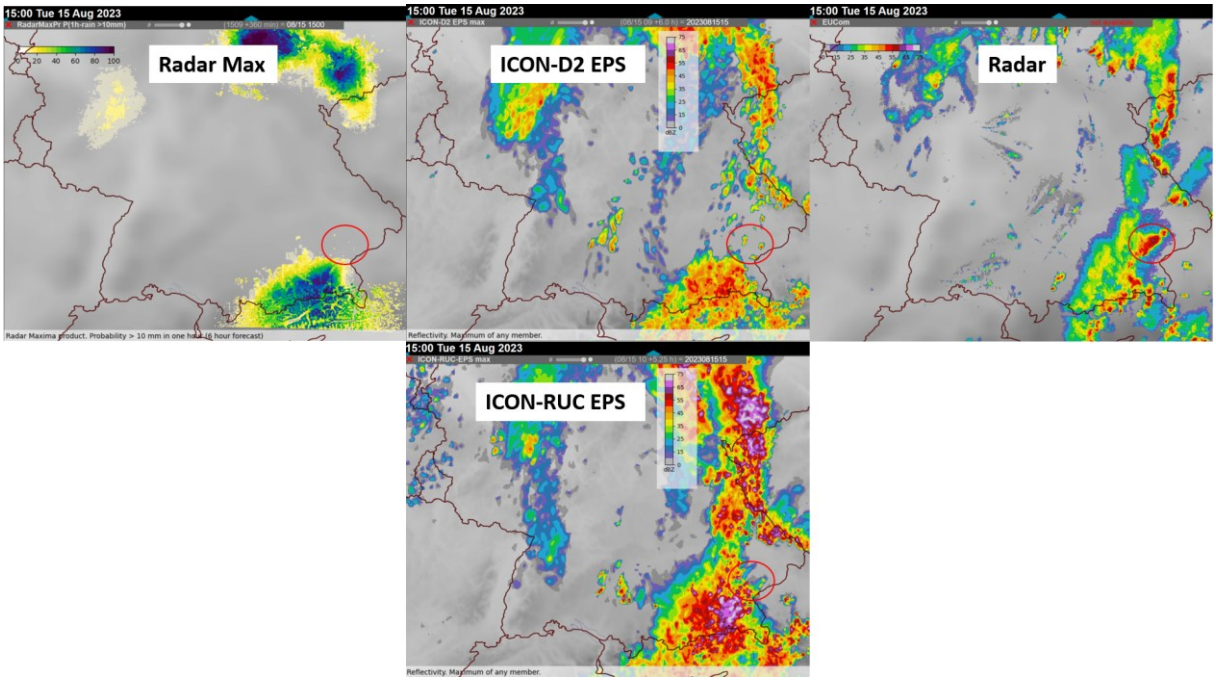
Question 3 was about the tendency to over- or under-forecast precipitation when using probabilities. Four groups filled in the evaluation using the options “Under-forecast”, “Forecast on spot”, “Over-forecast”, with three groups deciding for “under-forecast”, one group noted that the forecast “on spot”



2 August 2023: The forecast was on spot for fast-moving storms across Germany. Precipitation measurements indicate about 5 to slightly above 10 mm in one hour in the highlighted regions. In the southern part of the precip band, Radar Max seems to underestimate the intensity slightly.

The Testbed participants were asked to look for other products in comparison to the Radar Maxima product to rate the influence of the calibration. Two groups said that using the ICON-RUC updates reduced under-forecasting based on an older ICON-D2 EPS. In general, participants saw benefit from using the calibrated Radar Maxima product, five groups chose the answer “benefit in some cases” and one group “benefit in most of the cases”.

A case when ICON-RUC increased the forecast performance is of 15 August 2023. The Radar Maxima was not able to forecast intense rain in the highlighted area, and ICON-D2 EPS did not capture this precipitation. However, ICON-RUC forecasts of 10 and 11 UTC had a better forecast.



15 August 2023, 09 UTC Radar Maxima forecast of the probability of precipitation intensity exceeding 10 mm in one hour at 15 UTC across southern Germany (upper left), 09 UTC ICON-D2 EPS forecast of the maximum reflectivity of any member for this time (upper center), and radar display at 15 UTC (upper right). 10 UTC ICON-RUC EPS forecast for 15 UTC is displayed in the bottom.

Finally, when asked about gross failures in the Radar Maxima product that may occur from technical issues, 2 of 4 groups raised questions about the representation of precipitation in the complex topography of the Alps. Indeed, it is not sure if the influence of topography is that large in convective situations.

4 KONRAD3D-EPS and KONRAD3D-SINFONY

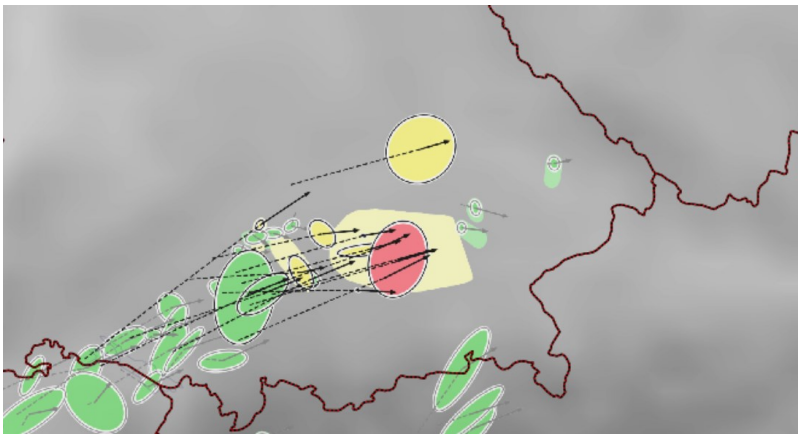
KONRAD3D-EPS is a nowcasting algorithm that enables an object-based ensemble prediction of convective cells detected in radar data. The basis for this is KONRAD3D cells that are provided with a temporal resolution of 5 min for the domain of the German radar composite. The goal is to provide a probabilistic prediction of intensity changes for convective cells and to realistically estimate the uncertainties associated with this prediction.

In a first step, an observation ensemble is generated from detected KONRAD3D cells. This is followed by an initial estimation of the maximum expected severity and lifetime, temporarily based on statistical distributions and later machine learning incorporating environmental data. Subsequently, a prediction of severity is made for each individual cell member. For this purpose, a parabolic model is used as preliminary investigations have shown that the severity of thunderstorm cells, on average, behaves like an inverted parabola. The prediction horizon is currently set to 2 hours.

KONRAD3D-SINFONY is an object-based forecasting system for convective cells that is based on the combination of nowcasting (NWC) and numerical weather prediction (NWP). NWC currently is superior to NWP on the very-short range up to about two hours in predicting convective cells while NWP performs better afterwards. The goal of KONRAD3D-SINFONY is to combine the advantages of both approaches and to generate a seamless prediction.

4.1 Visualisation

Both products were visualised either in the pseudo-member or the exceedance probability of cell severity form. The pseudo-member product displayed the location and intensity of a cell that was most “similar” to all the other cells in the ensemble. The visualisation included the original position of the cell, predicted position, motion vector and a hull around the matching members of the pseudo-member. Colour represented the intensity of the pseudo-member. Exceedance probability of cell severity was visualized using a colour scale used for the ICON-RUC probabilistic products.



An example of the visualisation of the pseudo-member product. Dashed lines show the track of the pseudo-member cells. Black arrows represent the storm motion and colour the intensity of the cell. Cells with black outlines have been detected by the radar and the cells with grey outlines are based on the model simulation.

4.2 Conclusions

- Both KONRAD3D-EPS and KONRAD3D-SINFONY **decreased the intensity** of well-organized storms, such as **supercells** or **bow-echoes faster than in reality**. In general, SINFONY decreased the cell intensity less rapidly than the EPS, but the constant intensity predictions of KONRAD3D or NowCastMIX were better for these types of storms.

- The initial **cell intensity** at the start of the nowcast **sometimes differed** between KONRAD-3D, -EPS and -SINFONY. It was not clear why this is the case.

- Forecasts of **cell intensity and cell lifetime only worked well for short-lived storms**. KONRAD3D-EPS forecasts were useful only for a short time. Regarding KONRAD3D-SINFONY, there was a **strong agreement that the ICON-RUC model adds some beneficial information** compared to the pure nowcasting. The benefit was found both for shorter and longer lead times, but most benefit was noted for long-lived storms and new cell development. It made KONRAD3D-SINFONY more valuable than KONRAD3D-EPS.

- **Probabilistic output was found to be more accessible than the pseudo-member visualization.** However, the probabilities clearly showed one of the biggest problems of the products: that the values decreased too quickly. Perhaps a low threshold of probability could be used by forecasters to draw attention to areas of interest.

- Noticeable artifacts were found both with transition from nowcast- to model-based forecasts and in the watch regions. **Jumpiness** in the cell location and

intensity and **overlapping** of multiple cells of different intensities was noted. This made the products often **confusing** and **difficult to understand**.

4.3 Recommendations

- For a future testbed, KONRAD3D-SINFONY is the preferred, better product of the two to be shown to the participants.
- The forecast of the life cycle of cells needs to be improved. The background environment should be taken into consideration, especially the vertical wind shear. In situations with weaker shear, cells can be expected to live shorter than in the situations with strong vertical wind shear.
- The product's visualisation could be simplified, for example by reducing the number of displayed pseudo-member cells or limiting the range of intensities displayed.
- The probabilistic product be smoothed as the probability areas now strongly copy the shapes of the simulated cells.
- The probabilistic product should be the primary product shown to the forecasters.

4.4 Questions about KONRAD-3D EPS

4.4.1 Do the amplitude of the severity predictions of KONRAD3D-EPS outperform the constant severity predictions of KONRAD3D?

11 participant groups rated this question using a star system, ranging from "not at all" (1 star), through "in some cases" (2 stars) to "in most or all cases" (3 stars). The average score was 1.9 stars and the most common answer was "in some cases". Three groups indicated "not at all" and two "in most or all cases".

The rating of this question typically depended on the type of the situation participants were facing. The lowest scores were assigned in situations with long-lived convective storms, especially supercells. Participants noted that the KONRAD3D-EPS decreases the severity of cells too fast and that the constant severity prediction by KONRAD3D were better. The best scores were given in situations with short-lived convective storms with disorganized convection, when the individual cells lived shorter than for 60 minutes.

Because the disorganized, short-lived storms are much more frequent than the organized, long-lived ones, it's natural that the algorithm is more suited for the prediction of the former type of storms. To better reflect the organized, long-lived storms, which are

also typically more severe, the algorithm must be able to recognize that such situation is present. This could be done by identifying the type of convective storm involved or using a large-scale environment in which storms form, such as the vertical wind shear. Statistical analysis of the life cycles and intensity evolution of cells could be then related to the type of the identified storm or the strength of the vertical wind shear.

4.4.2 Do the flexible lifetime predictions of KONRAD3D-EPS outperform the fixed lifetime prediction of KONRAD3D?

11 participant groups rated this question using a star system, ranging from "not at all" (1 star), through "in some cases" (2 stars) to "in most or all cases" (3 stars). The average score was 1.7 stars and the most common answer was "in some cases". Four groups indicated "not at all" and only one group "in most or all cases".

The responses to this question were very similar to the previous question. In general, the algorithm predicts the decay of the storms within one hour. For the longer-lived storms, such as bow-echoes and supercells, this was incorrect as many of them lived for several hours. In these cases, the KONRAD3D fixed lifetime prediction was better. For shorter-lived, disorganized storms, the flexible predictions of KONRAD3D-EPS were better.

4.4.3 How useful is the information provided by the KONRAD3D-EPS pseudo-member product?

10 participant groups rated this question using a star system, ranging from "not useful" (1 star), through "useful in some cases" (2 stars) to "very useful" (3 stars). The average score was 1.6 stars and the most common answer was "useful only in some cases". Four groups indicated the product was "not useful" and no group found it "very useful".

Relatively low scores attained for this question can be attributed to the fact that the predictions of the severity and the lifetime were incorrect for the cases with longer-lived, organized storms. One report indicated that the product helped them to concentrate on the area with the highest risk. Other reports, however, mention that the location of the pseudo-member cells were incorrect due to the changes to the overall motion of the storm system caused either by the deviation from the mean motion (in supercells), formation of new cells along a flank of the storm (in larger systems) or forward acceleration of the storm (in bow echoes).

4.4.4 How useful is the information provided by the KONRAD3D-EPS probability product?

11 participant groups rated this question using a star system, ranging from "not useful" (1 star), through "useful in some cases" (2 stars) to "very useful" (3 stars). The average score

was 2.2 stars and the most common answer was “useful only in some cases”. No groups indicated that the product was “not useful” and two group found it “very useful”.

The probability product was better received than the pseudo-member product. Two groups noted that it better reflects or visualizes the uncertainty of the forecast than the pseudo-member product. One group noted that while the probabilities for a long-lived convective system decrease too fast, it is still a better reflection of its potential movement than the pseudo-member product. One group indicated that the product was useless for a supercell (too fast decay), but useful for a cluster of storms. Interestingly, one group that dealt with short-lived and disorganized storms even wrote that the probabilities stayed “too strong for too long”.

4.4.5 What is the maximum lead time, for which you would trust a lifetime forecast in the case of long-lived cells? (current plots are delivered for max 2 h lead time)

Participants were able to choose from these options: *< 0.5 h*, *0.5 - 1 h*, *1 - 1.5 h*, *1.5 - 2 h* and *> 2 h*. It was possible to choose multiple options. 11 groups answered this question. No group has indicated the possibility of lead time of over 2 hours. Only one group indicated that the lead time could be between 1.5 and 2 hours. Other votes were almost equally distributed over the 3 lower lead time categories of up to 1.5 hours.

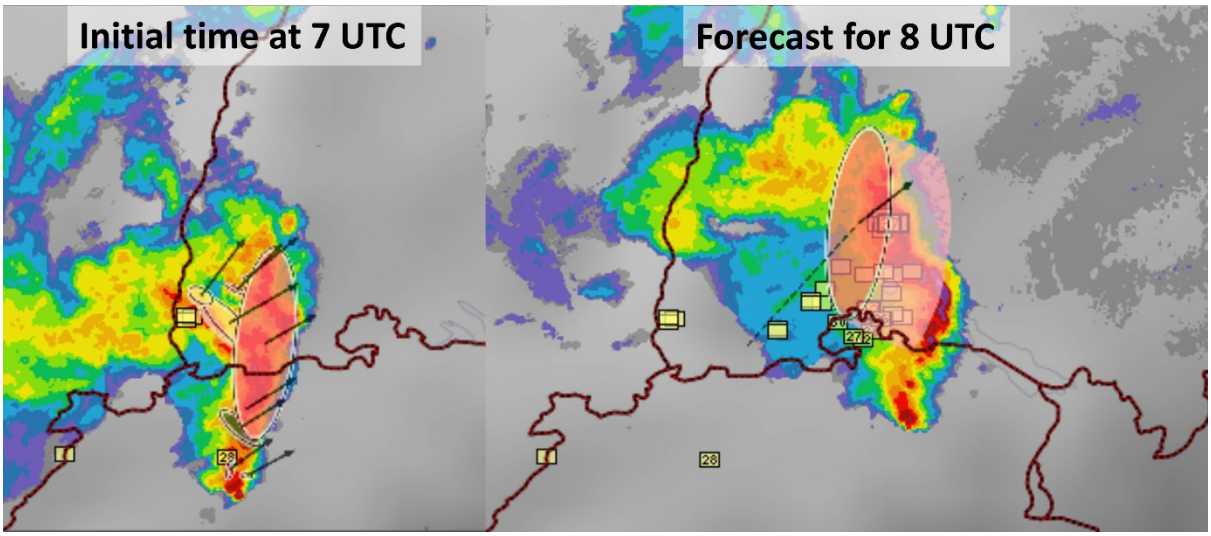
Two groups have indicated that the maximum trustworthy lead time for their cases was less than 0.5 hours. This concerned the supercell and the bow-echo, which may be surprising as in both cases, the storms lived for multiple hours. The reason is that the product predicted the storms to decay fast, while they either maintained or even increased the intensities. Only one group mentioned that the useful lead time extended up to 2 h, which is a maximum range currently displayed. The results acquired from the evaluation clearly indicate that at least in the current state of the product, it’s not useful to increase the lead time further beyond 2 hours.

4.5 Discussion illustrated with cases for KONRAD-3D EPS

4.5.1 21 June 2023

In the morning hours of 21 June 2023, a cluster of storms has travelled from France through Switzerland to Germany. The cluster has organized into a fast-moving bow-echo, producing a swath of damaging wind gusts towards the border with Czechia. The damaging stage of the storm lasted from 7:30 to 12 UTC. Two wind gusts over 40 m/s were measured. The case was unique from the perspective of a convective windstorm occurring so early in the day and was not well forecasted the day before the event, making the correct nowcast very important.

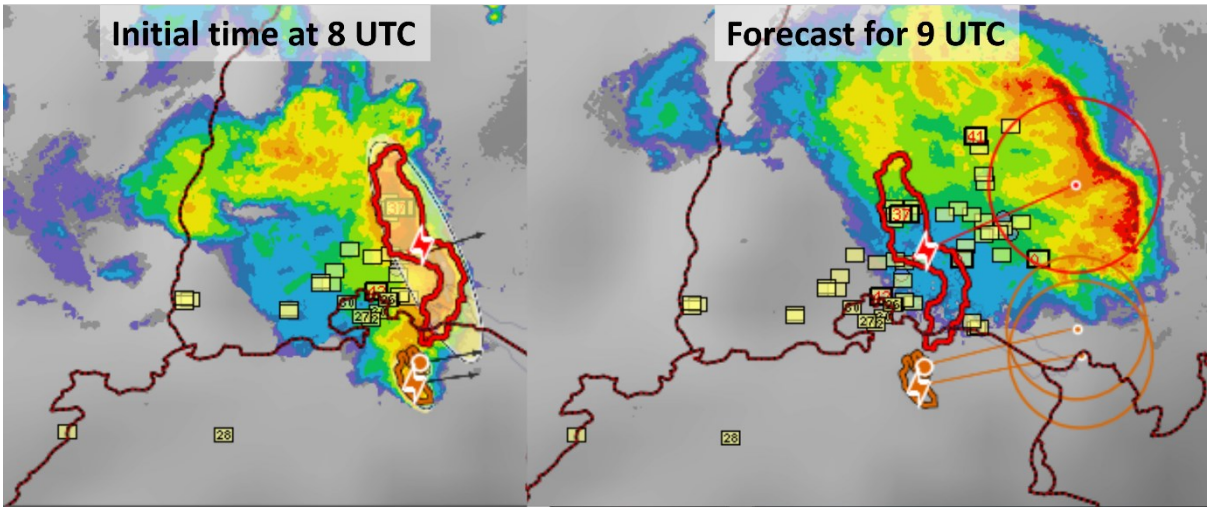
At 7 UTC, just before the bow-echo stage, KONRAD3D-EPS has detected the large cluster of cells and marked it as severe. The algorithm has maintained the intensity of the cluster as severe throughout 8 UTC. The system in fact became extremely severe, producing two gusts exceeding 32 m/s. The 8 UTC forecast places the pseudo member outside of the segment of the convective system that is undergoing the transition into the bow-echo.



21 June 2023 observed radar reflectivity, ESWD reports in last 3 hours and the KONRAD3D-EPS pseudo member product shown for the nowcast initial time at 7 UTC and forecast for 8 UTC.

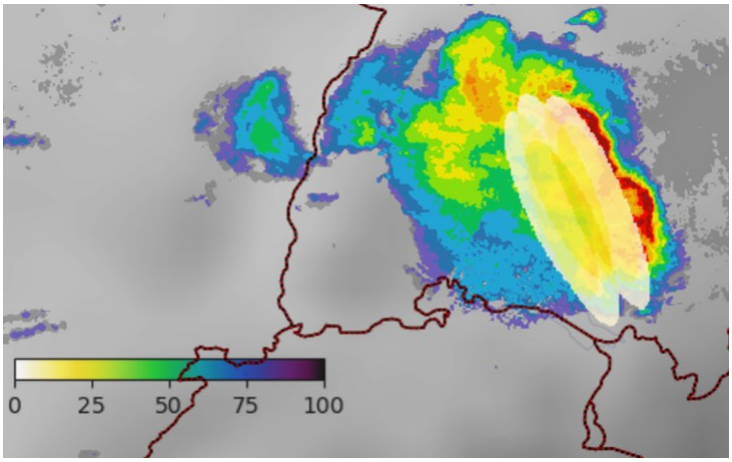
At 8 UTC, the product assigned, very surprisingly, only moderate intensity to the bow echo (yellow ellipse in the figure This contrasts with the KONRAD-3D, which marked the system as severe (red area with lightning inside). KONRAD3D-EPS forecasted the system to decay. The product expected the intensity of the cell to decrease to “any” by 8:40 UTC and completely decay by 9 UTC. By this time, the system was still extremely severe, producing a swath of severe wind gusts up to 41 m/s. While KONRAD-3D underestimated the intensity by one level, its forecast of the position of the system was very good. Also, it was correct to assume that the system will maintain its intensity in this case.

The same problem arose also with the 9 UTC nowcast. KONRAD3D-EPS identified the system, but only of the strong intensity. The system was forecast to decrease its intensity by 9:20 UTC and decay by 9:40 UTC. At 10 UTC, the system was in fact still producing severe wind gusts. Unlike at 8 UTC, KONRAD-3D also identified the cell as strong, not severe. Of course, it at least suggested that the system will maintain its intensity over the following hour. 10 UTC nowcast by the KONRAD3D-EPS had the same progression as the 9 UTC nowcast. The system was identified as being strong in intensity and was forecast to weaken by 10:20 UTC and decay by 10:50 UTC.



21 June 2023 observed radar reflectivity, ESWD reports in last 3 hours, KONRAD3D and KONRAD3D-EPS pseudo member product shown for the nowcast initial time at 7 UTC and forecast for 8 UTC.

The probabilistic information from the KONRAD3D-EPS hasn't improved the nowcasting potential. At the initialization time of 8 UTC, the product assigned probability of about 60% that the system is severe. However, within the first 10-minute forecast period, the probability decreased below 25% and dropped to 0 by 8:20 UTC. For the strong cell intensity, the probability was 100% at the analysis time, dropping to around 75% by 8:30 UTC and to 0 by 9 UTC. Furthermore, the convective system started moving ahead of the area of non-zero probability after 8:40 UTC (Fig X).



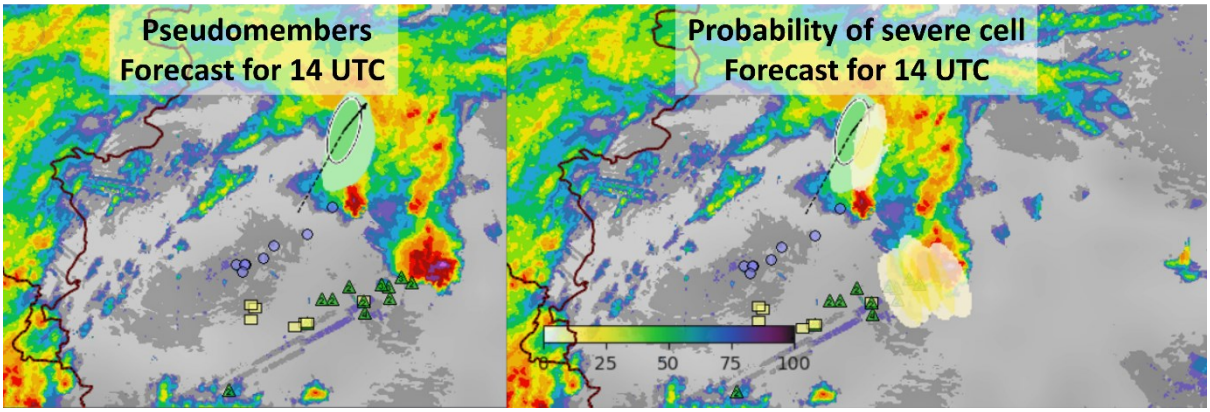
21 June 2023 8:50 UTC observed radar reflectivity and KONRAD3D-EPS forecast of strong cell probability initialized at 8 UTC.

In summary, KONRAD3D-EPS didn't correctly identify the intensity of the convective system and was forecasting too early decay of the storm, typically within 60 minutes of the nowcast start. It is noteworthy that only the 7 UTC nowcast, when the system was still strengthening, called for more than 60 min persistence. Comparing the 7 UTC to other nowcast on that day could shed more light on the performance of the product.

4.5.2 22 June 2023

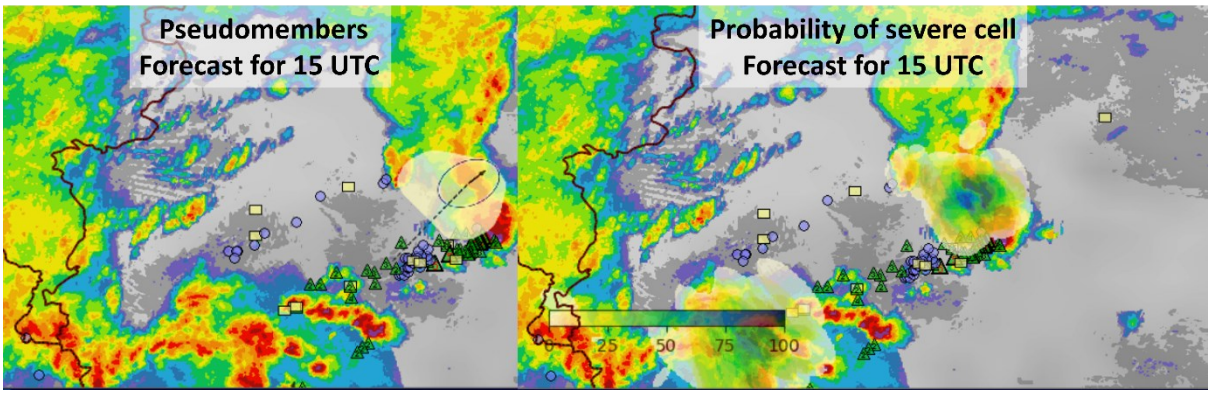
One day later, a long-lived supercell produced a swath of hail, wind and heavy rainfall across northern Germany. The most severe stage lasted from 14 to 15 UTC, when the storm produced wind-driven hail up to 6 cm across. Before 14 and after 15 UTC, the storm was still capable of large hail up to 4 cm and some wind damage. We explore two nowcast times, one from 13 and one from 14 UTC.

At 13 UTC, the supercell was identified as severe. By 13:40 UTC, its intensity was forecast to decrease to strong, by 13:50 UTC to weak and to decay by 14 UTC. The intensity of the cell was increasing throughout the time. Probabilistic visualization offered a better guidance. While the probability of severe cell decreased throughout the period, it was still around 25% during at 14 UTC. One cell member retained severe intensity till 14:50 UTC, albeit at the wrong location.



22 June 2023 14 UTC observed radar reflectivity, ESWD reports in the last 3h and KONRAD3D-EPS forecast of pseudo members and severe cell probability initialized at 13 UTC.

At 14 UTC, the supercell was identified again as severe. The pseudo member remained severe till 14:30 UTC but moving with more northerly component than the real supercell motion. The supercell remained in the "hull" of the pseudo member by that time. By 15 UTC, most of the supercell lied outside of the pseudo member and its hull. At that time, the intensity of the cell was forecast to be strong. Further weakening was forecast afterwards with decay by 15:20 UTC. It's concerning also that none of the other cells to the west are covered by any pseudo member. Probabilistic product showed high probability of severe cell, reaching 75%, even at 15 UTC. The area with the highest probability was displaced slightly to the north of the real supercell location. High probabilities of a severe cell also covered cluster of cells to the southwest, albeit only partly.



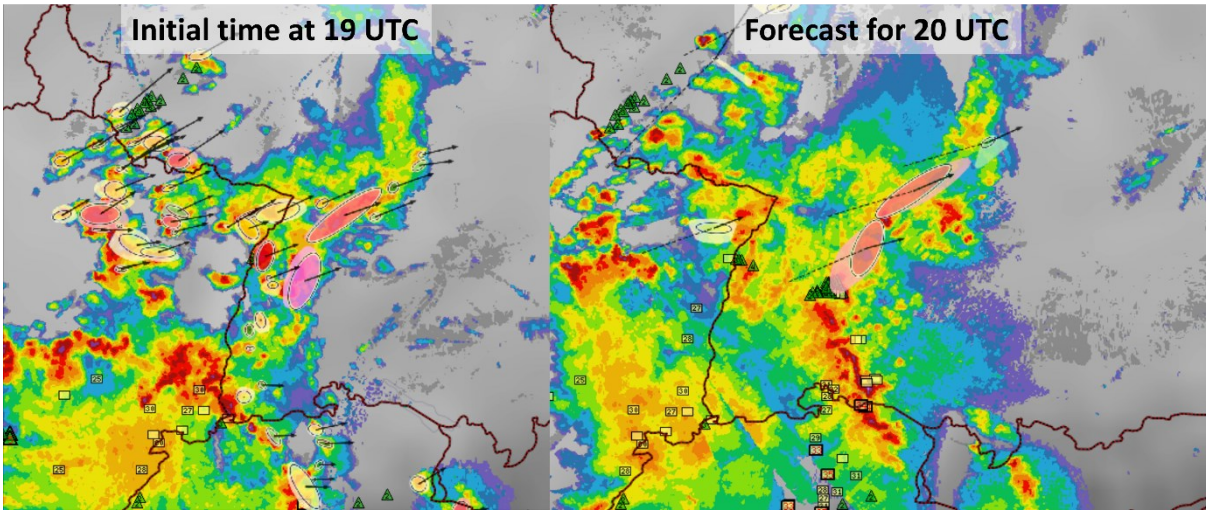
22 June 2023 15 UTC observed radar reflectivity, ESWD reports in the last 3h and KONRAD3D-EPS forecast of pseudo members and severe cell probability initialized at 13 UTC.

In summary, the pseudo member product offered less value in this forecast compared to the probabilistic product. The pseudo member weakened/decayed far too fast so that the product missed the most intense stage of the storm when looking at the 13 UTC forecast. The movement of both the pseudo member and of all the cells in the ensemble didn't follow the motion of the supercell.

4.5.3 11 July 2023

Another severe convective windstorm has crossed southern Germany on 11 July 2023. The system arrived in Germany in the evening hours from France and Switzerland, where it already produced wind gusts up to 35 m/s. Across southern Germany, the damaging stage of the storm lasted between 20 and 24 UTC, with a wide swath of wind damage reported and the strongest wind gust reaching 38 m/s. The windstorm continued over Austria even after midnight, still producing isolated gusts over 32 m/s. The convective system and environment conducive to severe wind gusts were well predictable even one or two days ahead.

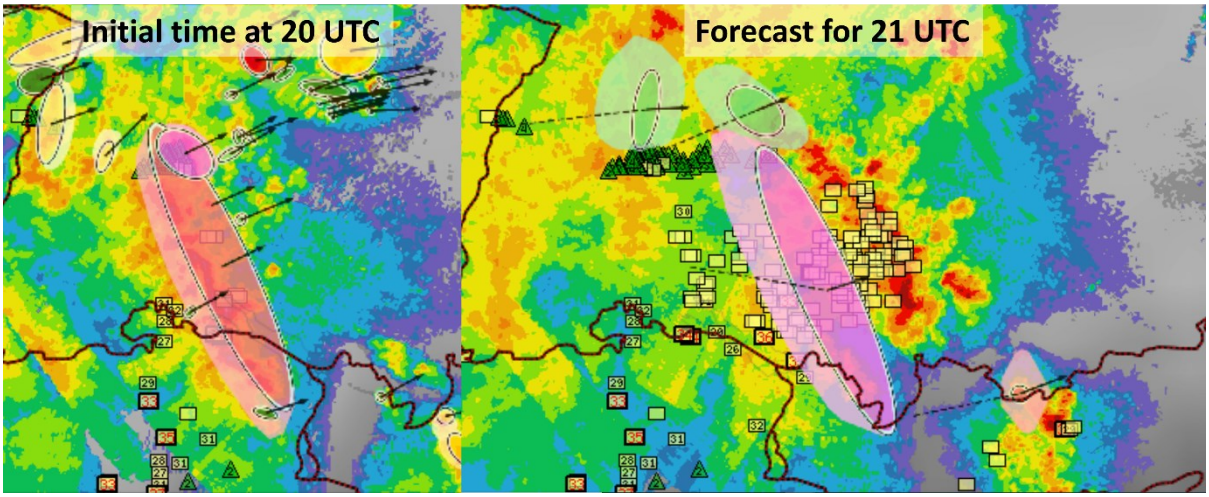
At 19 UTC, KONRAD3D-EPS didn't detect the cluster of storms arriving from France and Switzerland. Interestingly, it detected some of the much smaller and weaker cells ahead and to the north of the system. The system correctly picked up a very severe cell that would produce a swath of large hail, albeit the forecast of its motion was incorrect, as well as a forecast of a reduction in its intensity. By 20 UTC, the system entered Germany and produced the first hurricane-force wind gust.



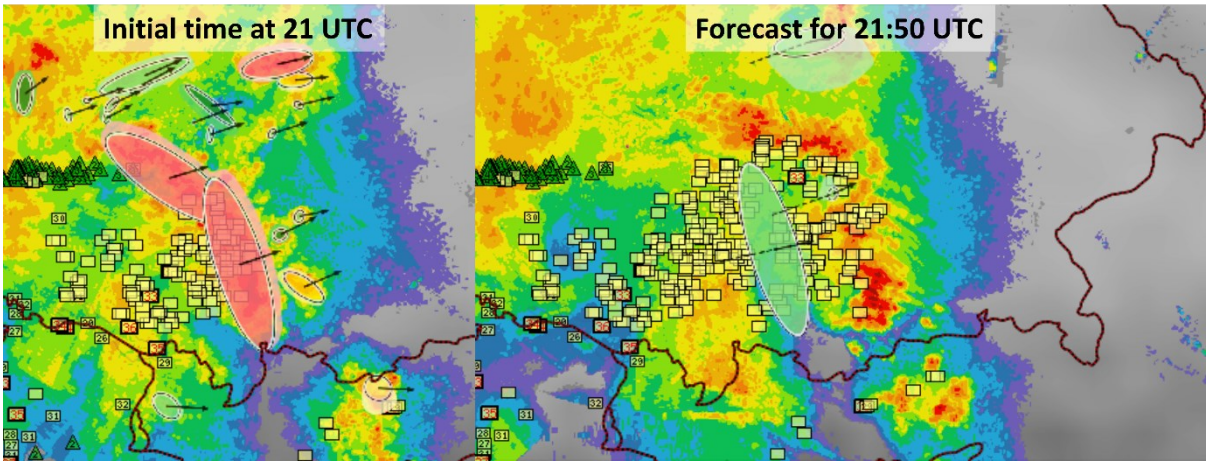
11 July 2023 observed radar reflectivity, ESWD reports in last 3 hours, and KONRAD3D-EPS pseudo member product shown for the nowcast initial time at 19 UTC and forecast for 20 UTC.

At 20 UTC, the product identified the storm system as severe and a supercell at its northern flank as extremely severe. In a strong contrast to the previous case of 21 June, the nowcast increased the intensity to extremely severe by 20:20 UTC, maintaining it throughout the remainder of the 120 minute nowcast till 22 UTC. The call for the very high intensity maintaining over the whole period was correct. However, the acceleration in the forward speed of the system resulted in an incorrect location of the pseudo member. The nowcast for supercell was less successful. The intensity was decreased to weak by 21 UTC with pseudo member located north of the real cell.

At 21 UTC, KONRAD3D-EPS identified the storm system as severe and predicted a decrease to strong intensity by 21:20 UTC, to weak intensity by 21:50 UTC and decaying the system completely by 22 UTC. Wind gusts up to 31 m/s were measured in the 21:50 to 22 UTC period and two wind gusts even over 32 m/s were measured in the 22 to 23 UTC period. Besides the wrong nowcast of the demise of the system, the movement of the pseudo member cell was also too slow, lagging significantly behind the true location of the system.



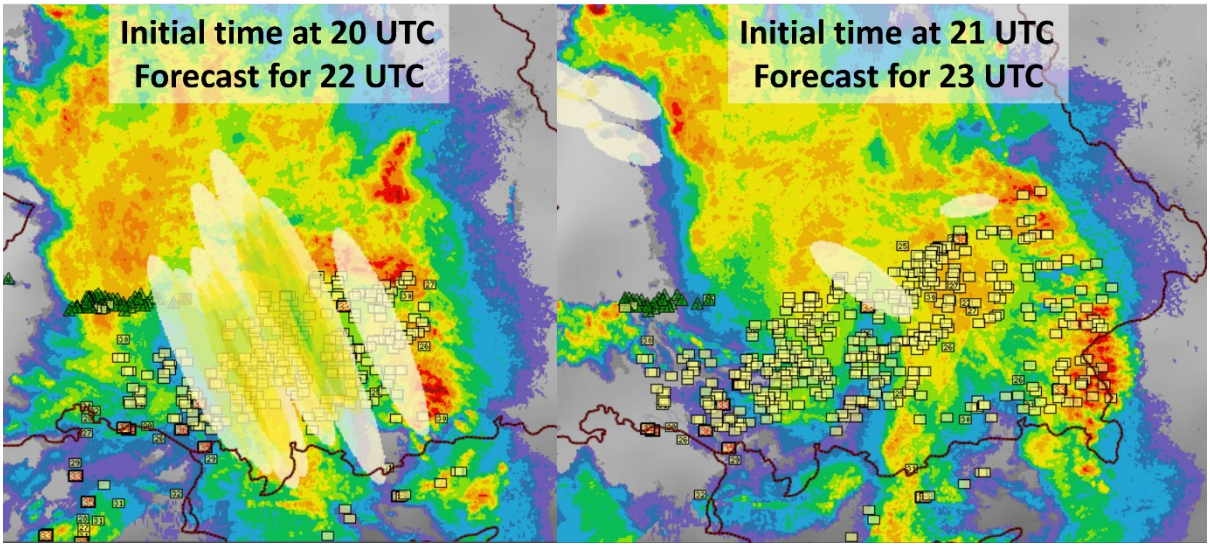
11 July 2023 observed radar reflectivity, ESWD reports in last 3 hours, and KONRAD3D-EPS pseudo member product shown for the nowcast initial time at 20 UTC and forecast for 21 UTC.



11 July 2023 observed radar reflectivity, ESWD reports in last 3 hours, and KONRAD3D-EPS pseudo member product shown for the nowcast initial time at 21 UTC and forecast for 21:50 UTC.

Probabilistic product also showed more confidence concerning the maintenance of the system intensity at 20 UTC initial time than at the 21 UTC. At 19 UTC, it also had no indication concerning the system. Two hour nowcast from 20 UTC still showed some members with severe intensity at 22 UTC, but with a very large spread in the location of the system. In the end, the system managed to travel further than even the "fastest" member of the ensemble. 21 UTC nowcast rapidly decreased the probability of severe cell but did show at least modest probability (< 25%) at 22 UTC. This would make this product setting more useful than the pseudo member one in this instance. By 23 UTC, the probability of severe cell decreased to 0. Comparing the probabilistic and the pseudo member product in the same display shows that the probabilistic product better displays the uncertainty in the forecast. Participants didn't understand the "hull" around the pseudo member and what information it gives concerning the certainty of the location of the pseudo member.

In summary, KONRAD3D-EPS failed to identify the severe convective system about to enter Germany at 19 UTC. Two subsequent forecasts, started at 20 and 21 UTC, showed very different behavior. 20 UTC output maintained the intensity of the system for much longer than the 21 UTC output. This is strange, as at both times, the convective system was very well developed and severe. Investigating further details that could make the difference for the two times could shed more light on the product behavior and potential avenues of its improvement.



11 July 2023 observed radar reflectivity, ESWD reports in last 3 hours, and KONRAD3D-EPS probability of severe cell shown for 22 UTC (as forecasted at 20 UTC) and for 23 UTC (as forecasted at 21 UTC).

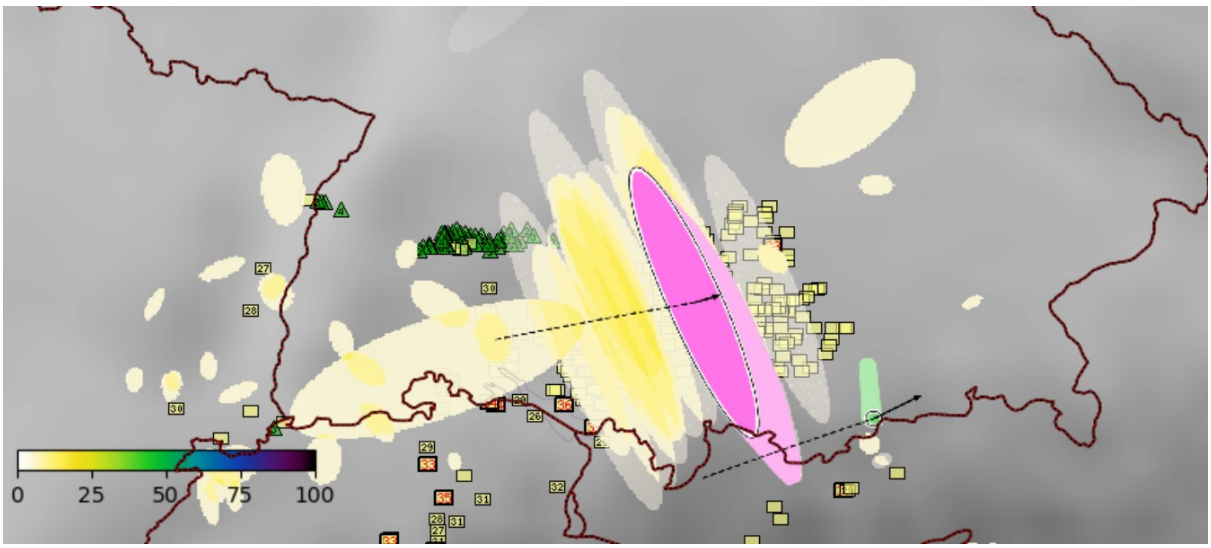


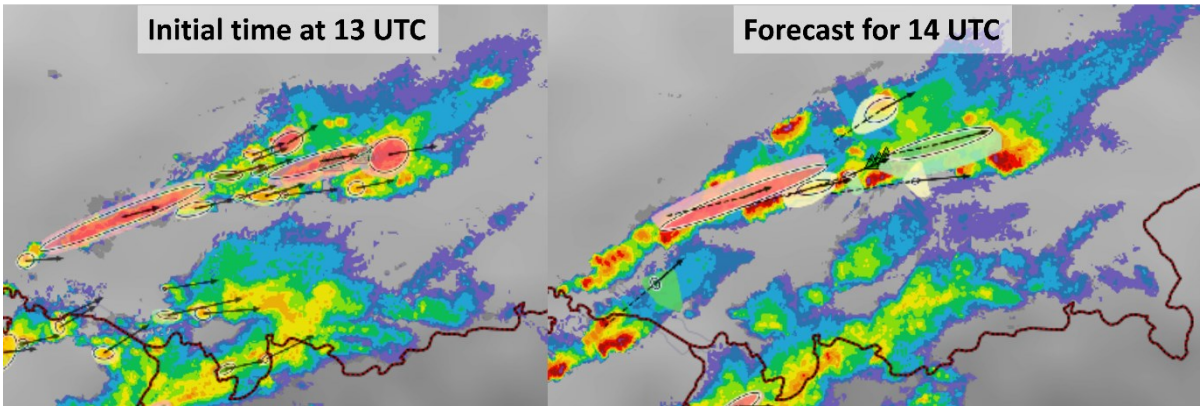
Fig X 11 July 2023 ESWD reports in last 3 hours, KONRAD3D-EPS pseudo member product and the probability of severe cell shown for 21:40 UTC (as forecasted at 20 UTC).

4.5.4 12 July 2023

On 12 July, a mix of multicells and supercells crossed southern Germany. Majority of storms had lifetimes higher than 60 minutes. Three supercells, two right-moving and one

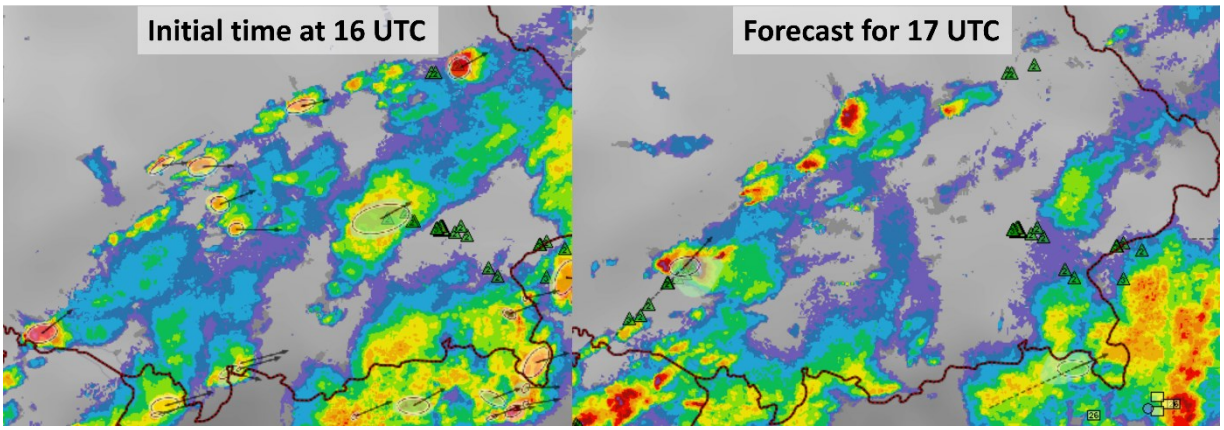
left-moving produced swaths of hail up to 5 cm in diameter. We explore the nowcasts of 13 and 16 UTC.

At 13 UTC, KONRAD3D-EPS assigned 4 cells a severe intensity. Three detections belonged to multicellular storms and one to supercell. The product decreased the intensity first for the right-moving supercell, to strong by 13:10 UTC, to weak by 13:30 UTC with no cell present at 13:40 UTC. At that time, the largest detected cell, a multicell, was still forecast to retain its severe intensity, decreasing to strong by 13:50 UTC. The probability of severe cell for the supercell dropped very rapidly. At 13:10 UTC, it was already below 50% and zero by 13:20 UTC. At the same time, the largest cell (multicell) still had 75% probability, which decreased to 25% by 14 UTC. Number of cells at 14 UTC had no detections, but these cells mostly formed after 13 UTC and couldn't be captured by the product.



12 July 2023 observed radar reflectivity, ESWD reports in last 3 hours, and KONRAD3D-EPS pseudo member product shown for the nowcast initial time at 13 UTC and forecast for 14 UTC.

At 16 UTC, a left-moving supercell was entering Germany from Switzerland with history of producing large hail. Several other cells were existing over southern Germany with one also having a history of producing large hail. Both cells with history of producing large hail were assigned severe intensity. The other cells were assigned strong intensity. The left-moving supercell was forecast to maintain severe intensity till 16:30 UTC, strong intensity till 16:50 UTC and decaying by 17:10 UTC. The probability of the severe cell for the left-moving supercell also decreased very quickly with less than 25% probability at 17 UTC. In reality, the supercell maintained the intensity and large hail production till 19 UTC. The other large hail producing cell was forecast to weaken faster with decay by 16:50 UTC on the border of Czechia and Germany. This prediction was correct.

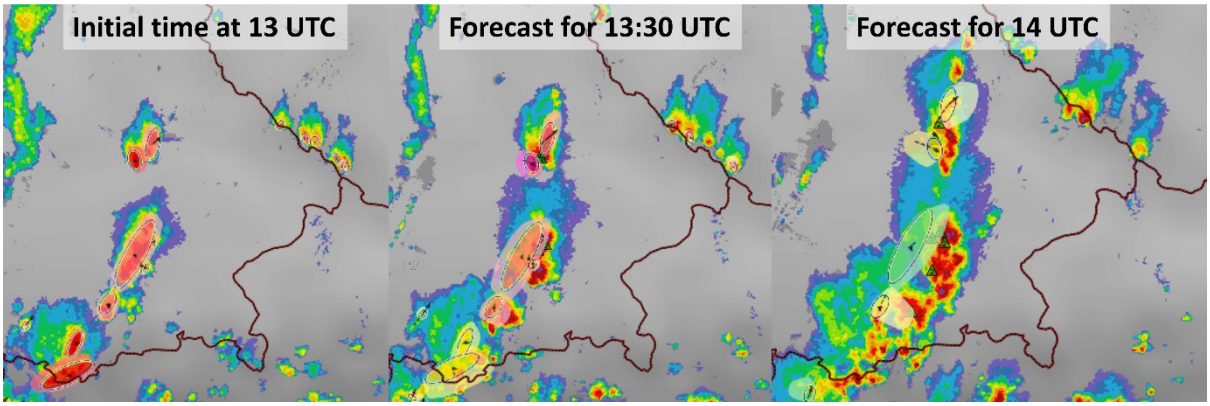


12 July 2023 observed radar reflectivity, ESWD reports in last 3 hours, and KONRAD3D-EPS pseudo member product shown for the nowcast initial time at 16 UTC and forecast for 17 UTC.

4.5.5 15 August 2023

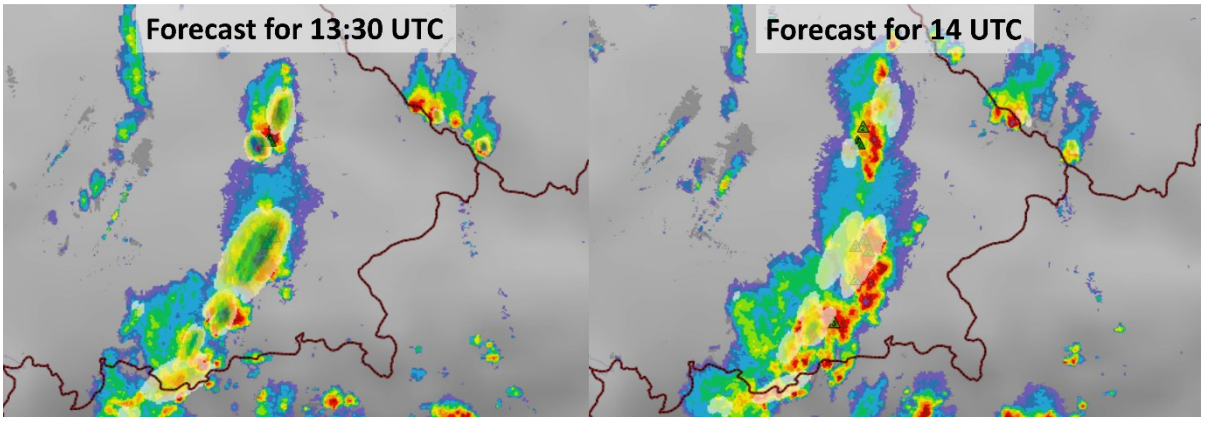
On 15 August, scattered storms formed over eastern Germany in weak mean flow and weak vertical wind shear. The storms existed in the form of rather short-lived cell and clusters. Thanks to the high CAPE, storms were capable of large hail and isolated severe wind gusts. The case gives a good opportunity to evaluate the performance of the product in low shear case, compared to the high shear cases typically investigated in the evaluation forms.

At 13 UTC, most of the cells over the area were marked as severe (Fig X). Perhaps the biggest problem of the forecast was that the storms went outside of the areas marked by the pseudo member cells. In this case, the mean flow was so weak that the individual cells were almost stationary. This led to the forecast of cells remaining basically at the same spot. However, as new cells formed at the flanks of the dying cells, this led to an apparent motion primarily to the southeast. This was, of course, not captured by the KONRAD3D-EPS. Other than that, the forecast of the decreasing intensity of cells in the next hour was correct. Similar pattern is visible also at the forecasted started at 14 UTC. The system correctly predicted the decay/weakening of the cells within one hour, but pseudo members didn't cover the development of new cells.



15 August 2023 observed radar reflectivity, ESWD reports in last 3 hours, and KONRAD3D-EPS pseudo member product shown for the nowcast initial time at 13 UTC and forecasts for 13:30 and 14 UTC.

Probabilistic product offered better guidance compared to the pseudo members, especially for the largest detected cell. While majority of the ensemble members also counted with quasistationary cell, some members moved the cell in the direction of the overall storm motion. For the northern cluster of cells, probabilistic product didn't show much improvement over the pseudo member visualization.



15 August 2023 observed radar reflectivity, ESWD reports in last 3 hours, and KONRAD3D-EPS probability of severe cell for 13:30 and 14 UTC, initialized at 13 UTC.

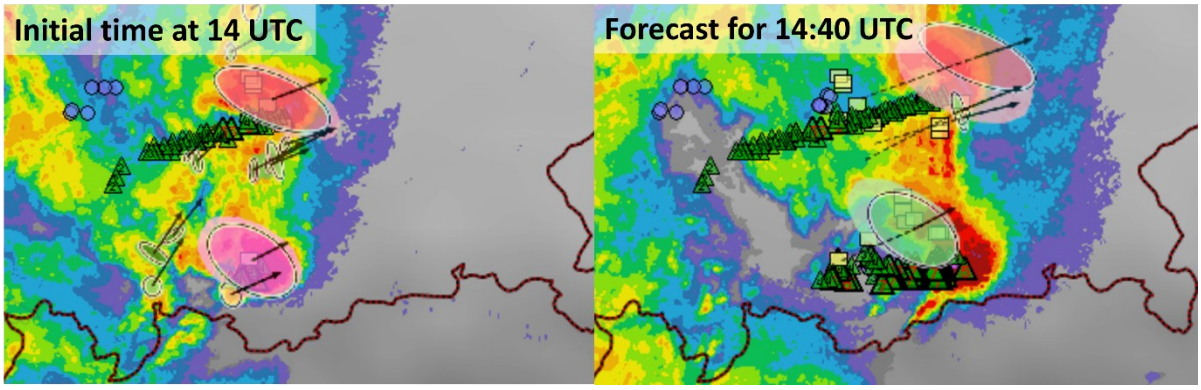
In summary, KONRAD3D-EPS correctly predicted fast decay of short-lived cells in this case but struggled with the storm motion induced by the formation of new cells at the flanks of the dying ones.

4.5.5.1 26 August

On 26 August, an outbreak of severe weather occurred over southern Germany. Two intense supercells developed around 13 UTC, creating swaths of hail and wind damage. The northern supercell produced hail up to 5 and the southern hail up to 10 cm in diameter. By 16 UTC, a bow-echo developed, which produced a swath of damaging winds with a wind gust of 35 m/s on Germany/Austria border. The event was well forecast one day ahead by high-resolution, convection allowing models.

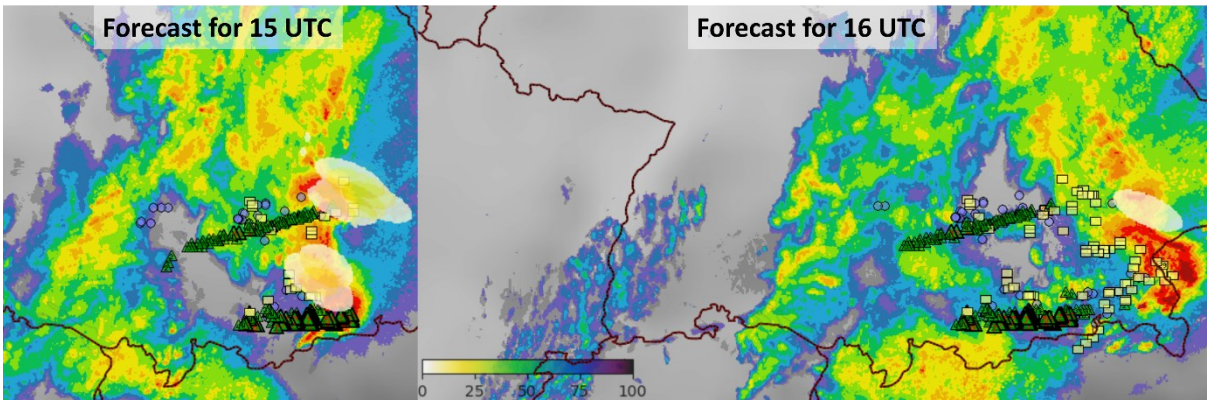
We investigate two nowcast times, one at 14 UTC, the supercell stage, and at 16 UTC, after the southern supercell merged with other cells and formed the bow-echo. The northern supercell had already produced the largest hail at 14 UTC and would weaken shortly after 15 UTC. The southern supercell had just started producing very large hail by 14 UTC, which would continue without interruption till 15 UTC.

At 14 UTC, the product assigned severe intensity to the northern and an extremely severe intensity to the southern storm, which was correct at least from the perspective of the size of the hail produced. However, the intensity of the southern supercell dropped to severe by 14:10 UTC, to strong by 14:30 UTC, to weak by 14:40 UTC and decaying by 14:50 UTC. Furthermore, the pseudo-member deviated to the north of the severe weather track. In contrast, the northern supercell was nowcasted quite well. The severe intensity was maintained till 14:40 UTC, decreasing to weak at 15:10 UTC. By this time the severe weather production had almost stopped in the storm.



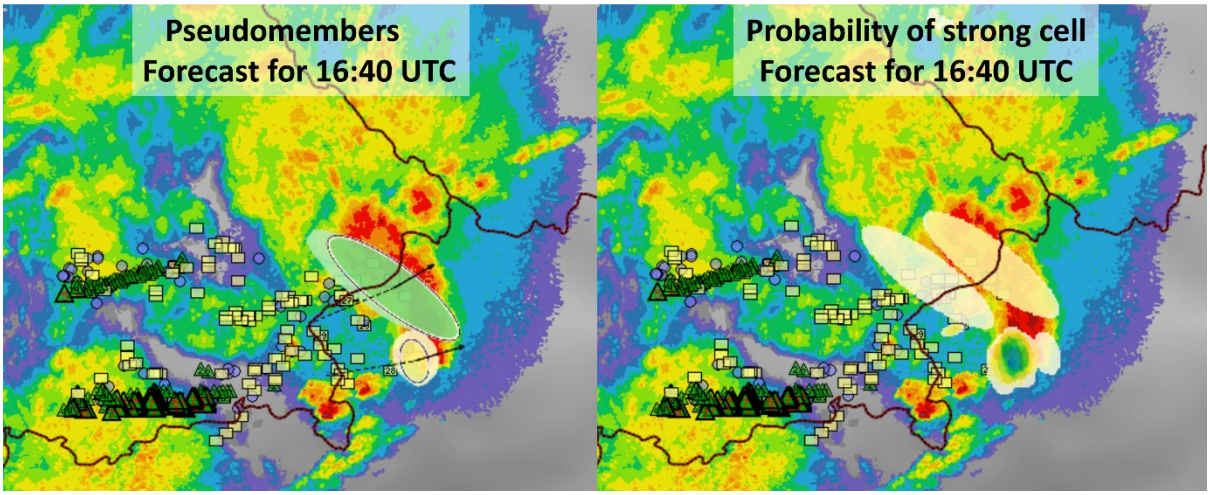
26 August 2023 observed radar reflectivity, ESWD reports in last 3 hours, and KONRAD3D-EPS pseudo member product shown for the initial time at 14 UTC and forecast for 14:40 UTC.

The probabilistic product didn't add more confidence concerning the maintenance of the high intensity of the southern cell. By 14:30 UTC, the probability dropped to 50% for the southern cell, remaining at almost 100% for the northern cell. Very small probability remained at 15 UTC, but only for the northern edge of the southern supercell. The hail swath left an area of non-zero probability of a severe cell by 14:40 UTC. Non-zero probability was maintained for the northern cell till 16 UTC, while no area of probability covered the southern supercell since 15:10 UTC, which started undergoing bow echo transition after 15 UTC.



26 August 2023 observed radar reflectivity, ESWD reports in last 3 hours, and KONRAD3D-EPS probability of severe cell forecast for 15 and 16 UTC, initiated at 14 UTC.

At 16 UTC, the bow echo at the Germany/Austria border started producing severe wind gusts, reaching up to 29 m/s by that time. KONRAD3D-EPS identified the system as strong, which was clearly an underestimation. The intensity of the system was forecast to decrease to weak by 16:40 UTC and to decay by 16:50 UTC. The cell at the southern edge of the system was forecast to persist longer, decaying by 17:30 UTC. The strongest wind gust, 35 m/s, was observed around 16:40 UTC, at the time when the pseudo-member of the system was predicted to have weak intensity. The probability of severe cell was very low already at the initial time, dropping to 0 by 16:30 UTC. The probability of a strong cell dropped rapidly from 100% at 16 UTC to less than 10% by 17 UTC.



26 August 2023 observed radar reflectivity, ESWD reports in last 3 hours, and KONRAD3D-EPS pseudo-member and probability of severe cell forecast for 16:40 UTC, initiated at 16 UTC.

4.6 Questions about KONRAD-3D SINFONY

4.6.1 How would you rate the quality of the combination of NWC and NWP?

11 participant groups rated this question using a star system, ranging from “very bad” (1 star), through “bad in most cases” (2 stars), “somewhat good” (3 stars), “good in most cases” (4 stars) to “very good” (5 stars). The average score was 3.1 stars, with 4 groups selecting “somewhat good” and 4 groups “good in most cases”. No groups indicated that the quality of combination was either “very bad” or “very good”.

The scores assigned from the participants suggest further potential for improvement. Interestingly, even for the same cases, ratings of the quality differ from one report to another. For example, for one case, three feedback forms rated the quality as “bad in most cases”, “somewhat good” and “good in most cases”. There was no pattern in the ratings concerning the type of situation participants encountered.

4.6.2 Are there noticeable artifacts in the forecasts that may result from the transition from the combination of NWC and NWP to forecasts that are purely based purely on NWP?

Noticeable artifacts have been mentioned by most of the submitted reports. Only two out of eleven reports mention no encountered artifacts. Two groups wrote that the artifacts appear, where there was no NWC-based reflectivity, but only simulated by the NWP. Another group that worked with a convective system mentioned an increase in the size of the convective system during the transition from the NWC to NWP based forecast. Jump of the cells from one location to another, a great increase in the number of the cells and the increase of the artifacts with the lead time has also been mentioned in the individual forms.

4.6.3 Are there artifacts in the watch regions (cells with gray borders) associated with the fact that the cells come from an NWP ensemble? E.g., cells too big, variety of different severities overlapping at the same place? If yes, does this add to the uncertainty in the forecast?

Only one out of 11 reports doesn’t mention any visible artifacts in the watch regions. 5 reports pointed at the jumpiness and sudden changes in location, intensity or size of the watch areas are problematic. 5 reports indicated that areas are too big and 5 reports stated that the overlapping of the cells of different sizes and intensities was problematic. One group states that the overlapping is “not necessarily a bad thing”. Word confusion was also used in two reports for this question.

ESSL recommends that the product gains more clarity in the feature. This could be achieved by reducing the number of the cells displayed for example by displaying only the more intense cells. Very large sizes of the cells representing simulated storm systems may be due to the ICON-RUC overestimating the reflectivity in the stratiform region of these systems (see the ICON-RUC evaluation).

4.6.4 Is there any additional value of the combined cells (black borders) in terms of life *cycle and severity compared to KONRAD3D-EPS?*

10 participant groups rated this question using a star system, ranging from "not at all" (1 star), through "only in some cases" (2 stars) to "yes, totally" (3 stars). The average score was 2.3 stars and the most common answer was "only in some cases". Four groups answered "yes, totally" and only one group "not at all".

For several cases, participants have mentioned an added value of the combined cells compared to KONRAD3D-EPS. The most value was noted for a long-lived supercell, followed by larger storm clusters and for lead times past 60 minutes. One report noted that the propagation of right-moving supercell was better captured by the KONRAD3D-EPS alone than by the combined cells. Another report mentioned an added trust to the forecast by using the combined cells.

Two groups also commented on what is easy or difficult to understand about this product. One group wrote that it takes time to understand the concept and another mentions that the product is complicated and that the forecasters didn't understand what different intensity levels mean.

4.6.5 How useful is the information provided by cells based purely on NWP predictions (grey borders) for shorter lead times of up to two hours?

10 participant groups rated this question using a star system, ranging from "not useful at all" (1 star), through "useful only in some cases" (2 stars) to "very useful" (3 stars). The average score was 2.3 stars and the most common answer was "very useful". Two groups found the information "not useful at all".

Two groups mentioned that the information is useful for forecasting where new cell initiation may take place. For a case of long-lived convective system, one group mentioned that the NWP-based cells better predicted the behavior of the convective system than the observed cells. Another group noted that the cells didn't move in the correct direction. One group noted that real storms very rarely occurred within the confines of the cells with grey borders.

Concerning how understandable the product is, the reports note the artifacts and the overlapping of cells with different intensities, which make it hard to follow.

4.6.6 How useful is the information provided by cells based purely on NWP predictions (grey borders) for longer lead times?

11 participant groups rated this question using a star system, ranging from "not useful at all" (1 star), through "useful only in some cases" (2 stars) to "very useful" (3 stars). The average score was 2.4 stars and the most common answer was "useful only in some cases". Four groups found the information "very useful" and no group selected the "not useful at all" choice.

Participants haven't written almost any qualitative feedback in this question. Comparing their ratings to the situation/environment/types of storms they encountered doesn't suggest there would be a strong link between the type of the setup and how useful the NWP predictions for longer lead times are.

4.6.7 How useful is the information provided by the KONRAD3D-SINFONY probability product? Consider that the probability product contains both the information from the combination of NWC and NWP and from pure NWP for future time steps.

10 participant groups rated this question using a star system, ranging from "not useful at all" (1 star), through "useful only in some cases" (2 stars) to "very useful" (3 stars). The average score was 2.2 stars and the most common answer was "very useful" with 4 groups selecting this answer. Two groups found the information "not useful at all".

4 groups remarked on the product being very useful, easy to understand or a good way how to display the probability of cell occurrence in a seamless way. One group didn't understand what the severity levels mean, which indeed are not explained and not directly connected to intensities of severe thunderstorm-related hazards. Interestingly, two groups that looked at the same situations evaluated the product very differently. While one group found it useful, another group remarked that the pure RUC information provides better information, and the product is too time consuming for a forecaster to really dig into.

4.6.8 Does KONRAD3D-SINFONY provide further beneficial information for the forecast in comparison to KONRAD3D-EPS?

11 participant groups rated this question using a star system, ranging from "no benefit from the product" (1 star), through "a little benefit or only in some cases" (2 stars) to "provides a lot of benefit" (3 stars). The average score was 2.7 stars with 8 groups

answering that it “provides a lot of benefit”. No group thought that there is “no benefit from the product”.

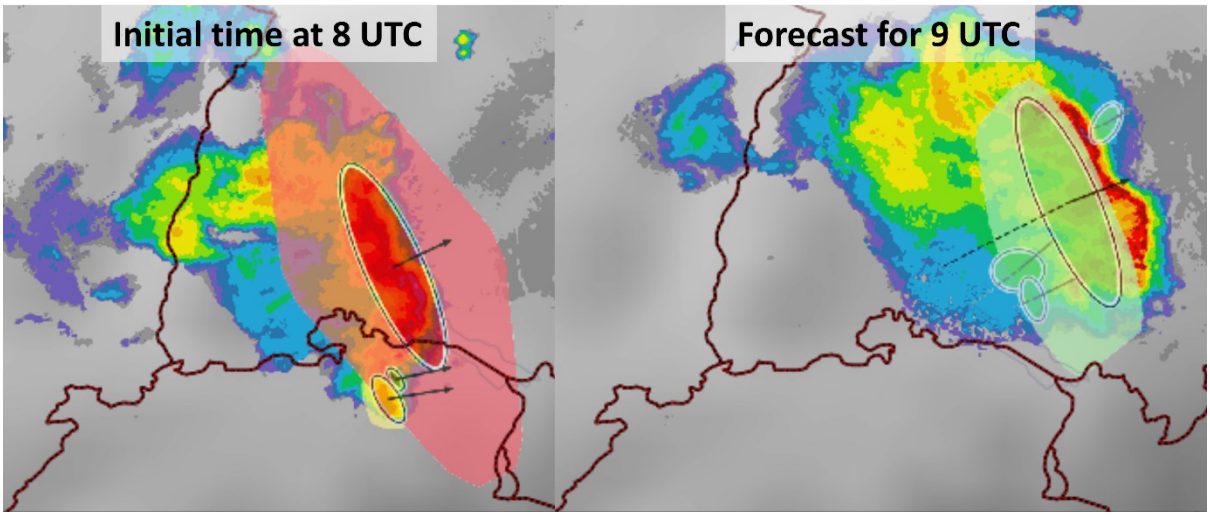
This question yielded the strongest agreement among the individual evaluation reports as the SINFONY provided at least some benefit in all the cases that were addressed during the evaluation period. Some groups have noted some limitations of the SINFONY for particular cases within this question, but these have already been discussed above.

4.7 Discussion illustrated with cases for SINFONY-EPS

4.7.1 21 June 2023

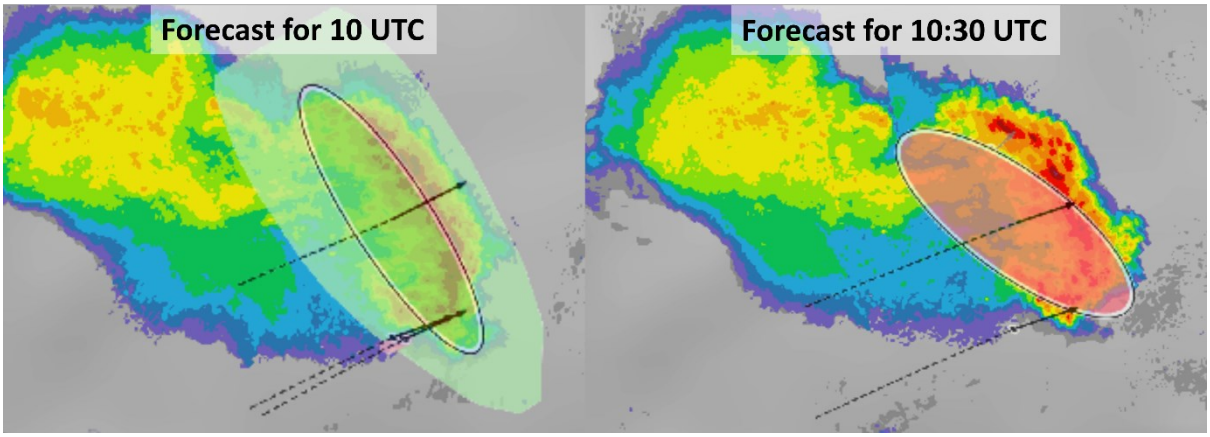
In the morning hours of 21 June 2023, a cluster of storms has travelled from France through Switzerland to Germany. The cluster has organized into a fast-moving bow-echo, producing a swath of damaging wind gusts towards the border with Czechia. The damaging stage of the storm lasted from 7:30 to 12 UTC. Two wind gusts over 40 m/s were measured. The case was unique from the perspective of a convective windstorm occurring so early in the day and was not well forecasted the day before the event, making the correct nowcast very important.

With KONRAD3D-EPS forecasting too fast decay of the bow echo compared to the reality, the main question is how the KONRAD3D-SINFONY compared to the pure nowcasting perspective. We will concentrate on the 8 and 9 UTC forecasts, when the system was at its peak and maintained high intensity. At 8 UTC, SINFONY detected the system as severe, compared to the strong intensity detected by the EPS, i.e. one level higher. The pseudo-member product also shows rapid decrease in the intensity of the system, to strong intensity by 8:10 UTC, to “any” intensity by 9 UTC and completely decaying the system by 9:20 UTC.



21 June 2023 observed radar reflectivity and KONRAD3D-SINFONY pseudo member product shown for the nowcast initial time at 8 UTC and forecast for 9 UTC.

The nowcast started at 9 UTC displayed more peculiar behavior with an artifact caused by the transition from the nowcast (NWC) to model (NWP) based forecast. The system was identified as strong by the product at 9 UTC, weakening by 9:30 UTC. An increase to strong intensity and then another decrease occurred in the 9:40 – 9:50 UTC period. At 10:30 UTC, the intensity of the system increased to severe, decreasing to strong at 10:40 UTC and completely disappearing by 11 UTC. The orientation of the detected cell and its size changed as well.



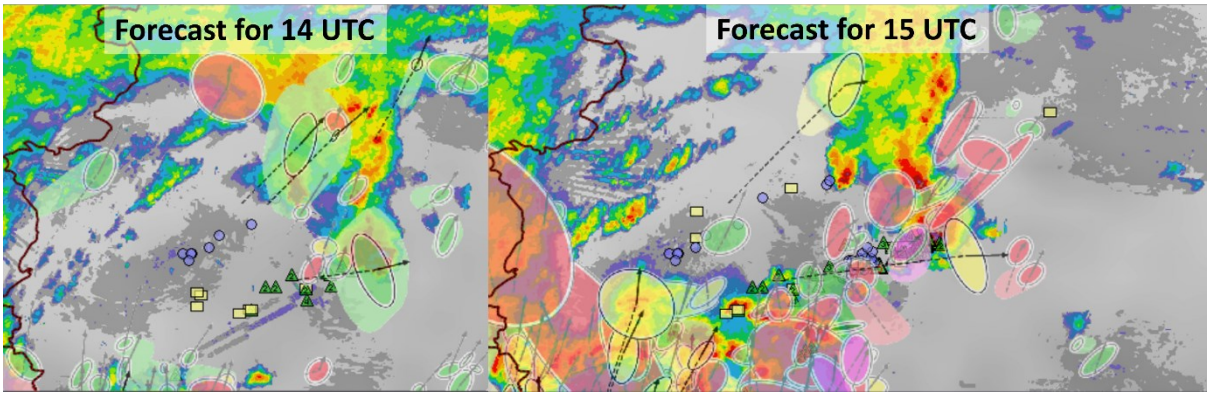
21 June 2023 observed radar reflectivity and KONRAD3D-SINFONY pseudo member product nowcast initiated at 9 UTC.

In summary, KONRAD3D-SINFONY maintained the system for longer time than the KONRAD-3D. However, in some nowcasts the intensity decreased, and the system decayed still too fast compared to the reality. In the other nowcasts, the intensity of the system jumped as an artifact of the product switching from NWC- to NWP-based forecast.

4.7.2 22 June 2023

One day later, a long-lived supercell produced a swath of hail, wind and heavy rainfall across northern Germany. The most severe stage lasted from 14 to 15 UTC, when the storm produced wind-driven hail up to 6 cm across. Before 14 and after 15 UTC, the storm was still capable of large hail up to 4 cm and some wind damage. We explore the nowcasting time from 13 UTC.

At 13 UTC, SINFONY pseudo member product identified the cell as severe. It was also predicted to weaken rather quickly, reaching weak intensity at 14 UTC. Compared to the EPS, the pseudo member of SINFONY followed the track of the supercell precisely. At 14:10 UTC, the intensity jumped up to severe, decreasing again at 14:20 UTC. At that time, NWP-based detections appeared next to the cell, one of them being of extremely severe intensity. The intensity and location of pseudo members of both NWC- and NWP- based detections around the supercell experienced many jumps towards 15 UTC and beyond. A cluster of cells to the southwest was covered by many overlapping cells of different sizes and intensities of pseudo member cells, limiting the usefulness of the product.

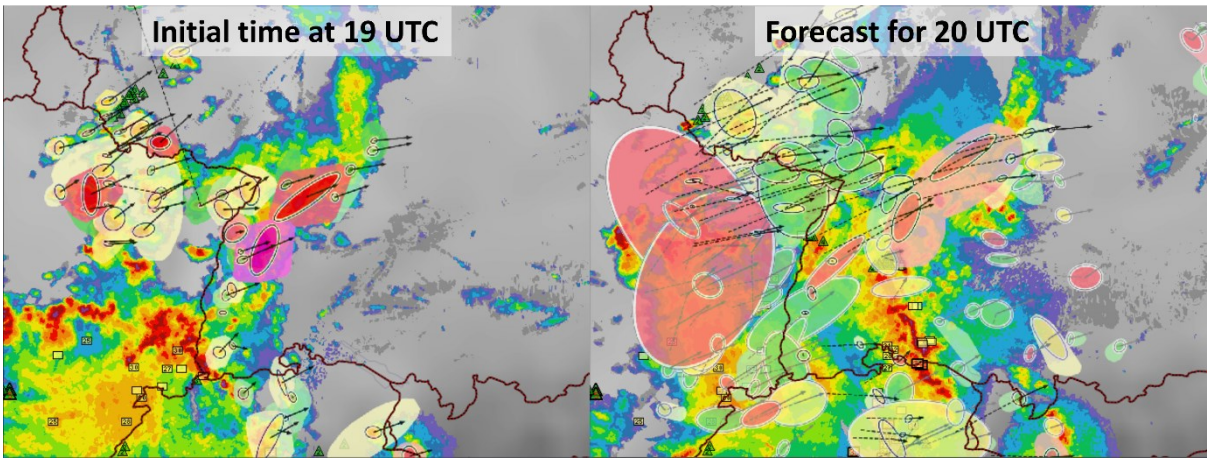


22 June 2023 observed radar reflectivity, ESWD reports in the last 3h and KONRAD3D-SINFONY pseudo member product nowcast initiated at 13 UTC.

4.7.3 11 July 2023

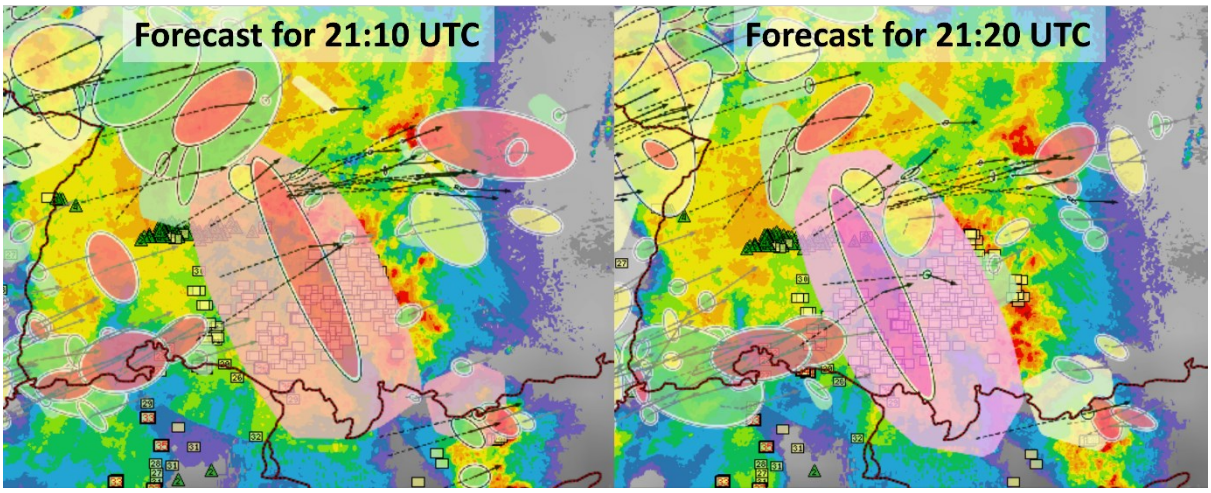
Another severe convective windstorm has crossed southern Germany on 11 July 2023. The system arrived in Germany in the evening hours from France and Switzerland, where it already produced wind gusts up to 35 m/s. Across southern Germany, the damaging stage of the storm lasted between 20 and 24 UTC, with a wide swath of wind damage reported and the strongest wind gust reaching 38 m/s. The windstorm continued over Austria even after midnight, still producing isolated gusts over 32 m/s.

We concentrate on the 19 to 21 UTC nowcasts. At 19 UTC, KONRAD3D-EPS didn't detect the incoming severe convective system. KONRAD3D-SINFONY detected only a small cell within the whole system with much better coverage of other storms in the area. Large numbers of additional cells (NWP based, grey borders) are introduced by the product as early as 19:20 UTC. At 20 UTC, the system is still not covered by a single cell while huge number of cells of different severities and sizes are detected at various times to the west and north of the system. Sudden jumps in the cell locations and intensities persist through the next hour till 21 UTC and beyond.



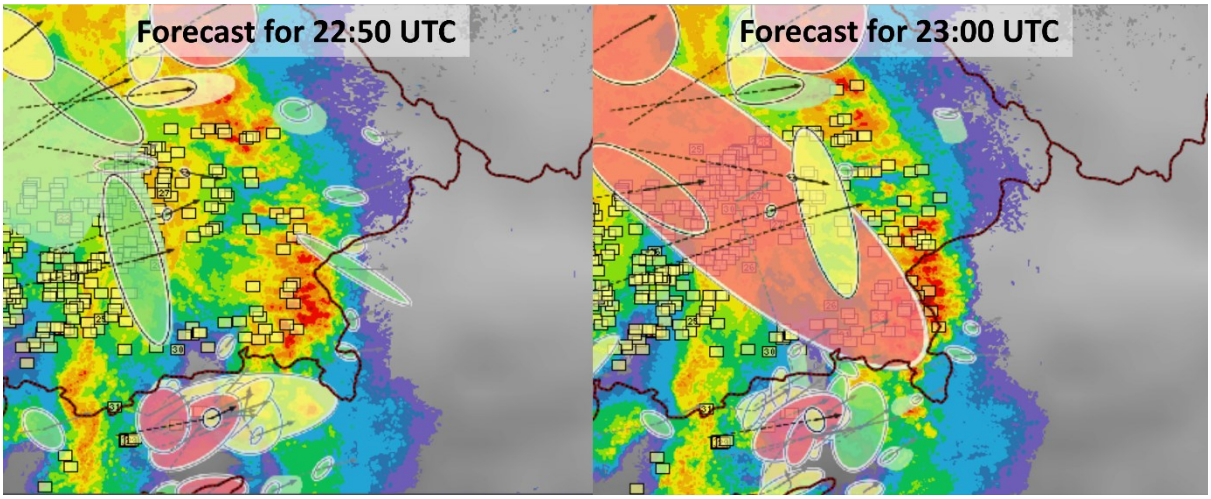
11 July 2023 observed radar reflectivity, ESWD reports in the last 3h and KONRAD3D-SINFONY pseudo member product shown for the nowcast initial time at 19 UTC and forecast for 20 UTC.

At 20 UTC, the system was detected, but large jumps were observed in the forecast of the position of the system's pseudo member in the next 2 hours. At some points the pseudo member was shifted west, while the movement of the system was due east. The intensity also changed several times from severe to extremely severe and back. As for the KONRAD3D-EPS, the system eventually outran even the very large uncertainty area surrounding the pseudo member.



11 July 2023 observed radar reflectivity, ESWD reports in the last 3h and KONRAD3D-SINFONY pseudo member product initiated at 20 UTC and displayed for 21:10 and 21:20 UTC.

At 21 UTC, the system was also detected as severe, but its intensity was forecast to quickly decrease to weak by 21:40 UTC. The intensity jumped back again to severe at 22 UTC, decreasing again to weak by 22:50 UTC and then jumping back again to severe with greatly expanded area by 23 UTC. Such erratic behavior puts less confidence in the forecast and is also not typical for the well-organized convective systems. Like the 20 UTC nowcast, the pseudo member and the uncertainty area started lagging the system considerably.



11 July 2023 observed radar reflectivity, ESWD reports in the last 3h and KONRAD3D-SINFONY pseudo member product initiated at 21 UTC and displayed for 22:50 and 23 UTC.

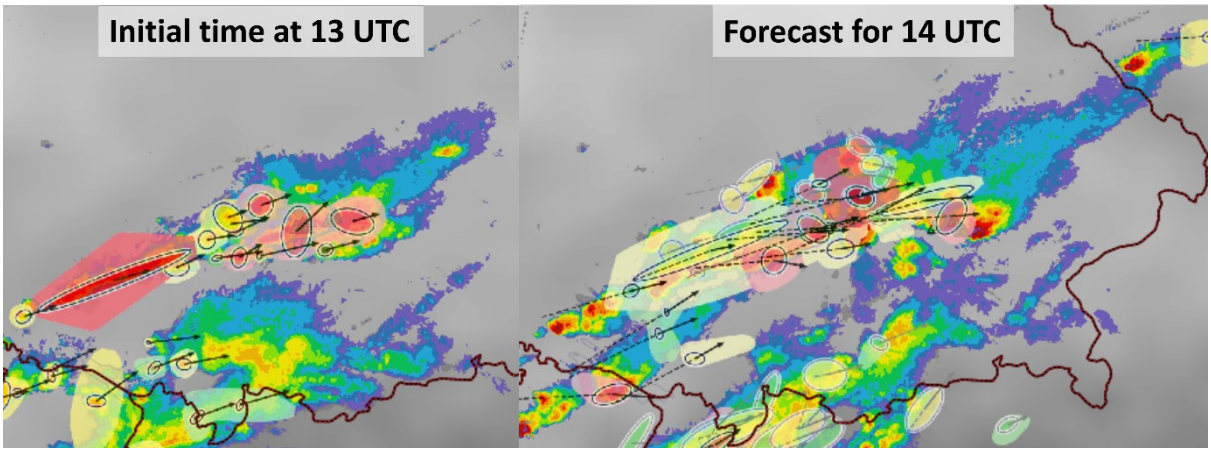
The probabilistic product was of no help at 19 UTC and showed too steep decrease of the probabilities of the severe cell both at 20 and 21 UTC nowcasts. For example, the probability of a severe cell was barely 50% at 21 UTC when using the 20 UTC nowcast. 21 UTC nowcast showed less than 20% of probability of severe cell by 22 UTC.

In summary, the KONRAD3D-SINFONY provided slightly better guidance concerning the maintenance of the high intensity of the convective system than the KONRAD3D-EPS. However, the intensity still decreased too fast. Furthermore, erratic behavior was noted with jumps in both location and the assigned severity of the pseudo member.

4.7.4 12 July 2023

On 12 July, a mix of multicells and supercells crossed southern Germany. Majority of storms had lifetimes higher than 60 minutes. Three supercells, two right-moving and one left-moving produced swaths of hail up to 5 cm in diameter. We explore the nowcasts of 13 and 16 UTC.

Compared to the KONRAD3D-EPS, more cells were assigned severe intensity in KONRAD3D-SINFONY at 13 UTC. The pseudo member representing the right-moving supercell had its intensity decreased to strong by 13:10 UTC and to weak by 13:30 UTC. Then, the intensity again increased to severe at 13:40 UTC, maintaining such intensity till 15 UTC. The increase in the intensity could be caused by the transition of the nowcast from NWC to NWP based. The location of the pseudo member was about 10 km away from the real location of the supercell at 14 UTC. Significant overlap of different cells with different sizes and intensities is visible for the cluster of the cells to the west of the supercell.



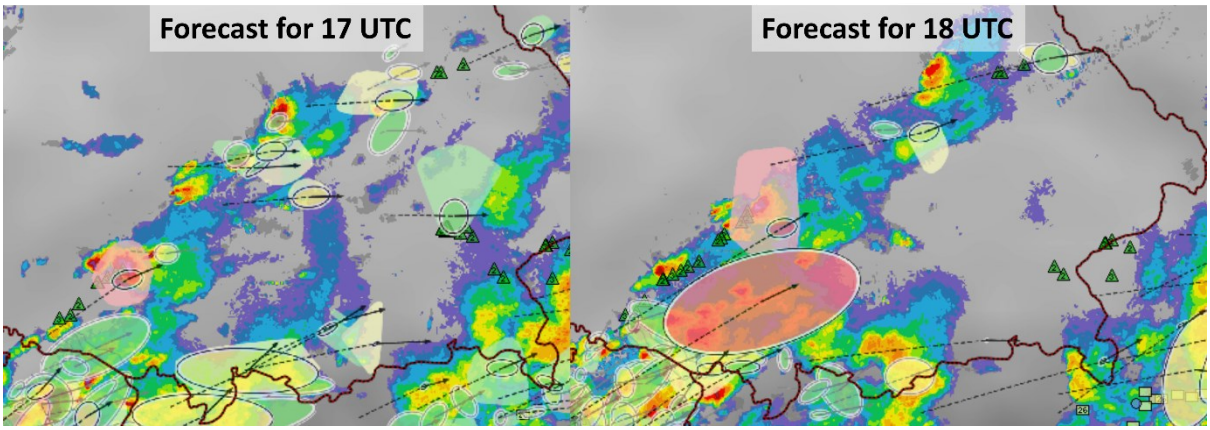
12 July 2023 observed radar reflectivity, ESWD reports in the last 3h and KONRAD3D-SINFONY pseudo member product shown for the nowcast initial time at 13 UTC and forecast for 14 UTC.

The 16 UTC nowcast maintained the intensity of the left-moving supercell for the whole two-hour period without any change. The pseudo member location was jumpy, but the supercell was always contained within the area of uncertainty. Some of the cells were not

picked up either by the NWC or NWP and were left undetected, both at 17 and 18 UTC timeframes. Neither of these storms produced severe weather in the end.

While the pseudo member visualization maintained the intensity as severe, the probability of severe cell decreased rapidly. The probability dropped to 50% already at 16:20 UTC and was below 20% at 17 UTC. After this period, the probability never exceeded 10%. Displaying the probability would thus make the forecaster much more uncertain concerning whether the storm would sustain its intensity over the next two hours.

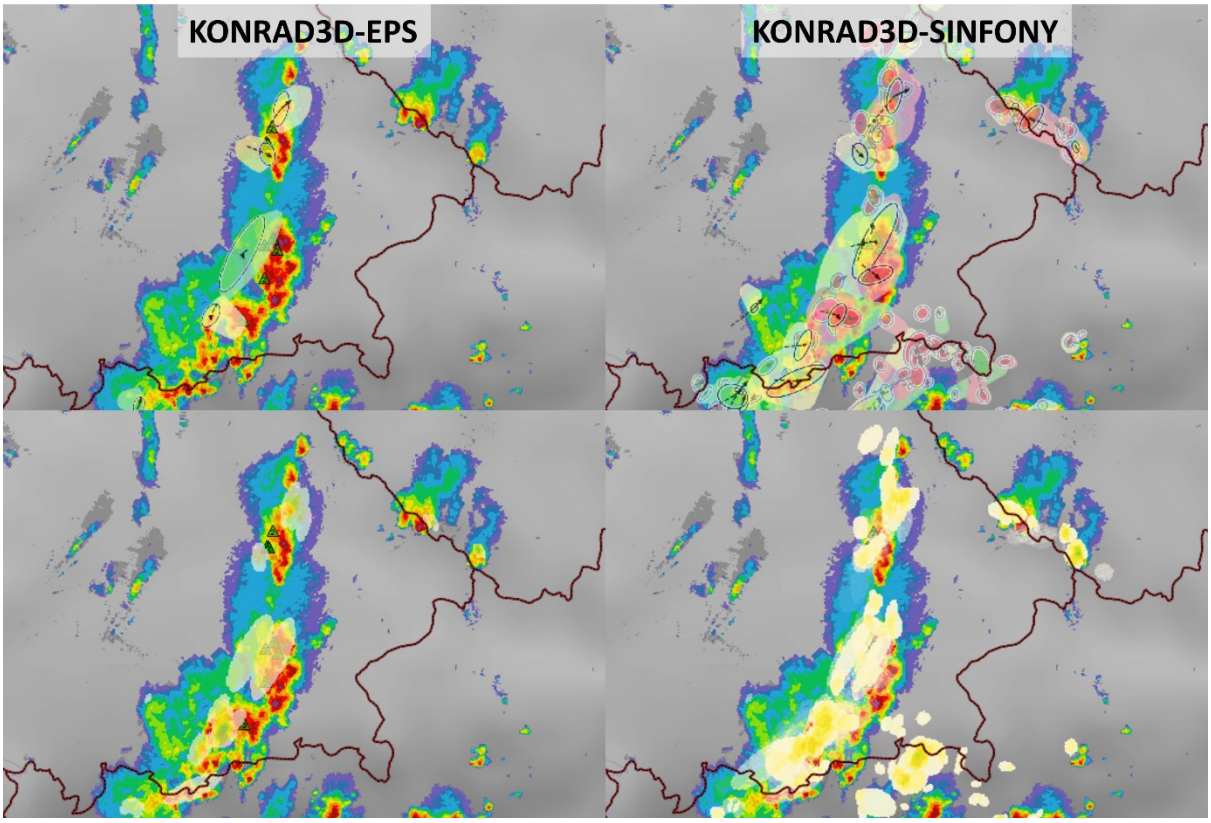
In summary, the pseudo member product of the KONRAD3D-SINFONY maintained the intensity of the supercells for longer time than the KONRAD3D-EPS. Jumps in the pseudo member location and overlap of different sized areas were noted as well.



12 July 2023 observed radar reflectivity, ESWD reports in the last 3h and KONRAD3D-SINFONY pseudo member product initiated at 16 UTC and displayed for 17 and 18 UTC.

4.7.5 15 August 2023

On 15 August, scattered storms formed over eastern Germany in weak mean flow and weak vertical wind shear. The storms existed in the form of rather short-lived cell and clusters. Thanks to the high CAPE, storms were capable of large hail and isolated severe wind gusts. On this day, KONRAD3D-EPS struggled with capturing the storm motion caused by the formation of new cells on the flanks of dying stationary cells. We investigate whether the NWP component of the SINFONY would alleviate the problem.



15 August 2023 observed radar reflectivity, ESWD reports in the last 3h and KONRAD3D-EPS and KONRAD3D-SINFONY pseudo member product and the probability of severe cell forecast for 14 UTC, initiated at 13 UTC.

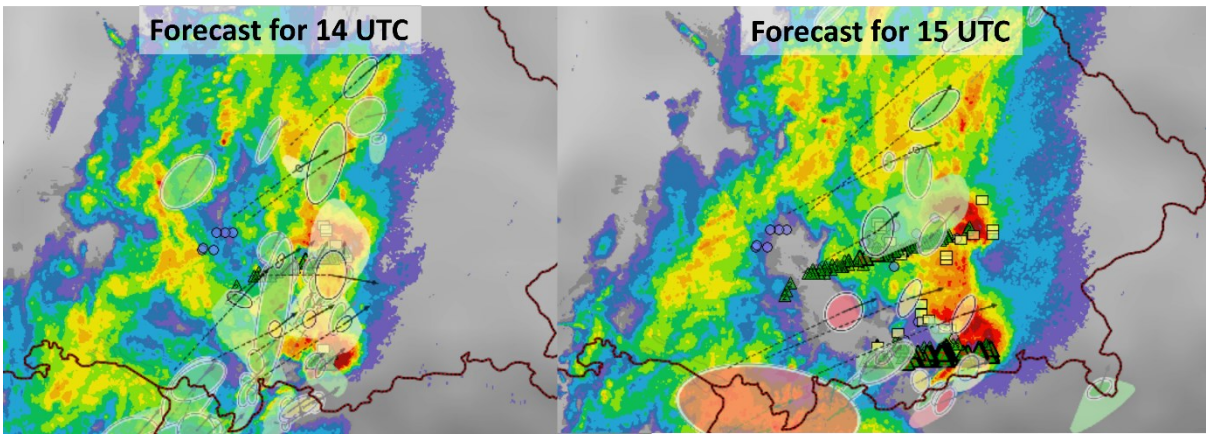
SINFONY tremendously improved the forecast compared to the EPS. 1h forecast started at 13 UTC had almost all storms covered by pseudo members and their hulls or by individual members in the probabilistic version of the forecast. NWP thus offers a lot of potential in situations when the simple extrapolation of the radar echo is not impossible, which is often the case in case of disorganized convective systems or systems that only move through development of new cells in a weak mean wind.

4.7.6 26 August

On 26 August, an outbreak of severe weather occurred over southern Germany. Two intense supercells developed around 13 UTC, creating swaths of hail and wind damage. The northern supercell produced hail up to 5 and the southern hail up to 10 cm in diameter. By 16 UTC, a bow-echo developed, which produced a swath of damaging winds with a wind gust of 35 m/s on Germany/Austria border. The event was well forecast one day ahead by high-resolution, convection allowing models.

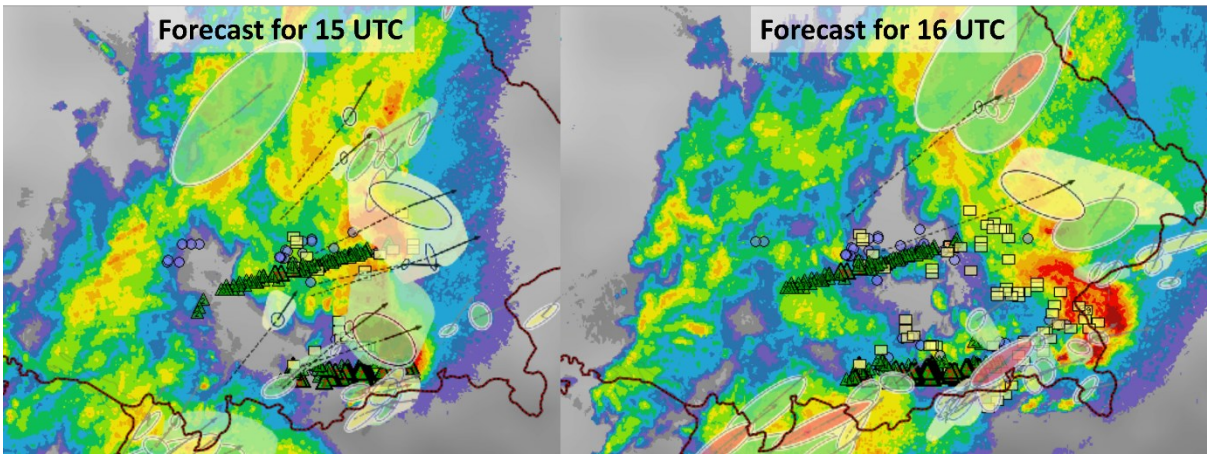
The first nowcast we investigate is 13 UTC. By that time, the northern supercell had already formed and produced large hail. The product identified it as an extremely severe cell. The southern supercell only initiated over the Alps and was still weak at that point. Therefore, this time frame offers a good opportunity to see if the NWP-based nowcast

could identify the intensification potential of the southern supercell. The northern supercell was predicted to weaken to severe by 13:40 UTC, while it remained extremely severe and produced the largest hail (5 cm) of its lifetime at that point. The southern cell became a supercell at 13:50 UTC with the first large hail reports. At 14 UTC forecast, it still wasn't covered by any detection. The northern cell also weakened to strong at that time and was forecast to remain weak by 15 UTC. A model-based severe cell was identified near the Alps at 14:30 UTC, remaining stationary and decaying by 15 UTC. The pseudo-member remained to the west of the severe weather track of the southern supercell.



26 August 2023 observed radar reflectivity, ESWD reports in the last 3h and KONRAD3D-SINFONY pseudo member product initiated at 13 UTC and displayed for 14 and 15 UTC.

KONRAD3D-EPS struggled with the nowcast of the southern supercell initiated at 14 UTC with too fast predicted decrease in its intensity. KONRAD3D-SINFONY offered just a slight improvement to that. While the cell lived longer than in the EPS, its weakening was still too fast in SINFONY. The analysis at 14 UTC was also different to the EPS. While the EPS correctly considered the southern supercell to be more intense than the northern one, SINFONY showed the southern one as severe and the northern one as extremely severe. EPS predicted the decay of the southern cell by 14:50 UTC. At that time, SINFONY still had the cell marked as strong. While this is clearly an improvement, the cell was in fact extremely severe and produced hail up to 10 cm around that time. The supercell outran weak and strong cell detections at 15:10 UTC. As the supercell started to produce severe wind gusts during the bow-echo transformation, no cell detections covered it. The NWP thus hasn't added too much information for this time. Probabilistic product didn't add much more help with probabilities of severe cell dropping below 25% after 14:30 UTC and to basically 0 at 15 UTC.



26 August 2023 observed radar reflectivity, ESWD reports in the last 3h and KONRAD3D-SINFONY pseudo member product initiated at 14 UTC and displayed for 15 and 16 UTC.

4.8 Questions about NowCastMIX

NowCastMIX has been thoroughly evaluated in the past Testbed editions. Here, we compare the performance of this, now standard operational tool of DWD, to the KONRAD3D-related products that are still under development.

4.8.1 Are the storm cell tracking vectors in NowCastMIX appropriate with respect to the resulting warning polygons (for the next 60 minutes)?

9 participant groups rated this question using a star system, ranging from “not at all” (1 star), through “in some cases” (2 stars) to “in most or all cases” (3 stars). The average score was 2.7 stars and the most common answer was “in most or all cases”. No group indicated “not at all”. Thus, the tracking vectors were found to be quite accurate on average by the participants.

In general, participants were happy with the tracking vectors in situations with longer-lived storms that moved relatively fast. This was for example the case of convective systems of 21 June or 11 July, or the supercells on 26 August. Worse performance was noted with lesser organized storms or in situations with slower storm motion, where the motion induced by new cell formation is the dominant component.

4.8.2 Are there systematic differences between the tracking vectors in NowCastMIX and KONRAD3D?

5 participant groups found no or only very small differences between the two products. 3 groups indicated that the NowCastMIX was better and 1 group found the KONRAD3D vector better. However, in case of one answer, the participants didn’t understand how to visualize the future tracking vector of KONRAD3D, which was the reason for them picking up the NowCastMIX as the better product.

Better or slightly better tracking vectors in NowCastMIX were mentioned especially in longer-lived and well-organized convective storms, such as the bow-echoes on 21 June or 11 July, or the supercells on 26 August. KONRAD3D was better in the situation with slower storm motion, such as 12 July.

4.8.3 Are the assessments of storm severity in NowCastMIX, as seen in the warning polygon level, appropriate?

9 participant groups rated this question using a star system, ranging from “not at all” (1 star), through “in some cases” (2 stars) to “in most or all cases” (3 stars). The average score was 2.7 stars, and the most common answer was “in most or all cases”. No group indicated “not at all”. Thus, the warning polygon levels were found to be appropriate on average by the participants.

Underestimation of the storm severity by NowCastMIX was noted in the later stages of the long-lived convective systems on 21 June or 11 July. In case of supercells on 26 August, NowCastMIX better identified the severity of the most intense storm. An overuse of the highest warning category was noted for the disorganized storms on 15 August.

4.8.4 Are there systematic differences between NowCastMIX and KONRAD3D in terms of storm severity?

5 participants noted small or no differences between the products. 3 groups noted that the NowCastMIX typically showed higher intensity and 1 group that the KONRAD3D showed higher intensity. For some cases, the storm severity assessment by the products differed during the storm life cycle.

4.8.5 Are the assessments of storm attributes appropriate in NowCastMIX?

Some of the participant groups answered this question. However, the product didn’t include the assessment of intensities of different phenomena and thus we don’t analyze the answers in this report.

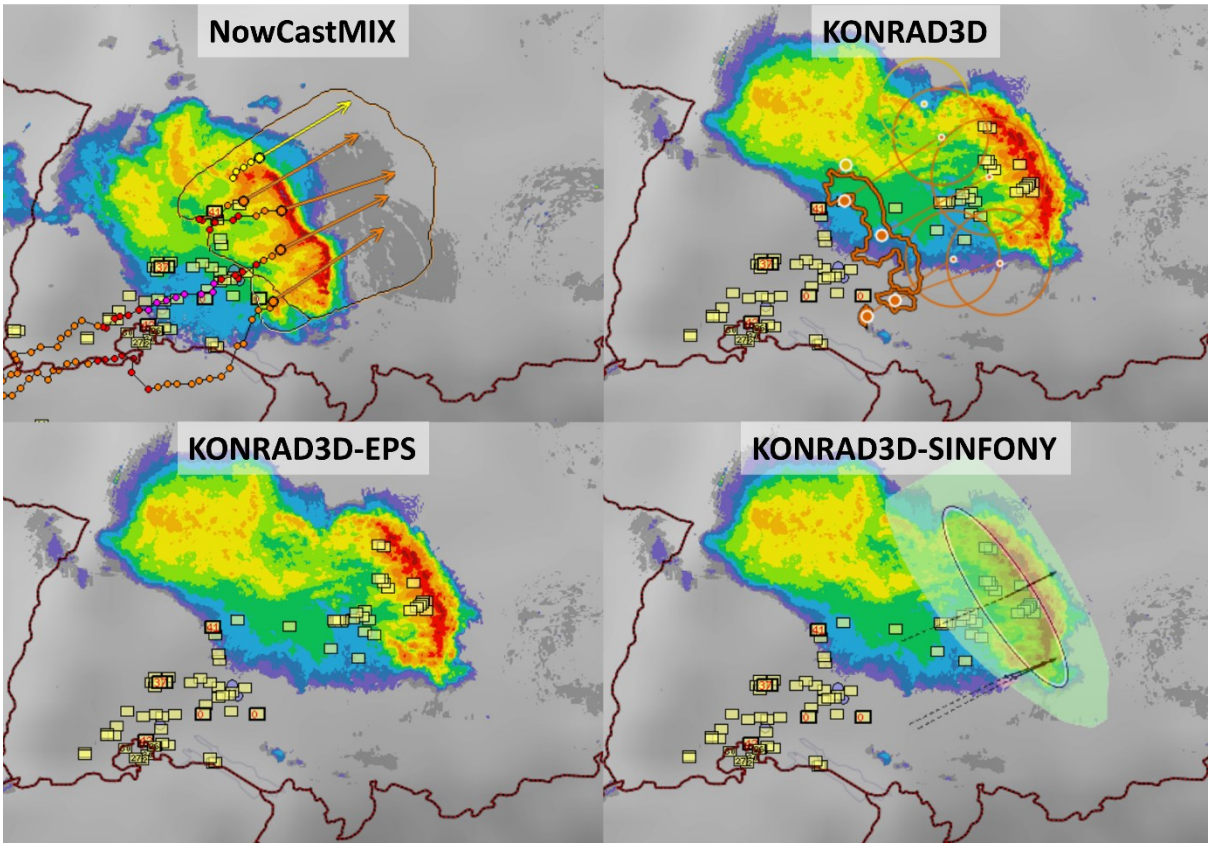
4.9 Cases for NowcastMIX

We look at two 3 cases of extremely severe storms, for which we discuss the differences among the 4 investigated nowcasting tools (NowCastMIX, KONRAD3D, KONRAD3D-EPS and KONRAD3D-SINFONY).

4.9.1 21 June 2023

In the morning hours of 21 June 2023, a cluster of storms has travelled from France through Switzerland to Germany. The cluster has organized into a fast-moving bow-echo, producing a swath of damaging wind gusts towards the border with Czechia. The damaging stage of the storm lasted from 7:30 to 12 UTC. Two wind gusts over 40 m/s were measured. The case was unique from the perspective of a convective windstorm occurring so early in the day and was not well forecasted the day before the event, making the correct nowcast very important.

At 9 UTC, the convective system was past its most intense stage, but was still a textbook bow-echo and capable of severe wind gusts. Both NowCastMIX and KONRAD3D identified the system as strong at 9 UTC. This was an underestimation by one level, as the system was still capable of severe weather. Both NowCastMIX and KONRAD3D provide a good nowcast concerning the track of the storm. KONRAD3D-EPS predicted the cell to decay by 10 UTC and KONRAD3D-SINFONY predicted the cell to be weak by 10 UTC. The prediction of the future path of the storm was correct. To summarize, the nowcast of the products that assumed the constant intensity of the storm throughout the was more successful.

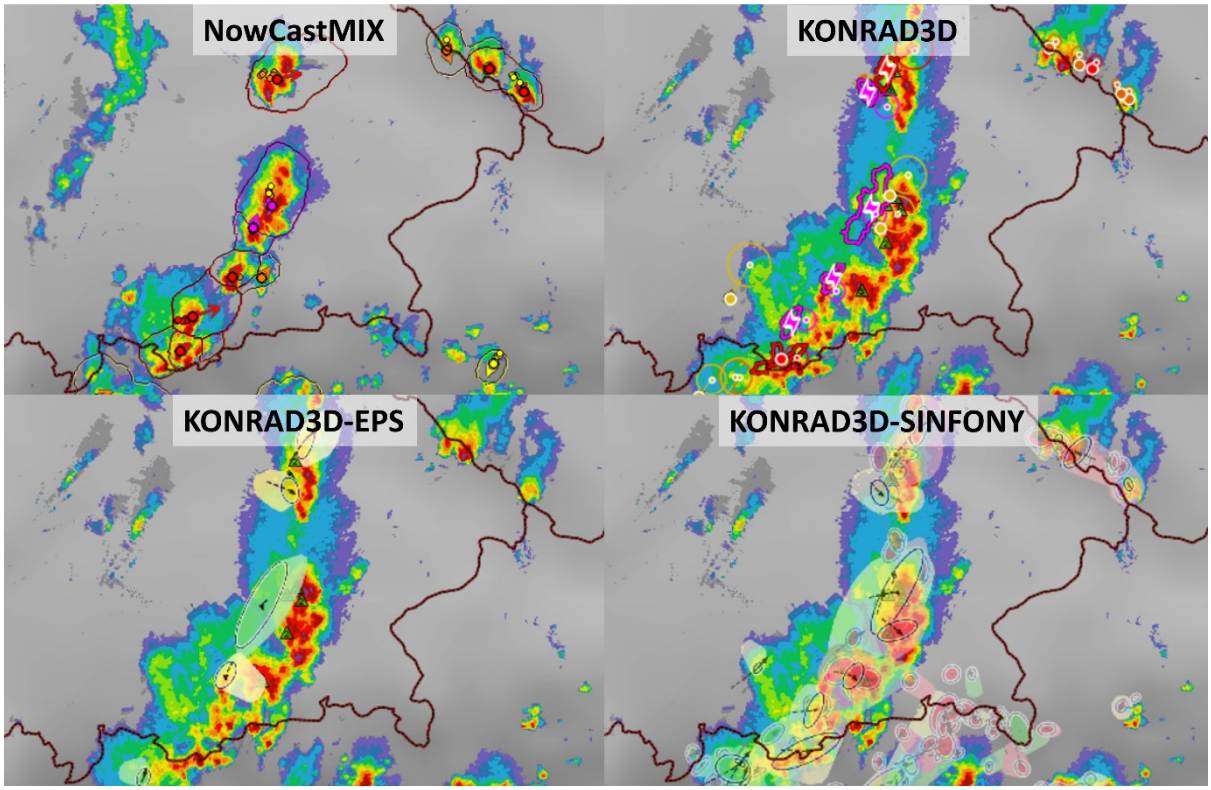


21 June 2023 observed radar reflectivity, ESWD reports in the last 3h, and NowCastMIX warning polygon at 9 UTC, and KONRAD3D, KONRAD3D-EPS, KONRAD3D-SINFONY pseudo member product 10 UTC forecast initiated at 9 UTC.

4.9.2 15 August 2023

On 15 August, scattered storms formed over eastern Germany in weak mean flow and weak vertical wind shear. The storms existed in the form of rather short-lived cell and clusters. Thanks to the high CAPE, storms were capable of large hail and isolated severe wind gusts.

Due to the weak prevailing flow, the storm motion was primarily due to the formation of the new cells at the flanks of the dying ones. This proved to be a challenge to the nowcasting products. At 13 UTC, several clusters of cells were detected over southern Germany. The northern one was identified as severe by NowCastMIX and severe/extremely severe by KONRAD3D. Large hail was observed shortly after 13 UTC, but not of extremely severe intensity. NowCastMIX provided better warning polygon than KONRAD3D, which missed the future location of cells. KONRAD3D-EPS both missed the future intensity and location of cells within this cluster. KONRAD3D-SINFONY underestimated the intensity, but at least covered the future location of cells.



15 August 2023 observed radar reflectivity, ESWD reports in the last 3h, and NowCastMIX warning polygon at 13 UTC, and KONRAD3D, KONRAD3D-EPS, KONRAD3D-SINFONY pseudo member product forecast of 14 UTC, initiated at 13 UTC.

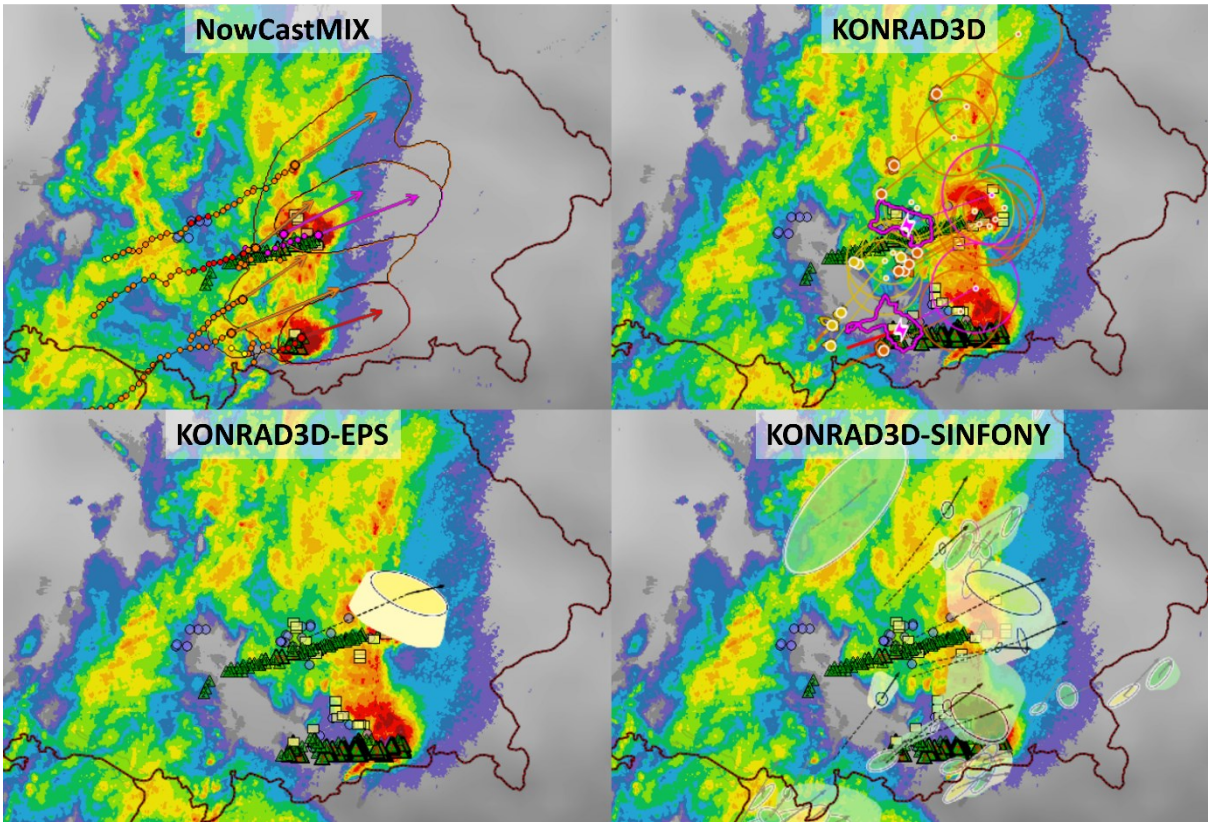
The central cluster extremely severe intensity assigned both by NowCastMIX and KONRAD3D. This was a slight overestimation, as the storm only matched severe criteria. For an unknown reason, KONRAD3D nowcasted the storm as strong, two levels weaker compared to the analysis. The storm stayed severe through 14 UTC. Only KONRAD3D-

SINFONY managed to correctly predict the motion of the storm over the next hour. It also managed to maintain higher storm intensity. KONRAD3D-EPS both failed to forecast the maintenance of the intensity of the storm and the future storm motion. An additional value of the NWP for forecasting future cell development was crucial in this case.

4.9.3 26 August 2023

On 26 August, an outbreak of severe weather occurred over southern Germany. Two intense supercells developed around 13 UTC, creating swaths of hail and wind damage. The northern supercell produced hail up to 5 and the southern hail up to 10 cm in diameter. By 16 UTC, a bow-echo developed, which produced a swath of damaging winds with a wind gust of 35 m/s on Germany/Austria border. At 13 UTC, northern supercell already produced a long swath of wind-driven hail and a southern supercell has just produced first reports of hail > 5 cm. NowCastMIX marked the northern supercell as extremely severe and southern supercell as severe. KONRAD3D marked both as extremely severe. At this point, the southern supercell was more severe than the northern one.

The northern supercell weakened to some degree past 13 UTC but was still capable of large hail by 14 UTC. The southern supercell maintained extremely severe intensity all the way through 14 UTC. This makes the constant intensity predictions of NowCastMIX and KONRAD3D better than both KONRAD3D-EPS and KONRAD3D-SINFONY. KONRAD3D-EPS pseudo-member product was especially misleading by decaying the southern supercell by 14 UTC. The path of the northern supercell was well captured both by NowCastMIX and KONRAD3D. The pseudo-member of the KONRAD3D-EPS was northwest of the real position of cell by 14 UTC. The path of the southern supercell deviated further south than the products predicted, with hailswath eventually moving slightly out of the warned areas.



26 August 2023 observed radar reflectivity, ESWD reports in the last 3h, and NowCastMIX warning polygon at 13 UTC, and KONRAD3D, KONRAD3D-EPS, KONRAD3D-SINFONY pseudo member product forecast of 14 UTC, initiated at 13 UTC.

#1

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Wednesday, June 28, 2023 2:50:38 PM
Last Modified:	Wednesday, June 28, 2023 4:30:41 PM
Time Spent:	01:40:02
IP Address:	80.109.154.58

Page 1: Evaluation

Q1

Does RUC simulate more realistic reflectivity patterns of convective storms compared to the ICON-D2? Think about squall lines, supercells, updraft cores, stratiform regions, and their shapes.

☆

A bit more realistic

Motivate your answer::

21.06.2023. 6 UTC model output has quite good results, however the 00z output only reasonable results at ICON-RUC-EPS (not ICON D2 nor normal RUC)

Q2

Compare the timing of the storm initiation in the RUC to the reality.

On point,

1 hour too late,

Further comments::

RUC is slighly slower than the real system but stays closer to the storm than D2

Q3

For cold-pool-driven storms, can you say something about the simulated cold pools in comparison to observations)?

About right,

Too slow moving,

Moving at correct speed

Q4

Does RUC correctly forecast the duration (life-cycle) of the storms?

Squall lines life cycle simulated correctly,

Squall lines persist too long

Q5

Are the ensemble forecasts of the RUC reliable (reflectivity, updraft helicity)? In the case of intense and small-scale convection, we have to expect low exceedance probabilities for the higher thresholds over a wider region, especially for longer lead times. Do highlighted regions cover the observed cells?

Sensible balance of misses, hits and false alarms,

Elaborate further::

regions fairly good the actual cell for most of the area

Q6

Have you noticed a large difference in a specific parameter (such as updraft helicity or reflectivity) in RUC that is clearly better or worse than ICON-D2?

init 6UTC - the RUC had higher reflectivities than D2 especially later; wind gust at RUC did not match the observation close to the init, but caught up later (boarder to czech republic) while D2 did not have much there
0UTC init, RUC-EPS did catch the storm, while D2 failed in time and space with the storm

Q7

In the actual weather situation that you look at, does the rapid hourly forecast update help in any way to give more precise warnings?

Motivate your answer:

not clear for that case, because important time would be around midnight, where they are missing

#2

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Friday, June 30, 2023 9:44:26 AM
Last Modified: Friday, June 30, 2023 10:16:09 AM
Time Spent: 00:31:43
IP Address: 80.109.154.58

Page 1: Evaluation

Q1

Does RUC simulate more realistic reflectivity patterns of convective storms compared to the ICON-D2? Think about squall lines, supercells, updraft cores, stratiform regions, and their shapes.

☆

Much more realistic

Motivate your answer::

In the case of 22.6.2023 It showed a little bit more convective initiation.

Q2

Compare the timing of the storm initiation in the RUC to the reality.

1 hour too early,

On point,

2+ hours too late,

Further comments::

In Austria it could not predict storm initiation on time, in eastern Germany it actually worked quite well (overestimation of dry and hot air behind the mountains due to fohn?)

Q3

For cold-pool-driven storms, can you say something about the simulated cold pools in comparison to observations)?

About right

Q4

Does RUC correctly forecast the duration (life-cycle) of the storms?

Single cell life cycle simulated correctly,

Supercells persist too long

Q5

Are the ensemble forecasts of the RUC reliable (reflectivity, updraft helicity)? In the case of intense and small-scale convection, we have to expect low exceedance probabilities for the higher thresholds over a wider region, especially for longer lead times. Do highlighted regions cover the observed cells?

Sensible balance of misses, hits and false alarms,

Elaborate further::

We think that in this particual situation ICON RUC EPS was quite reliable in predicting the highest probability of storm development and movement. More than the deterministic version.

Q6

Have you noticed a large difference in a specific parameter (such as updraft helicity or reflectivity) in RUC that is clearly better or worse than ICON-D2?

RUC initiates convection far more easily. In some runs too much.

Q7

In the actual weather situation that you look at, does the rapid hourly forecast update help in any way to give more precise warnings?

Yes, both for shorter and longer lead times,

Motivate your answer:

The model was able to predict quite soon, that the storms in the area of western Bohemia and Austria will not be as strong as was predicted before.

#3

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Wednesday, July 05, 2023 4:01:34 PM
Last Modified: Wednesday, July 05, 2023 4:35:58 PM
Time Spent: 00:34:24
IP Address: 80.109.154.58

Page 1: Evaluation

Q1

Does RUC simulate more realistic reflectivity patterns of convective storms compared to the ICON-D2? Think about squall lines, supercells, updraft cores, stratiform regions, and their shapes.

☆

Much more realistic

Motivate your answer::

22 of June 2023 at 18 UTC: RUC showed much better simul. radar signal then D2. D2 was way too weak.

Q2

Compare the timing of the storm initiation in the RUC to the reality.

On point,

Further comments::

In our case (2 hour forecast) it was more or less on point.

Q3

For cold-pool-driven storms, can you say something about the simulated cold pools in comparison to observations)?

Too strong,

Moving at correct speed

Q4

Does RUC correctly forecast the duration (life-cycle) of the storms?

Single cell life cycle simulated correctly,

Supercells life cycle simulated correctly

Q5

Are the ensemble forecasts of the RUC reliable (reflectivity, updraft helicity)? In the case of intense and small-scale convection, we have to expect low exceedance probabilities for the higher thresholds over a wider region, especially for longer lead times. Do highlighted regions cover the observed cells?

Sensible balance of misses, hits and false alarms,

Elaborate further::

We only considered very short lead time of 2 hours.

Q6

Have you noticed a large difference in a specific parameter (such as updraft helicity or reflectivity) in RUC that is clearly better or worse than ICON-D2?

Simulated reflectivity was clearly better in RUC than in D2.
Wind gusts were too strong in both models.

Q7

In the actual weather situation that you look at, does the rapid hourly forecast update help in any way to give more precise warnings?

Yes, but only for shorter lead times (up to 3h)

#4

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Saturday, September 09, 2023 12:13:01 PM
Last Modified: Saturday, September 09, 2023 1:53:08 PM
Time Spent: 01:40:07
IP Address: 94.142.239.106

Page 1: Evaluation

Q1

Does RUC simulate more realistic reflectivity patterns of convective storms compared to the ICON-D2? Think about squall lines, supercells, updraft cores, stratiform regions, and their shapes.

☆

Motivate your answer::

A bit more realistic

11 July 2023 E France/S Germany: Both ICON-D2 and RUC simulate the transition from isolated supercells to a large convective system (bow echo). The bow echo looks considerably weaker in ICON-D2 than in the RUC. The storm cores are weaker than reality in ICON-D2, but stronger in the RUC. This concerns both the supercell cores, squall line cores, and the stratiform precipitation regions.

Q2

Compare the timing of the storm initiation in the RUC to the reality.

On point,

Further comments::

The timing was absolutely on point when using the 13 UTC ensemble forecast.

Q3

For cold-pool-driven storms, can you say something about the simulated cold pools in comparison to observations)?

Too strong,

Moving at correct speed

Q4

Does RUC correctly forecast the duration (life-cycle) of the storms?

Squall lines life cycle simulated correctly,

Supercells life cycle simulated correctly

Q5

Are the ensemble forecasts of the RUC reliable (reflectivity, updraft helicity)? In the case of intense and small-scale convection, we have to expect low exceedance probabilities for the higher thresholds over a wider region, especially for longer lead times. Do highlighted regions cover the observed cells?

Sensible balance of misses, hits and false alarms,

Elaborate further::

For a longer lead time, the RUC ensemble had a lower probability than ICON-D2. For shorter lead time, RUC showed higher probabilities.

Q6

Have you noticed a large difference in a specific parameter (such as updraft helicity or reflectivity) in RUC that is clearly better or worse than ICON-D2?

The updraft helicity values are higher in RUC than in ICON-D2

Q7

In the actual weather situation that you look at, does the rapid hourly forecast update help in any way to give more precise warnings?

Yes, both for shorter and longer lead times

#5

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Monday, September 11, 2023 4:03:57 PM
Last Modified: Monday, September 11, 2023 5:28:56 PM
Time Spent: 01:24:58
IP Address: 94.142.239.106

Page 1: Evaluation

Q1

Does RUC simulate more realistic reflectivity patterns of convective storms compared to the ICON-D2? Think about squall lines, supercells, updraft cores, stratiform regions, and their shapes.

☆

Much more realistic

Motivate your answer::

15 August: ICON-RUC produces much more realistic storms than ICON-D2, especially in terms of the intense reflectivity cores. That said, RUC simulates too-intense hail cores. Cold pools are too strong in the RUC especially when intense isolated cores are present.

Q2

Compare the timing of the storm initiation in the RUC to the reality.

On point,

Further comments::

Fluctuating between 1 hour too early and 1 hour too late.

Q3

For cold-pool-driven storms, can you say something about the simulated cold pools in comparison to observations)?

Too strong,

Moving at correct speed

Q4

Does RUC correctly forecast the duration (life-cycle) of the storms?

Single cell life cycle simulated correctly,

Squall lines life cycle simulated correctly

Q5

Are the ensemble forecasts of the RUC reliable (reflectivity, updraft helicity)? In the case of intense and small-scale convection, we have to expect low exceedance probabilities for the higher thresholds over a wider region, especially for longer lead times. Do highlighted regions cover the observed cells?

Sensible balance of misses, hits and false alarms

Q6

Have you noticed a large difference in a specific parameter (such as updraft helicity or reflectivity) in RUC that is clearly better or worse than ICON-D2?

Reflectivity definitely better in RUC than ICON-D2.

Q7

In the actual weather situation that you look at, does the rapid hourly forecast update help in any way to give more precise warnings?

Yes, both for shorter and longer lead times

#6

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Wednesday, September 13, 2023 10:25:20 PM
Last Modified: Wednesday, September 13, 2023 11:51:02 PM
Time Spent: 01:25:42
IP Address: 160.45.74.5

Page 1: Evaluation

Q1

Does RUC simulate more realistic reflectivity patterns of convective storms compared to the ICON-D2? Think about squall lines, supercells, updraft cores, stratiform regions, and their shapes.

☆

Much more realistic

Motivate your answer::

Case July 12 2023 11-14 UTC, S Germany, convective line: ICON-D2 does not capture the mode properly as it indicates rather isolated convection. RUC is better, it indicates the linear mode much more.

Q2

Compare the timing of the storm initiation in the RUC to the reality.

On point,

Further comments::

July 12 2023, S Germany 11 UTC, CI along a front: RUC was really good with respect to the time of initiation. At some locations, RUC is too early as convection failed to further develop in reality, some other storms were inicated a little too late.

Q3

For cold-pool-driven storms, can you say something about the simulated cold pools in comparison to observations)?

About right,

Moving at correct speed

Q4

Does RUC correctly forecast the duration (life-cycle) of the storms?

Single cell life cycle simulated correctly,

Squall lines life cycle simulated correctly,

Supercells life cycle simulated correctly

Q5

Are the ensemble forecasts of the RUC reliable (reflectivity, updraft helicity)? In the case of intense and small-scale convection, we have to expect low exceedance probabilities for the higher thresholds over a wider region, especially for longer lead times. Do highlighted regions cover the observed cells?

Sensible balance of misses, hits and false alarms,

Elaborate further::

Case July 12, 2023, large hail occurred in areas highlighted by UH and high reflectivity.

Q6

Have you noticed a large difference in a specific parameter (such as updraft helicity or reflectivity) in RUC that is clearly better or worse than ICON-D2?

On July 12, 2023, 12 UTC +2h, RUC UH looked better than ICON-D2.

Q7

In the actual weather situation that you look at, does the rapid hourly forecast update help in any way to give more precise warnings?

Yes, but only for shorter lead times (up to 3h),

Motivate your answer:

Best results for the following 2 hours.

#7

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Monday, September 18, 2023 6:34:46 PM
Last Modified: Monday, September 18, 2023 6:46:19 PM
Time Spent: 00:11:33
IP Address: 62.46.104.209

Page 1: Evaluation

Q1

Does RUC simulate more realistic reflectivity patterns of convective storms compared to the ICON-D2? Think about squall lines, supercells, updraft cores, stratiform regions, and their shapes.

☆

Much more realistic

Motivate your answer::

For the case of 10 July the performance of RUC was better then D2. Shapes and intensities of cells looked closer to reality than.

Q2

Compare the timing of the storm initiation in the RUC to the reality.

On point

Q3

For cold-pool-driven storms, can you say something about the simulated cold pools in comparison to observations)?

Respondent skipped this question

Q4

Does RUC correctly forecast the duration (life-cycle) of the storms?

Single cell life cycle simulated correctly,

Supercells life cycle simulated correctly

Q5

Are the ensemble forecasts of the RUC reliable (reflectivity, updraft helicity)? In the case of intense and small-scale convection, we have to expect low exceedance probabilities for the higher thresholds over a wider region, especially for longer lead times. Do highlighted regions cover the observed cells?

Sensible balance of misses, hits and false alarms,

Elaborate further::

Difficult forecast for 10 July was well-balanced over Bavaria and W Czech Republic.

Q6

Have you noticed a large difference in a specific parameter (such as updraft helicity or reflectivity) in RUC that is clearly better or worse than ICON-D2?

On 10 July 2023 RUC seems to be clearly better in general, as the convective regime was better reflected in the ensemble compared to the D2 eps.

Q7

In the actual weather situation that you look at, does the rapid hourly forecast update help in any way to give more precise warnings?

Yes, both for shorter and longer lead times,

Motivate your answer:

there is a clear positive effect of the RUC for 9 and 12 UTC runs on 10 July 2023

#8

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Tuesday, September 19, 2023 3:14:27 PM
Last Modified:	Tuesday, September 19, 2023 3:27:18 PM
Time Spent:	00:12:51
IP Address:	62.46.104.209

Page 1: Evaluation

Q1

Does RUC simulate more realistic reflectivity patterns of convective storms compared to the ICON-D2? Think about squall lines, supercells, updraft cores, stratiform regions, and their shapes.

☆

A bit more realistic

Motivate your answer::

For case of 2 August NW Germany tornadoes around 12:40 and 13:00 UTC: both models were quite good, RUC slightly better but a bit overestimating reflectivity. RUC reflectivity patterns slightly more realistic.

Q2

Compare the timing of the storm initiation in the RUC to the reality.

On point,

1 hour too late

Q3

For cold-pool-driven storms, can you say something about the simulated cold pools in comparison to observations)?

Moving at correct speed

Q4

Does RUC correctly forecast the duration (life-cycle) of the storms?

Single cell life cycle simulated correctly

Q5

Are the ensemble forecasts of the RUC reliable (reflectivity, updraft helicity)? In the case of intense and small-scale convection, we have to expect low exceedance probabilities for the higher thresholds over a wider region, especially for longer lead times. Do highlighted regions cover the observed cells?

Sensible balance of misses, hits and false alarms

Q6

Have you noticed a large difference in a specific parameter (such as updraft helicity or reflectivity) in RUC that is clearly better or worse than ICON-D2?

No clear differences for 6 and 9 UTC runs. Both EPS had the highest density of UH signals in the wider region around the times and places of the tornadoes.

Q7

In the actual weather situation that you look at, does the rapid hourly forecast update help in any way to give more precise warnings?

Yes, both for shorter and longer lead times,

Motivate your answer:

rapid hourly forecast updates can confirm the region of highest risk

#1

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Thursday, July 06, 2023 4:01:02 PM
Last Modified:	Thursday, July 06, 2023 4:34:54 PM
Time Spent:	00:33:51
IP Address:	80.109.154.58

Page 1: Evaluation

Q1

Does the product provide a useful overview of the precipitation forecast?

☆

in some cases

Is it more useful in convective or stratiform situations?:

In convective situations

Q2

Is the coverage meaningful? Note, the size of the coverage reflects theuncertainty of the ensemble when showing precipitation at different locations.

Coverage far too small,

Further comments::

It was not far too small, but in some areas to coverage was a little bit too small, for example in Bavaria and Baden Württemberg.

Q3

Is there a tendency to over- or under-forecast precipitation when using probabilities?

Under-forecast,

Further comments::

On our forecast day (30 June 2023) it showed a little to less precipitation in some areas.

Q4

How do you assess the Radar Maxima compared to the uncalibrated similar product based on SINFONY/RUC/Intense? Do you see a benefit in calibration using radar data?

☆

Benefit in some cases

Further comments::

We looked at RUC in comparison to the radar maxima product and RUC showed some better signals in southern Germany than the maxima product.

Q5

Did you encounter gross failures in the Radar Maxima product that may occur from technical issues, such as spots with unrealistic high values, holes, lines, or other non-meteorological artifacts?

We found some major problems in the alps, where the radar maxima product shows unrealistic signals due to topography footprints. It could possibly be connected with the usage of model fields in a certain height, which show no data because the datafield is inside the modelorography.

#2

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Saturday, September 09, 2023 2:46:52 PM
Last Modified:	Monday, September 11, 2023 10:13:42 AM
Time Spent:	Over a day
IP Address:	94.142.239.106

Page 1: Evaluation

Q1

Does the product provide a useful overview of the precipitation forecast?

☆

in some cases

Is it more useful in convective or stratiform situations?:

11 July 2023. Very good forecast for the afternoon supercells over France and a bow-echo over S Germany. Worse forecast of afternoon isolated cells/supercells over Germany.

Q2

Is the coverage meaningful? Note, the size of the coverage reflects theuncertainty of the ensemble when showing precipitation at different locations.

Coverage OK,

Further comments::

With reservations, see the point above.

Q3

Is there a tendency to over- or under-forecast precipitation when using probabilities?

Under-forecast,

Further comments::

Underforecast for some storms, see above. Also looks like the precipitation may have been underestimated over S Germany, but this needs to be thoroughly verified with rain gauge network.

Q4

How do you assess the Radar Maxima compared to the uncalibrated similar product based on SINFONY/RUC/Intense? Do you see a benefit in calibration using radar data?

☆

Benefit in some cases

Q5

Did you encounter gross failures in the Radar Maxima product that may occur from technical issues, such as spots with unrealistic high values, holes, lines, or other non-meteorological artifacts?

Problems over mountains.

#3

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Monday, September 18, 2023 8:49:58 PM
Last Modified:	Monday, September 18, 2023 8:52:11 PM
Time Spent:	00:02:12
IP Address:	62.46.104.209

Page 1: Evaluation

Q1

Does the product provide a useful overview of the precipitation forecast?

☆

in some cases

Is it more useful in convective or stratiform situations?:

On the convective day of 10 July the product looked ok for the lower thresholds

Q2

Is the coverage meaningful? Note, the size of the coverage reflects theuncertainty of the ensemble when showing precipitation at different locations.

Coverage OK

Q3

Is there a tendency to over- or under-forecast precipitation when using probabilities?

Under-forecast,

Further comments::

No prob for larger 10 mm found for 10 July

Q4

How do you assess the Radar Maxima compared to the uncalibrated similar product based on SINFONY/RUC/Intense? Do you see a benefit in calibration using radar data?

☆

Benefit in some cases

Q5

Did you encounter gross failures in the Radar Maxima product that may occur from technical issues, such as spots with unrealistic high values, holes, lines, or other non-meteorological artifacts?

not found on 10 July

#4

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Tuesday, September 19, 2023 4:16:25 PM
Last Modified:	Tuesday, September 19, 2023 4:21:35 PM
Time Spent:	00:05:09
IP Address:	62.46.104.209

Page 1: Evaluation

Q1

Does the product provide a useful overview of the precipitation forecast?

☆ **in some cases**

Is it more useful in convective or stratiform situations?: for both

Q2

Is the coverage meaningful? Note, the size of the coverage reflects theuncertainty of the ensemble when showing precipitation at different locations.

Coverage OK

Q3

Is there a tendency to over- or under-forecast precipitation when using probabilities?

Forecast on spot,

Further comments::

seems ok for 2 August with fast moving cells

Q4

How do you assess the Radar Maxima compared to the uncalibrated similar product based on SINFONY/RUC/Intense? Do you see a benefit in calibration using radar data?

☆ **Benefit in most or all cases**

Further comments:: sure, there is a benefit from the calibration

Q5

Did you encounter gross failures in the Radar Maxima product that may occur from technical issues, such as spots with unrealistic high values, holes, lines, or other non-meteorological artifacts?

looks ok after comparison with ICON-D2 EPS upscale reflectivity

#5

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Tuesday, September 19, 2023 4:24:28 PM
Last Modified:	Tuesday, September 19, 2023 4:30:08 PM
Time Spent:	00:05:39
IP Address:	62.46.104.209

Page 1: Evaluation

Q1

Does the product provide a useful overview of the precipitation forecast?

☆ **in some cases**

Is it more useful in convective or stratiform situations?: for case of 15 Aug the EPS of 6 and 9 UTC only showed low probability for initiation over central Bavaria, resulting also in low probabilities within the radar maxima product

Q2

Is the coverage meaningful? Note, the size of the coverage reflects theuncertainty of the ensemble when showing precipitation at different locations.

Coverage OK

Q3

Is there a tendency to over- or under-forecast precipitation when using probabilities?

Respondent skipped this question

Q4

How do you assess the Radar Maxima compared to the uncalibrated similar product based on SINFONY/RUC/Intense? Do you see a benefit in calibration using radar data?

☆ **Benefit in some cases**

Further comments:: RUC with ini times at 10 and 11 UTC performed better

Q5

Did you encounter gross failures in the Radar Maxima product that may occur from technical issues, such as spots with unrealistic high values, holes, lines, or other non-meteorological artifacts?

Respondent skipped this question

#6

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Thursday, October 05, 2023 6:27:01 PM
Last Modified:	Thursday, October 05, 2023 6:31:17 PM
Time Spent:	00:04:15
IP Address:	94.142.239.106

Page 1: Evaluation

Q1

Does the product provide a useful overview of the precipitation forecast?

☆

in some cases

Is it more useful in convective or stratiform situations?:

15 August 2023, some overestimation was noted over central Germany and underestimation over Austria and W Czechia.

Q2

Is the coverage meaningful? Note, the size of the coverage reflects theuncertainty of the ensemble when showing precipitation at different locations.

Coverage far too large,

Further comments::

A bit too large

Q3

Is there a tendency to over- or under-forecast precipitation when using probabilities?

Further comments::

Depends on the region actually in this case.

Q4

How do you assess the Radar Maxima compared to the uncalibrated similar product based on SINFONY/RUC/Intense? Do you see a benefit in calibration using radar data?

☆

Benefit in some cases

Further comments::

This is tough to evaluate, because Radar Maxima hasn't been available with such a high temporal frequency to be used for nowcasting purposes.

Q5

Did you encounter gross failures in the Radar Maxima product that may occur from technical issues, such as spots with unrealistic high values, holes, lines, or other non-meteorological artifacts?

Respondent skipped this question

#1

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Wednesday, June 28, 2023 3:57:57 PM
Last Modified:	Wednesday, June 28, 2023 4:26:33 PM
Time Spent:	00:28:35
IP Address:	80.109.154.58

Page 1: Evaluation

Q1

Do the amplitude of the severity predictions of KONRAD3D-EPS outperform the constant severity predictions of KONRAD3D?

☆ **not at all**

Reasons:: In case of derecho od 21.6.2023 it constantly underestimated the severity of the system in its predictions.

Q2

Do the flexible lifetime predictions of KONRAD3D-EPS outperform the fixed lifetime prediction of KONRAD3D?

☆ **not at all**

Reasons:: No. Conrad EPS constantly expects the storm system to die - but it is not what happens. Conrad 3D was, however quite more succesfull.

Q3

How useful is the information provided by the KONRAD3D-EPS pseudo-member product?

☆ **useful only in some cases**

What is difficult or easy to understand from this product?: we did not find anything that could help us in this case, but it might be helpful in other cases.

Q4

How useful is the information provided by the KONRAD3D-EPS probability product?

☆ **useful only in some cases**

What is difficult and easy to understand from this product?: The scale is quite nice and understandable so it might be useful.

Q5

What is the maximum lead time, for which you would trust a lifetime forecast in the case of long-lived cells? (current plots are delivered for max 2 h lead time)

< 0.5 h,

Reasons::

1. It cannot keep up with the fast moving, long living storm system with forward propagation - especially in the morning. It constantly neglects the propagation of the system and counts only with the storm being advected by the flow - at least it seems that way 2. It constantly underestimates the severity of the system. We guess it is because the system in the morning with this strength does not 'fit in' the classical statistical archetype of a storm cell. Because statistically, these storms are far more rare, than ordinary multicell clusters or so. For isolated cells in the afternoon it was giving better results.

#2

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Friday, June 30, 2023 9:27:58 AM
Last Modified:	Friday, June 30, 2023 10:11:35 AM
Time Spent:	00:43:36
IP Address:	80.109.154.58

Page 1: Evaluation

Q1

Do the amplitude of the severity predictions of KONRAD3D-EPS outperform the constant severity predictions of KONRAD3D?

☆ **in most or all cases**

Reasons:: 22.06.23 18UTC: 3 Cells around munich the decay is well forecasted. Howerever new cells in front developed and the EPS took longer for the decay than the actual cells, so from a forecaster view the longer lifetime was a good indication that more can happen.

Q2

Do the flexible lifetime predictions of KONRAD3D-EPS outperform the fixed lifetime prediction of KONRAD3D?

☆ **in most or all cases**

Reasons:: From EPS is gets clear that the cells are in decaying phase. Especially the one over munich is well captured.

Q3

How useful is the information provided by the KONRAD3D-EPS pseudo-member product?

☆ **useful only in some cases**

What is difficult or easy to understand from this product?: in most case you only look at the pseudo-member and do not care too much of the shaded area around it. Especially with the southern cell decays within the shaded area.

Q4

How useful is the information provided by the KONRAD3D-EPS probability product?

☆ **useful only in some cases**

What is difficult and easy to understand from this product?: it helps for the area information but stays too long too strong (especially for the southern cell)

Q5 What is the maximum lead time, for which you would trust a lifetime forecast in the case of long-lived cells? (current plots are delivered for max 2 h lead time)	**< 0.5 h,** **0.5 - 1 h,** Reasons:: mainly looked at decaying phase

#3

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Wednesday, July 05, 2023 3:30:34 PM
Last Modified: Wednesday, July 05, 2023 4:21:44 PM
Time Spent: 00:51:10
IP Address: 80.109.154.58

Page 1: Evaluation

Q1

Do the amplitude of the severity predictions of KONRAD3D-EPS outperform the constant severity predictions of KONRAD3D?

☆ **in some cases**

Reasons:: KONRAD3D-EPS showed too much decay in long-lived supercells. For the first hour, intensity and core position was (much) better estimated by KONRAD3D.

Q2

Do the flexible lifetime predictions of KONRAD3D-EPS outperform the fixed lifetime prediction of KONRAD3D?

☆ **in some cases**

Reasons:: KONRAD3D-EPS showed too much decay in long-lived supercells. Assuming a more or less constant intensity seemed better for the next two hours for this type of storms. For less organized severe cells, KONRAD3D-EPS did give a useful signal of decay after around 1 hour.

Q3

How useful is the information provided by the KONRAD3D-EPS pseudo-member product?

☆ **not useful**

What is difficult or easy to understand from this product?: For the particular cell we focused on, the pseudo-member was on the wrong side of the uncertainty area. Whereas the real cell motion was south of the uncertainty area, the pseudo-member cell was positioned on the north side.

Q4

How useful is the information provided by the KONRAD3D-EPS probability product?

☆ **useful only in some cases**

What is difficult and easy to understand from this product?: It gives some useful indication of possible decay of an existing cell. However, it seems to completely ignore a deviant cell motion that was already going on for at least 30-45 minutes.

Q5

What is the maximum lead time, for which you would trust a lifetime forecast in the case of long-lived cells? (current plots are delivered for max 2 h lead time)

< 0.5 h,

0.5 - 1 h,

Reasons::

KONRAD3D-EPS decreases the severity of a long-lived supercell from severe to moderate after only 30 minutes, and from moderate to week after 70 minutes. However, in reality this cell remained severe for at least 4 hours.

#4

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Saturday, September 09, 2023 1:53:21 PM
Last Modified: Saturday, September 09, 2023 2:15:41 PM
Time Spent: 00:22:20
IP Address: 94.142.239.106

Page 1: Evaluation

Q1

Do the amplitude of the severity predictions of KONRAD3D-EPS outperform the constant severity predictions of KONRAD3D?

☆ **in some cases**

Reasons::

11 July. At 20 UTC, the algorithm predicted increasing intensity of the convective system, which was correct. In other cases (such as 21 UTC), the predicted intensity decreased too fast. At 22 UTC, KONRAD doesn't even detect the cells in the path of the strong outflow with weak convection.

Q2

Do the flexible lifetime predictions of KONRAD3D-EPS outperform the fixed lifetime prediction of KONRAD3D?

☆ **not at all**

Q3

How useful is the information provided by the KONRAD3D-EPS pseudo-member product?

☆ **useful only in some cases**

What is difficult or easy to understand from this product?:

The pseudo-member product helped to concentrate on the area with the highest risk of severe weather. The movement speed of the pseudo-member cell was too slow in the case of an accelerating system. Bad performance at 19 UTC, when no detection of severe cell was made for the system coming out of Switzerland that was already producing damaging wind gusts..

Q4

How useful is the information provided by the KONRAD3D-EPS probability product?

☆

useful only in some cases

What is difficult and easy to understand from this product?:

In some cases, the probability better outlines the potential movement of the convective system (e.g. 20 UTC). However, the probabilities decrease too fast for well-organized storms.

Q5

1 - 1.5 h

What is the maximum lead time, for which you would trust a lifetime forecast in the case of long-lived cells? (current plots are delivered for max 2 h lead time)

#5

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Monday, September 18, 2023 6:46:47 PM
Last Modified: Monday, September 18, 2023 8:36:39 PM
Time Spent: 01:49:52
IP Address: 62.46.104.209

Page 1: Evaluation

Q1

Do the amplitude of the severity predictions of KONRAD3D-EPS outperform the constant severity predictions of KONRAD3D?

☆ **in most or all cases**

Reasons:: Yes, on 10 July (10 to 13 UTC) the decaying model was in good correlation with the typical life cycles. This was due to relatively short-lived thunderstorms on that day.

Q2

Do the flexible lifetime predictions of KONRAD3D-EPS outperform the fixed lifetime prediction of KONRAD3D?

☆ **in some cases**

Reasons:: Yes, for 10 July rather short-lived thunderstorms this was the case.

Q3

How useful is the information provided by the KONRAD3D-EPS pseudo-member product?

☆ **useful only in some cases**

What is difficult or easy to understand from this product?: On 10 July the pseudo-member provided useful information.

Q4

How useful is the information provided by the KONRAD3D-EPS probability product?

☆ **very useful**

What is difficult and easy to understand from this product?: easy to understand on 10 July

Q5

What is the maximum lead time, for which you would trust a lifetime forecast in the case of long-lived cells? (current plots are delivered for max 2 h lead time)

1 - 1.5 h

#6

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Tuesday, September 19, 2023 3:40:59 PM
Last Modified: Tuesday, September 19, 2023 3:50:02 PM
Time Spent: 00:09:03
IP Address: 62.46.104.209

Page 1: Evaluation

Q1

Do the amplitude of the severity predictions of KONRAD3D-EPS outperform the constant severity predictions of KONRAD3D?

Reasons::

based on the cases of 2 Aug 2023 over NW Germany no clear answer can be given

Q2

Do the flexible lifetime predictions of KONRAD3D-EPS outperform the fixed lifetime prediction of KONRAD3D?

Reasons::

not in general

Q3

Respondent skipped this question

How useful is the information provided by the KONRAD3D-EPS pseudo-member product?

Q4

How useful is the information provided by the KONRAD3D-EPS probability product?

☆

useful only in some cases

What is difficult and easy to understand from this product?:

in principle it makes sense

Q5

What is the maximum lead time, for which you would trust a lifetime forecast in the case of long-lived cells? (current plots are delivered for max 2 h lead time)

< 0.5 h,

0.5 - 1 h,

Reasons::

for the case of 2 August the predictability is rather short

#7

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Wednesday, September 20, 2023 11:21:52 AM
Last Modified: Wednesday, September 20, 2023 11:47:29 AM
Time Spent: 00:25:36
IP Address: 89.14.207.254

Page 1: Evaluation

Q1

Do the amplitude of the severity predictions of KONRAD3D-EPS outperform the constant severity predictions of KONRAD3D?

☆ **in some cases**

Reasons:: 12 July 2023, 13 UTC, S Germany: The EPS looks much more precise with pretty small risk areas. However, some intense cells moved out of these regions. KONRAD3D had larger warning areas what was a better choice in this case with clustering and different storm movement vectors, although it overestimated the affected area.

Q2

Do the flexible lifetime predictions of KONRAD3D-EPS outperform the fixed lifetime prediction of KONRAD3D?

☆ **in some cases**

Reasons:: Indeed, some cells in S Germany disappeared after a short time what was correctly predicted by the EPS, and not by CONRAD3D. However, the supercell in the east of the convective area outran the EPS warning area, whereas CONRAD3D captured it.

Q3

How useful is the information provided by the KONRAD3D-EPS pseudo-member product?

☆ **not useful**

What is difficult or easy to understand from this product?: I do not see additional information as long as you have a time series that you can loop. The probabilities still indicate the expected storm struture in the beginning.

Q4

How useful is the information provided by the KONRAD3D-EPS probability product?

☆ **useful only in some cases**

What is difficult and easy to understand from this product?: Some of the storms moved away from the probability areas.

Q5

What is the maximum lead time, for which you would trust a lifetime forecast in the case of long-lived cells? (current plots are delivered for max 2 h lead time)

1 - 1.5 h,

Reasons::

After 1 hour, probabilities were low, and the number of misses increased.

#8

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Thursday, September 21, 2023 9:42:21 AM
Last Modified: Thursday, September 21, 2023 9:58:40 AM
Time Spent: 00:16:18
IP Address: 78.54.69.8

Page 1: Evaluation

Q1

Do the amplitude of the severity predictions of KONRAD3D-EPS outperform the constant severity predictions of KONRAD3D?

☆

in some cases

Reasons::

12 July 2023, 13 UTC, S Germany, comparing a linear cluster to a supercell at the eastern edge of the cluster: EPS captures the development of the linear cluster in a much better way compared to KONRAD3D, especially the linear structure is better displayed regarding the most intense parts. For the SC, EPS gives a bad prediction compared to CONRAD3D: Movement vector and severety are better predicted by CONRAD3D. In particular, the SC weakens extremely fast compared to the linear cluster to the west in the EPS prediction.

Q2

Do the flexible lifetime predictions of KONRAD3D-EPS outperform the fixed lifetime prediction of KONRAD3D?

☆

in some cases

Reasons::

12 July 2023, 13 UTC, S Germany, comparing a linear cluster to a supercell at the eastern edge of the cluster: EPS captures the development of the linear cluster in a much better way compared to KONRAD3D, especially the longevity of the linear structure is better displayed. For the SC, EPS gives a bad prediction compared to CONRAD3D: The lifetime is much better predicted by CONRAD3D. In particular, the SC weakens extremely fast compared to the linear cluster to the west in the EPS prediction.

Q3

How useful is the information provided by the KONRAD3D-EPS pseudo-member product?

☆

not useful

What is difficult or easy to understand from this product?:

12 July 2023, 13 UTC, S Germany, comparing a linear cluster to a supercell at the eastern edge of the cluster: For the linear cluster, it provides the same information as the probability display. For the SC, you can see that rapid weakening of the pseudo-member what gives indeed some additional information, but I did not find it useful.

Q4

How useful is the information provided by the KONRAD3D-EPS probability product?

☆

useful only in some cases

What is difficult and easy to understand from this product?:

12 July 2023, 13 UTC, S Germany, comparing a linear cluster to a supercell at the eastern edge of the cluster: It is useful for the convective cluster. For the SC, it looks that the EPS does only indicate a weak probability after a short time what is not too useful in this case. To me it seems that KONRAD3D outperforms the EPS for isolated, long-lived cells, whereas the EPS is better clusters.

Q5

What is the maximum lead time, for which you would trust a lifetime forecast in the case of long-lived cells? (current plots are delivered for max 2 h lead time)

< 0.5 h,

Reasons::

12 July 2023, 13 UTC, S Germany, comparing a linear cluster to a supercell at the eastern edge of the cluster: For the long-lived SC, the forecast was not good and I would not trust the forecast, in particular regarding the severity, but also the movement vector and longevity.

#9

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Thursday, October 05, 2023 5:51:16 PM
Last Modified: Thursday, October 05, 2023 6:06:50 PM
Time Spent: 00:15:33
IP Address: 94.142.239.106

Page 1: Evaluation

Q1

Do the amplitude of the severity predictions of KONRAD3D-EPS outperform the constant severity predictions of KONRAD3D?

☆ **in some cases**

Reasons:: 15 August weak mean flow situation over S Germany. EPS does suggest the weakening of the cells, which occurred.

Q2

Do the flexible lifetime predictions of KONRAD3D-EPS outperform the fixed lifetime prediction of KONRAD3D?

☆ **in some cases**

Q3

How useful is the information provided by the KONRAD3D-EPS pseudo-member product?

☆ **useful only in some cases**

What is difficult or easy to understand from this product?: New cell development through backbuilding along the cold pools or along the orography, which was difficult to capture for the system. However, it gave better information on the possible cell location than the KONRAD3D.

Q4

How useful is the information provided by the KONRAD3D-EPS probability product?

☆ **useful only in some cases**

What is difficult and easy to understand from this product?: See above, but definitely outperformed KONRAD3D in terms of the future cell locations.

Q5

What is the maximum lead time, for which you would trust a lifetime forecast in the case of long-lived cells? (current plots are delivered for max 2 h lead time)

0.5 - 1 h,

Reasons::

In this case, the possible lead time was rather short due to the nature of the situation (weaker mean flow and weaker shear)

#10

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Saturday, October 07, 2023 9:47:09 AM
Last Modified: Saturday, October 07, 2023 10:22:10 AM
Time Spent: 00:35:00
IP Address: 160.45.74.5

Page 1: Evaluation

Q1

Do the amplitude of the severity predictions of KONRAD3D-EPS outperform the constant severity predictions of KONRAD3D?

☆ **not at all**

Reasons:: 26 August 2023, 13-14 UTC, S Germany: The EPS version gave only few indications that a cell close to the Alps will increase its severity ot kep it (14 UTC). It did not really outperform Konrad3D. At 14 UTC, Konrad3D even looked better as it kept the severe intensity, whereas the EPS indicated a decay.

Q2

Do the flexible lifetime predictions of KONRAD3D-EPS outperform the fixed lifetime prediction of KONRAD3D?

☆ **not at all**

Reasons:: 26 August 2023, 13-14 UTC, S Germany: EPS indicated a decay that was too fast compared to reality. Konrad3D looked better.

Q3

How useful is the information provided by the KONRAD3D-EPS pseudo-member product?

☆ **useful only in some cases**

What is difficult or easy to understand from this product?: 26 August 2023, 13-14 UTC: The pseudo-member did not give much additional information in this case. However, the pseudo-member has a tendency for faster propagation (northern cell) and more eastward propagation (southern cell) compared to a EPS.

Q4

How useful is the information provided by the KONRAD3D-EPS probability product?

☆

useful only in some cases

What is difficult and easy to understand from this product?:

26 August 2023, 13-14 UTC: Although Konrad3D EPS did not perform very well in this situation, I still think that the EPS probabilities can help in forecasting, since the uncertainty is better vizualized.

Q5

What is the maximum lead time, for which you would trust a lifetime forecast in the case of long-lived cells? (current plots are delivered for max 2 h lead time)

1 - 1.5 h,

Reasons::

26 August 2023, 13-14 UTC: In this case, the EPS performed worse compared to Konrad3D. For the southern cell, only 1 hour of lifetime forecast. For the larger northern system, it looked better (1.5 hours, later even 2 hours).

#11

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Saturday, October 07, 2023 10:24:02 AM
Last Modified: Saturday, October 07, 2023 10:37:05 AM
Time Spent: 00:13:03
IP Address: 160.45.74.5

Page 1: Evaluation

Q1

Do the amplitude of the severity predictions of KONRAD3D-EPS outperform the constant severity predictions of KONRAD3D?

☆ **not at all**

Reasons:: 26 August 2023, 15 UTC: EPS indicated a decay after 1 hour, what was a worse forecast compared to Konrad3D.

Q2

Do the flexible lifetime predictions of KONRAD3D-EPS outperform the fixed lifetime prediction of KONRAD3D?

☆ **not at all**

Reasons:: 26 August 2023, 15 UTC: To fast decay in EPS. Konrad3D was better with the fixed lifetimes.

Q3

How useful is the information provided by the KONRAD3D-EPS pseudo-member product?

☆ **not useful**

What is difficult or easy to understand from this product?: 26 August 2023, 15 UTC: In this case of a fast propagating bow echo, the pseudo-member did not indicate useful additional information. The location and movement speed of the EPS looked better compared to the pseudo-member.

Q4

How useful is the information provided by the KONRAD3D-EPS probability product?

☆ **very useful**

What is difficult and easy to understand from this product?: 26 August 2023, 15 UTC: In this case, the probabilities really had some value: They indicated where to expect severe wind gusts (especially the probability for severe between 14:30 and 15 with high performance). The original EPS gives larger areas for severe, so that the probs where more accurate.

Q5

What is the maximum lead time, for which you would trust a lifetime forecast in the case of long-lived cells? (current plots are delivered for max 2 h lead time)

1.5 - 2 h,

Reasons::

26 August 2023, 15 UTC: For the severe bow echo, EPS was accurate to slightly beyond 1 hour lifetime forecast.

#12

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Saturday, October 07, 2023 10:37:26 AM
Last Modified:	Saturday, October 07, 2023 10:40:50 AM
Time Spent:	00:03:23
IP Address:	160.45.74.5

Page 1: Evaluation

Q1

Do the amplitude of the severity predictions of KONRAD3D-EPS outperform the constant severity predictions of KONRAD3D?

☆

in some cases

Reasons::

26 August 2023, 15 UTC: One addition to the evaluation: While the severity predictions did not look better than Konrad3D, the movement speed in the EPS looked better (the system moved fast compared to Konrad3D, and the EPS had a better performance, almost perfect).

Q2

Respondent skipped this question

Do the flexible lifetime predictions of KONRAD3D-EPS outperform the fixed lifetime prediction of KONRAD3D?

Q3

Respondent skipped this question

How useful is the information provided by the KONRAD3D-EPS pseudo-member product?

Q4

Respondent skipped this question

How useful is the information provided by the KONRAD3D-EPS probability product?

Q5

Respondent skipped this question

What is the maximum lead time, for which you would trust a lifetime forecast in the case of long-lived cells? (current plots are delivered for max 2 h lead time)

#1

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Wednesday, June 28, 2023 3:00:00 PM
Last Modified:	Wednesday, June 28, 2023 3:57:20 PM
Time Spent:	00:57:19
IP Address:	80.109.154.58

Page 1: Evaluation

Q1

How would you rate the quality of the combination of NWC and NWP?

☆ **Bad in most cases**

Q2

Are there noticeable artifacts in the forecasts that may result from the transition from the combination of NWC and NWP to forecasts that are purely based purely on NWP?

The ICON-Model had problems to forecast the whole system (0UTC-run). In the nowcast, it seems that the product is strongly influnced by the NWP, becasue the system is weakening in the Konrad-3Dsinfony but not in the Radar.

Q3

Are there artifacts in the watch regions (cells with gray borders) associated with the fact that the cells come from an NWP ensemble? E.g., cells too big, variety of different severities overlapping at the same place? If yes, does this add to the uncertainty in the forecast?

the size is ok, but maybe the overlapping of the cell can confuse, especially when the weaker one is winning

Q4

Is there any additional value of the combined cells (black borders) in terms of life cycle and severity compared to KONRAD3D-EPS?

☆ **only in some cases**

What is difficult or easy to understand about this product?: not r

Q5

How useful is the information provided by cells based purely on NWP predictions (grey borders) for shorter lead times of up to two hours?

☆ **not useful at all**

What is difficult or easy to understand about this product?: the cases are rare when the grey borders occure

Q6

How useful is the information provided by cells based purely on NWP predictions (grey borders) for longer lead times?

☆ **useful only in some cases**

Q7

How useful is the information provided by the KONRAD3D-SINFONY probability product? Consider that the probability product contains both the information from the combination of NWC and NWP and from pure NWP for future time steps.

☆ **not useful**

Q8

Does KONRAD3D-SINFONY provide further beneficial information for the forecast in comparison to KONRAD3D-EPS?

☆ **A little benefit or only in some cases**

#2

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Wednesday, July 05, 2023 3:29:24 PM
Last Modified:	Wednesday, July 05, 2023 4:28:41 PM
Time Spent:	00:59:17
IP Address:	80.109.154.58

Page 1: Evaluation

Q1

How would you rate the quality of the combination of NWC and NWP?

☆ **Good in most cases**

Q2

Are there noticeable artifacts in the forecasts that may result from the transition from the combination of NWC and NWP to forecasts that are purely based purely on NWP?

Yes - at 18:30 on 22 June 2023. It was confusing for us.

Q3

Are there artifacts in the watch regions (cells with gray borders) associated with the fact that the cells come from an NWP ensemble? E.g., cells too big, variety of different severities overlapping at the same place? If yes, does this add to the uncertainty in the forecast?

Yes. See Q2. It is more confusing than adding uncertainty. These artifacts probably should be filtered out.

Q4

Is there any additional value of the combined cells (black borders) in terms of life cycle and severity compared to KONRAD3D-EPS?

☆ **yes, totally**

What is difficult or easy to understand about this product?: Slightly complicated at first. Users need first to become familiar with the product.

Q5

How useful is the information provided by cells based purely on NWP predictions (grey borders) for shorter lead times of up to two hours?

☆ **useful only in some cases**

What is difficult or easy to understand about this product?: Artifacts need to be cleaned.

Q6

How useful is the information provided by cells based purely on NWP predictions (grey borders) for longer lead times?

☆	**very useful**

Q7

How useful is the information provided by the KONRAD3D-SINFONY probability product? Consider that the probability product contains both the information from the combination of NWC and NWP and from pure NWP for future time steps.

What is difficult and easy to understand from this product?:	The product was not available.

Q8

Does KONRAD3D-SINFONY provide further beneficial information for the forecast in comparison to KONRAD3D-EPS?

☆	**Provides a lot of benefit**
Comments::	The combination of NWC and NWP looks promising. We appreciate the combination of both.

#3

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Thursday, July 06, 2023 4:06:40 PM
Last Modified:	Thursday, July 06, 2023 4:52:04 PM
Time Spent:	00:45:23
IP Address:	80.109.154.58

Page 1: Evaluation

Q1

How would you rate the quality of the combination of NWC and NWP?

☆ **Somewhat good**

Q2

Are there noticeable artifacts in the forecasts that may result from the transition from the combination of NWC and NWP to forecasts that are purely based purely on NWP?

Yes, there are noticeable artefacts that appear in places where there was no extrapolated reflectivity, but where model simulated significant reflectivity.

Q3

Are there artifacts in the watch regions (cells with gray borders) associated with the fact that the cells come from an NWP ensemble? E.g., cells too big, variety of different severities overlapping at the same place? If yes, does this add to the uncertainty in the forecast?

Yes, some big poligons appear from NWP ensemble and it adds to the uncertainty of the forecast. They are also quite jumpy.

Q4

Is there any additional value of the combined cells (black borders) in terms of life cycle and severity compared to KONRAD3D-EPS?

☆ **yes, totally**

What is difficult or easy to understand about this product?: The product is quite complicated (a lot of jumpy visual information) so the visualization should be simplified in some way and made more usable for operational usage. Forecasters would also like to have a better idea about the meaning of different severity levels of the polygons.

Q5

How useful is the information provided by cells based purely on NWP predictions (grey borders) for shorter lead times of up to two hours?

☆ **very useful**

What is difficult or easy to understand about this product?: It seems to be quite useful, but it strongly depends on the situation (if the model correctly predict evollution of convection over the area).

Q6

How useful is the information provided by cells based purely on NWP predictions (grey borders) for longer lead times?

☆ **very useful**

Q7

How useful is the information provided by the KONRAD3D-SINFONY probability product? Consider that the probability product contains both the information from the combination of NWC and NWP and from pure NWP for future time steps.

☆ **very useful**

What is difficult and easy to understand from this product?: It is difficult to understand what exactly severity levels mean in probability product (strong, severe, extreme). Forecasters should have some idea how are they related to convective threats.

Q8

Does KONRAD3D-SINFONY provide further beneficial information for the forecast in comparison to KONRAD3D-EPS?

☆ **Provides a lot of benefit**

Comments:: The product is quite complicated and to evaluate it properly more experience is required in using the product. Also forecasters should have a good training (to understand how the product works and how is it supposed to be used) before using it and evaulating it.

#4

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Saturday, September 09, 2023 2:15:52 PM
Last Modified: Saturday, September 09, 2023 2:28:05 PM
Time Spent: 00:12:12
IP Address: 94.142.239.106

Page 1: Evaluation

Q1

How would you rate the quality of the combination of NWC and NWP?

☆ **Bad in most cases**

Q2

Are there noticeable artifacts in the forecasts that may result from the transition from the combination of NWC and NWP to forecasts that are purely based purely on NWP?

Yes, for example the 11 July 20 UTC forecast shows the pseudo-member for the MCS jump from one position to another in subsequent timeframes.

Q3

Are there artifacts in the watch regions (cells with gray borders) associated with the fact that the cells come from an NWP ensemble? E.g., cells too big, variety of different severities overlapping at the same place? If yes, does this add to the uncertainty in the forecast?

Yes, there are far too many cells overlapping each other, changing intensity and location rapidly. This makes product problematic to use.

Q4

Is there any additional value of the combined cells (black borders) in terms of life cycle and severity compared to KONRAD3D-EPS?

☆ **not at all**

What is difficult or easy to understand about this product?: It doesn't seem that the model simulation offered any additional guidance for the observed cells in this case.

Q5

How useful is the information provided by cells based purely on NWP predictions (grey borders) for shorter lead times of up to two hours?

☆ **useful only in some cases**

What is difficult or easy to understand about this product?: There has been some useful information from the grey cells appearing shortly after the forecast initialization. But they often disappear too quickly.

Q6

How useful is the information provided by cells based purely on NWP predictions (grey borders) for longer lead times?

☆ **useful only in some cases**

Q7

How useful is the information provided by the KONRAD3D-SINFONY probability product? Consider that the probability product contains both the information from the combination of NWC and NWP and from pure NWP for future time steps.

☆ **useful only in some cases**

What is difficult and easy to understand from this product?: Useful in some cases and better covers the area that can be impacted by storms than the pseudo-members alone. However, the probabilities decrease too quickly even if model simulate still an intense system.

Q8

Does KONRAD3D-SINFONY provide further beneficial information for the forecast in comparison to KONRAD3D-EPS?

☆ **A little benefit or only in some cases**

Comments:: Only a little benefit. Probabilities still decrease too quickly and pseudomembers provide confusing forecasts with far too many cells overlapping each other. Especially problematic was the forecast from 19 UTC with no coverage of severe linear segment arriving from Switzerland.

#5

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Monday, September 18, 2023 8:36:55 PM
Last Modified:	Monday, September 18, 2023 8:44:55 PM
Time Spent:	00:07:59
IP Address:	62.46.104.209

Page 1: Evaluation

Q1

How would you rate the quality of the combination of NWC and NWP?

☆ **Good in most cases**

Q2

Are there noticeable artifacts in the forecasts that may result from the transition from the combination of NWC and NWP to forecasts that are purely based purely on NWP?

Not found on 10 July

Q3

Are there artifacts in the watch regions (cells with gray borders) associated with the fact that the cells come from an NWP ensemble? E.g., cells too big, variety of different severities overlapping at the same place? If yes, does this add to the uncertainty in the forecast?

Not found on 10 July

Q4

Is there any additional value of the combined cells (black borders) in terms of life cycle and severity compared to KONRAD3D-EPS?

Respondent skipped this question

Q5

How useful is the information provided by cells based purely on NWP predictions (grey borders) for shorter lead times of up to two hours?

Respondent skipped this question

Q6

How useful is the information provided by cells based purely on NWP predictions (grey borders) for longer lead times?

☆ **very useful**

Q7

How useful is the information provided by the KONRAD3D-SINFONY probability product? Consider that the probability product contains both the information from the combination of NWC and NWP and from pure NWP for future time steps.

☆	**very useful**
What is difficult and easy to understand from this product?:	I found it quite easy to understand

Q8

Does KONRAD3D-SINFONY provide further beneficial information for the forecast in comparison to KONRAD3D-EPS?

☆	**Provides a lot of benefit**
Comments::	transition vom NWC to NWF in one display

#6

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Tuesday, September 19, 2023 3:50:14 PM
Last Modified: Tuesday, September 19, 2023 4:03:48 PM
Time Spent: 00:13:33
IP Address: 62.46.104.209

Page 1: Evaluation

Q1

How would you rate the quality of the combination of NWC and NWP?

☆ **Somewhat good**

Q2

Are there noticeable artifacts in the forecasts that may result from the transition from the combination of NWC and NWP to forecasts that are purely based purely on NWP?

not found on 2 August

Q3

Are there artifacts in the watch regions (cells with gray borders) associated with the fact that the cells come from an NWP ensemble? E.g., cells too big, variety of different severities overlapping at the same place? If yes, does this add to the uncertainty in the forecast?

onverlapping isn't necessarily a bad thing

Q4

Is there any additional value of the combined cells (black borders) in terms of life cycle and severity compared to KONRAD3D-EPS?

☆ **only in some cases**

What is difficult or easy to understand about this product?: if both in NWC and NWP present, this adds to the trust in the forecast

Q5

How useful is the information provided by cells based purely on NWP predictions (grey borders) for shorter lead times of up to two hours?

☆ **very useful**

What is difficult or easy to understand about this product?: It shows the potential for initiation according to the model. I like having the information displayed. Maybe it would be good to mark them better for the eye, like also using a lighter colour and not only the grey border, which is difficult to see at times

Q6

How useful is the information provided by cells based purely on NWP predictions (grey borders) for longer lead times?

☆	**very useful**

Q7

How useful is the information provided by the KONRAD3D-SINFONY probability product? Consider that the probability product contains both the information from the combination of NWC and NWP and from pure NWP for future time steps.

☆	**very useful**
What is difficult and easy to understand from this product?:	this is a good way to display the likelyhood of convective cells in a seamless way of the next hours

Q8

Does KONRAD3D-SINFONY provide further beneficial information for the forecast in comparison to KONRAD3D-EPS?

☆	**Provides a lot of benefit**

#7

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Thursday, September 21, 2023 10:16:59 AM
Last Modified:	Thursday, September 21, 2023 11:16:57 AM
Time Spent:	00:59:57
IP Address:	78.54.69.8

Page 1: Evaluation

Q1

How would you rate the quality of the combination of NWC and NWP?

☆ **Good in most cases**

Q2

Are there noticeable artifacts in the forecasts that may result from the transition from the combination of NWC and NWP to forecasts that are purely based purely on NWP?

12 July 2023, 13 UTC, S Germany, comparing a linear cluster to a supercell at the eastern edge of the cluster: The linear structure of the cluster clearly comes from the radar data, whereas the NWP product leads to a broad zone of higher probabilities, and you can see this transition. The SC is not captured well by the radar-based forecast, but ICON-D2 EPS includes it in some way. Here, a stronger signal from NWP leads to better results. Again, the transition is clearly visible with respect to the intensity, location, and longevity.

Q3

Are there artifacts in the watch regions (cells with gray borders) associated with the fact that the cells come from an NWP ensemble? E.g., cells too big, variety of different severities overlapping at the same place? If yes, does this add to the uncertainty in the forecast?

12 July 2023, 13 UTC, S Germany, comparing a linear cluster to a supercell at the eastern edge of the cluster: Within the cluster, NWP predictions show a rather broad area instead of linear segments: This adds to uncertainty, however, in this case, it is not too useful, and a more confined area would be appreciated. Additionally, there are all severity levels overlaping in this cluster, what reflects that there are some stronger cells embedded (useful information). For the SC, the cells is slightly too large.

Q4

Is there any additional value of the combined cells (black borders) in terms of life cycle and severity compared to KONRAD3D-EPS?

☆ **yes, totally**

What is difficult or easy to understand about this product?:

12 July 2023, 13 UTC, S Germany, comparing a linear cluster to a supercell at the eastern edge of the cluster: Within the cluster, radar-based predictions show decaying severity, whereas the NWP-based forecasts indicate more severe convection in the same area. This is useful information. Additionally, backbuilding is included in the NWP, what is a very good addition to CONRAD3D-EPS that does not capture it. For the SC, the benefit of using the NWP ensemble is clearly visible and in this case, NWP could be weighted much more even in the very short term since CONRAD3D-EPS has a bad handling with this cell.

Q5

How useful is the information provided by cells based purely on NWP predictions (grey borders) for shorter lead times of up to two hours?

☆ **very useful**

What is difficult or easy to understand about this product?:

You have to get used to the combination of watches and combined cells. The overlap of different severity levels is not so nice to work with.

Q6

How useful is the information provided by cells based purely on NWP predictions (grey borders) for longer lead times?

☆ **useful only in some cases**

Q7

How useful is the information provided by the KONRAD3D-SINFONY probability product? Consider that the probability product contains both the information from the combination of NWC and NWP and from pure NWP for future time steps.

☆ **useful only in some cases**

What is difficult and easy to understand from this product?:

12 July 2023, 13 UTC, S Germany, comparing a linear cluster to a supercell at the eastern edge of the cluster: The probability product seems to show the movement and severity of the SC less pronounced as I would expect. The backbuilding of the linear cluster is not well connected to the NWC watches. Here, it looks like new storms will develop after some time, whereas in reality, there was some continuous backbuilding.

Q8

Does KONRAD3D-SINFONY provide further beneficial information for the forecast in comparison to KONRAD3D-EPS?

☆

Provides a lot of benefit

Comments::

12 July 2023, 13 UTC, S Germany, comparing a linear cluster to a supercell at the eastern edge of the cluster: Clearly, NWP data can improve the forecast in this case.

#8

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Thursday, October 05, 2023 6:08:20 PM
Last Modified:	Thursday, October 05, 2023 6:14:13 PM
Time Spent:	00:05:53
IP Address:	94.142.239.106

Page 1: Evaluation

Q1

How would you rate the quality of the combination of NWC and NWP?

☆ **Somewhat good**

Q2

Are there noticeable artifacts in the forecasts that may result from the transition from the combination of NWC and NWP to forecasts that are purely based purely on NWP?

15 August 2023, S Germany, afternoon hours. Yes, the number of cells greatly increases with a combination of very small and very large cells.

Q3

Are there artifacts in the watch regions (cells with gray borders) associated with the fact that the cells come from an NWP ensemble? E.g., cells too big, variety of different severities overlapping at the same place? If yes, does this add to the uncertainty in the forecast?

Yes, there are strong jumps both in the intensity and size of the cells.

Q4

Is there any additional value of the combined cells (black borders) in terms of life cycle and severity compared to KONRAD3D-EPS?

☆ **only in some cases**

Q5

How useful is the information provided by cells based purely on NWP predictions (grey borders) for shorter lead times of up to two hours?

☆ **very useful**

What is difficult or easy to understand about this product?:

In this case this information was very important as it showed regions, where new cells may develop along the mountains or outflow boundaries.

Q6

How useful is the information provided by cells based purely on NWP predictions (grey borders) for longer lead times?

☆ **useful only in some cases**

Q7

How useful is the information provided by the KONRAD3D-SINFONY probability product? Consider that the probability product contains both the information from the combination of NWC and NWP and from pure NWP for future time steps.

☆ **useful only in some cases**

What is difficult and easy to understand from this product?: Jumpiness reduces the usefulness of this product, but in this case, SINFONY was better than the EPS and both were better for KONRAD3D for nowcasting.

Q8

Does KONRAD3D-SINFONY provide further beneficial information for the forecast in comparison to KONRAD3D-EPS?

☆ **Provides a lot of benefit**

Comments:: In this case there would be strong benefit because of the simulation of new cells in situation, when most of the cells were generally short lived.

#9

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Saturday, October 07, 2023 10:41:29 AM
Last Modified: Saturday, October 07, 2023 11:12:00 AM
Time Spent: 00:30:30
IP Address: 160.45.74.5

Page 1: Evaluation

Q1

How would you rate the quality of the combination of NWC and NWP?

☆ **Somewhat good**

Q2

Are there noticeable artifacts in the forecasts that may result from the transition from the combination of NWC and NWP to forecasts that are purely based purely on NWP?

26 August 2023, 13 UTC, S Germany: You can see the transition at a fofecast lead time of 90 minutes (14:30 UTC): Close to the Alps, the decayed NWC information that was somewhat misleading is replaced by NWP information that indicates more potential close to the Alps.

Q3

Are there artifacts in the watch regions (cells with gray borders) associated with the fact that the cells come from an NWP ensemble? E.g., cells too big, variety of different severities overlapping at the same place? If yes, does this add to the uncertainty in the forecast?

26 August 2023, 13 UTC, S Germany: Yes. It also adds to the uncertainty. Since NWC did not perform very well in this case, the NWP ensemble was a good addition. Some areas where to big, especially in the west.

Q4

Is there any additional value of the combined cells (black borders) in terms of life cycle and severity compared to KONRAD3D-EPS?

☆ **yes, totally**

What is difficult or easy to understand about this product?:

26 August 2023, 13 UTC, S Germany: the EPS did not perform well with respect to the lifetime. The combined cells where much better, e.g. between 14 and 14:30 UTC (forecast lead time 60-90 minutes).

Q5

How useful is the information provided by cells based purely on NWP predictions (grey borders) for shorter lead times of up to two hours?

☆	**very useful**
What is difficult or easy to understand about this product?:	26 August 2023, 13 UTC, S Germany: the EPS did not perform well with respect to the lifetime. The NWP was better, e.g. between 14 and 14:30 UTC (forecast lead time 60-90 minutes).

Q6

How useful is the information provided by cells based purely on NWP predictions (grey borders) for longer lead times?

☆	**useful only in some cases**

Q7

How useful is the information provided by the KONRAD3D-SINFONY probability product? Consider that the probability product contains both the information from the combination of NWC and NWP and from pure NWP for future time steps.

☆	**useful only in some cases**
What is difficult and easy to understand from this product?:	Rather easy to use.

Q8

Does KONRAD3D-SINFONY provide further beneficial information for the forecast in comparison to KONRAD3D-EPS?

☆	**Provides a lot of benefit**
Comments::	26 August 2023, 13 UTC, S Germany: Overall, it improves the forecast, although there are some misleading indications of severe development in the NWP product (severe development in the west of the area).

#10

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Saturday, October 07, 2023 11:22:04 AM
Last Modified: Saturday, October 07, 2023 11:47:35 AM
Time Spent: 00:25:31
IP Address: 160.45.74.5

Page 1: Evaluation

Q1

How would you rate the quality of the combination of NWC and NWP?

☆ **Good in most cases**

Q2

Are there noticeable artifacts in the forecasts that may result from the transition from the combination of NWC and NWP to forecasts that are purely based purely on NWP?

26 August 2023, 15 UTC, S Germany: With greater forecast lead time (16:30 UTC), there are more of these artifacts.

Q3

Are there artifacts in the watch regions (cells with gray borders) associated with the fact that the cells come from an NWP ensemble? E.g., cells too big, variety of different severities overlapping at the same place? If yes, does this add to the uncertainty in the forecast?

26 August 2023, 15 UTC, S Germany: At a forecast lead time of 40 minutes, there is a large yellow area that disappears in the next step. It does not add to the uncertainty in this case: The NWP ensemble area is too big, and there are several different severities overapping.

Q4

Is there any additional value of the combined cells (black borders) in terms of life cycle and severity compared to KONRAD3D-EPS?

☆ **only in some cases**

What is difficult or easy to understand about this product?: 26 August 2023, 15 UTC, S Germany: Severity is the same, propagation is less accurate compared to Konrad3D EPS.

Q5

How useful is the information provided by cells based purely on NWP predictions (grey borders) for shorter lead times of up to two hours?

☆ **useful only in some cases**

What is difficult or easy to understand about this product?: 26 August 2023, 15 UTC, S Germany: Looking at the movement of the BE system in southern Germany, NWP-based information performs better.

Q6

How useful is the information provided by cells based purely on NWP predictions (grey borders) for longer lead times?

☆ **useful only in some cases**

Q7

How useful is the information provided by the KONRAD3D-SINFONY probability product? Consider that the probability product contains both the information from the combination of NWC and NWP and from pure NWP for future time steps.

☆ **very useful**

What is difficult and easy to understand from this product?: 26 August 2023, 15 UTC, S Germany: Especially when looking at severe cells, the product gives a good guidance for the next 2 hours.

Q8

Does KONRAD3D-SINFONY provide further beneficial information for the forecast in comparison to KONRAD3D-EPS?

☆ **A little benefit or only in some cases**

Comments:: 26 August 2023, 15 UTC, S Germany: In this case, there are almost no differences.

#11

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Thursday, October 12, 2023 3:00:45 PM
Last Modified:	Thursday, October 12, 2023 3:45:10 PM
Time Spent:	00:44:24
IP Address:	80.109.154.58

Page 1: Evaluation

Q1

How would you rate the quality of the combination of NWC and NWP?

☆ **Bad in most cases**

Q2

Are there noticeable artifacts in the forecasts that may result from the transition from the combination of NWC and NWP to forecasts that are purely based purely on NWP?

Yes, there are noticeable artefacts that appear in places where there was no extrapolated reflectivity, but where model simulated significant reflectivity. For example base time 202308261500 and forecast for 1710 UTC.

Q3

Are there artifacts in the watch regions (cells with gray borders) associated with the fact that the cells come from an NWP ensemble? E.g., cells too big, variety of different severities overlapping at the same place? If yes, does this add to the uncertainty in the forecast?

Yes, some big polygons appear from NWP ensemble and it adds to the uncertainty of the forecast. They are also quite jumpy.

Q4

Is there any additional value of the combined cells (black borders) in terms of life cycle and severity compared to KONRAD3D-EPS?

☆ **only in some cases**

What is difficult or easy to understand about this product?: life cycle of cells is better represented but still need polishing. Also new cell development is captured better. But there a lot of artifiacts which are not realistic.

Q5

How useful is the information provided by cells based purely on NWP predictions (grey borders) for shorter lead times of up to two hours?

☆ **not useful at all**

What is difficult or easy to understand about this product?: In our case the cells coming from the NWP moved in a totally different direction compared to the reality.

Q6

How useful is the information provided by cells based purely on NWP predictions (grey borders) for longer lead times?

☆ **useful only in some cases**

Q7

How useful is the information provided by the KONRAD3D-SINFONY probability product? Consider that the probability product contains both the information from the combination of NWC and NWP and from pure NWP for future time steps.

☆ **not useful**

What is difficult and easy to understand from this product?: For a forecaster for example pure ICON RUC model data provides better information which is understandable more quickly. As an operational forecaster there is not the time to capture all features provided from the KONRAD 3D sinfony.

Q8

Does KONRAD3D-SINFONY provide further beneficial information for the forecast in comparison to KONRAD3D-EPS?

☆ **Provides a lot of benefit**

Comments:: life cycle of cells is better represented but still need polishing. Also new cell development is captured better. But there a lot of artifiacts which are not realistic.

#1

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Wednesday, June 28, 2023 2:48:50 PM
Last Modified:	Wednesday, June 28, 2023 4:06:43 PM
Time Spent:	01:17:52
IP Address:	80.109.154.58

Page 1: Evaluation

Q1

Are the storm cell tracking vectors in NowCastMIX appropriate with respect to the resulting warning polygons (for the next 60 minutes)?

☆

in most or all cases

Reasons::

Nowcastmix can effectively track the existing cells and simulate the cell motion (velocity). Overall, the size of the polygons was good. There can be some problems with the identification of merging and splitting cells.

Q2

Are there systematic differences between the tracking vectors in NowCastMIX and KONRAD3D?

NowCastMIX is typically better,

Further comments (too slow vs too fast motion, not accounting for propagation...):

KONRAD3D only detects the cells, whereas Nowcastmix calculates a tracking vector for the next 60 minutes.

Q3

Are the assessments of storm severity in NowCastMIX, as seen in the warning polygon level, appropriate?

☆

In some cases

Further comments::

In some cases the severity is underestimated when comparing it with the radar reflectivity. For exmple in the event on 21.06.2023, there was only one window (0740 - 0815) where the severity was "Extreme".

Q4

Are there systematic differences between NowCastMIX and KONRAD3D in terms of storm severity?

No or small differences,

Further comments::

In some cases Nowcastmix shows higher intensity (7:40 - 8:15 UTC), in some other cases Konrad3d shows higher intensity (7:00 - 7:10 UTC).

Q5

Are the assessments of storm attributes (hail, heavy rain, gusts) appropriate in NowCastMIX?

Are there some systematic differences to KONRAD3D?:	It is not possible to answer, because we do not know what is the exact phenomenon (gusts, hail or rain) causing the severity of the storm.

#2

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Thursday, July 06, 2023 4:18:42 PM
Last Modified: Thursday, July 06, 2023 4:59:48 PM
Time Spent: 00:41:05
IP Address: 80.109.154.58

Page 1: Evaluation

Q1

Are the storm cell tracking vectors in NowCastMIX appropriate with respect to the resulting warning polygons (for the next 60 minutes)?

☆ **in some cases**

Reasons::

For the orange cell could be appropiate, but not for the red ones.

Q2

Are there systematic differences between the tracking vectors in NowCastMIX and KONRAD3D?

About the same,

Further comments (too slow vs too fast motion, not accounting for propagation...):

For the severe storms the vector should be long enough so it covers for at least 1h of the cells lifetime.

Q3

Are the assessments of storm severity in NowCastMIX, as seen in the warning polygon level, appropriate?

☆ **In most or all cases**

Q4

Are there systematic differences between NowCastMIX and KONRAD3D in terms of storm severity?

NowCastMIX typically shows higher intensity,

Further comments::

NowCastMIX gets better identification of the storm

Q5

Are the assessments of storm attributes (hail, heavy rain, gusts) appropriate in NowCastMIX?

☆ **In some cases**

#3

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Saturday, September 09, 2023 2:28:14 PM
Last Modified:	Saturday, September 09, 2023 2:46:21 PM
Time Spent:	00:18:06
IP Address:	94.142.239.106

Page 1: Evaluation

Q1

Are the storm cell tracking vectors in NowCastMIX appropriate with respect to the resulting warning polygons (for the next 60 minutes)?

☆ **in most or all cases**

Reasons::

11 July 2023. Yes, in most cases with the exception of an accelerating convective system, where the system outran the warning polygon.

Q2

Are there systematic differences between the tracking vectors in NowCastMIX and KONRAD3D?

NowCastMIX is typically better,

Further comments (too slow vs too fast motion, not accounting for propagation...):

NowCastMIX performed better as it was able to better detect the 19 UTC incoming system from Switzerland and also better outlined the track of the storms. KONRAD3D circles were too small compared to the size of the linear segment.

Q3

Are the assessments of storm severity in NowCastMIX, as seen in the warning polygon level, appropriate?

☆ **In some cases**

Further comments::

Some underestimation was noted in the later stages of a convective system, when strong outflow continued producing severe wind gusts. Overall, it was pretty good.

Q4

Are there systematic differences between NowCastMIX and KONRAD3D in terms of storm severity?

NowCastMIX typically shows higher intensity

Q5

Are the assessments of storm attributes (hail, heavy rain, gusts) appropriate in NowCastMIX?

Are there some systematic differences to KONRAD3D?:	Impossible to assess

#4

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Monday, September 18, 2023 6:12:28 PM
Last Modified: Monday, September 18, 2023 8:17:16 PM
Time Spent: 02:04:47
IP Address: 160.45.74.5

Page 1: Evaluation

Q1

Are the storm cell tracking vectors in NowCastMIX appropriate with respect to the resulting warning polygons (for the next 60 minutes)?

☆

in some cases

Reasons::

12 July 2023 12 UTC: Cells close to the Alps are expected to move NE, what is not the case since the they stick to the Alps. Cells farther N are predicted much better (direction and speed), although back-building is missed.

Q2

Are there systematic differences between the tracking vectors in NowCastMIX and KONRAD3D?

KONRAD3D is typically better,

Further comments (too slow vs too fast motion, not accounting for propagation...):

12 July 2023, 12 UTC: KONRAD3D picks some cells better than NowCastMIX. CLose to the Alps, one cell is predicted to move NE by Nowcast-Mix, and the E-movement of KONRAD3D is better. Over central Germany, a cell is not devected by NowcastMIX but by KONRAD3D.

Q3

Are the assessments of storm severity in NowCastMIX, as seen in the warning polygon level, appropriate?

☆

In most or all cases

Q4

Are there systematic differences between NowCastMIX and KONRAD3D in terms of storm severity?

No or small differences,

Further comments::

12 July 2023, 12 UTC

Q5

Are the assessments of storm attributes (hail, heavy rain, gusts) appropriate in NowCastMIX?

☆	**In most or all cases**
Are there some systematic differences to KONRAD3D?:	12 July 2023 12 UTC. Hail of 2 cm occurred in red polygons, whereas weaker storms generally only had yellow and orange flags.

#5

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Monday, September 18, 2023 8:52:55 PM
Last Modified:	Monday, September 18, 2023 9:04:33 PM
Time Spent:	00:11:38
IP Address:	62.46.104.209

Page 1: Evaluation

Q1

Are the storm cell tracking vectors in NowCastMIX appropriate with respect to the resulting warning polygons (for the next 60 minutes)?

☆ **in most or all cases**

Reasons:: looks good for 10 July

Q2

About the same

Are there systematic differences between the tracking vectors in NowCastMIX and KONRAD3D?

Q3

Are the assessments of storm severity in NowCastMIX, as seen in the warning polygon level, appropriate?

☆ **In most or all cases**

Further comments:: 10 July

Q4

No or small differences,

Are there systematic differences between NowCastMIX and KONRAD3D in terms of storm severity?

Further comments::

KONRAD3D seem to react a bit faster to changes in intensity than NowCastMIX on 10 July

Q5

Are the assessments of storm attributes (hail, heavy rain, gusts) appropriate in NowCastMIX?

Are there some systematic differences to KONRAD3D?: were not displayed for 10 July

#6

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Tuesday, September 19, 2023 4:04:18 PM
Last Modified: Tuesday, September 19, 2023 4:16:13 PM
Time Spent: 00:11:54
IP Address: 62.46.104.209

Page 1: Evaluation

Q1

Are the storm cell tracking vectors in NowCastMIX appropriate with respect to the resulting warning polygons (for the next 60 minutes)?

☆ **in most or all cases**

Reasons:: for the 2 tornado cases of 2 August NowCastMIX has performed well

Q2

Are there systematic differences between the tracking vectors in NowCastMIX and KONRAD3D?

About the same

Q3

Are the assessments of storm severity in NowCastMIX, as seen in the warning polygon level, appropriate?

☆ **In most or all cases**

Further comments:: for the cases of 2 August very good

Q4

Are there systematic differences between NowCastMIX and KONRAD3D in terms of storm severity?

NowCastMIX typically shows higher intensity,

Further comments:: for the cases of 2 August

Q5

Are the assessments of storm attributes (hail, heavy rain, gusts) appropriate in NowCastMIX?

Respondent skipped this question

#7

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Thursday, October 05, 2023 6:14:22 PM
Last Modified: Thursday, October 05, 2023 6:26:44 PM
Time Spent: 00:12:21
IP Address: 94.142.239.106

Page 1: Evaluation

Q1

Are the storm cell tracking vectors in NowCastMIX appropriate with respect to the resulting warning polygons (for the next 60 minutes)?

☆ **in some cases**

Reasons::

15 August 2023, S Germany, afternoon hours. Problems with decaying cells and new development along the outflow boundaries and orography.

Q2

Are there systematic differences between the tracking vectors in NowCastMIX and KONRAD3D?

NowCastMIX is typically better,

Further comments (too slow vs too fast motion, not accounting for propagation...):

Because covered larger and more smoothed areas.

Q3

Are the assessments of storm severity in NowCastMIX, as seen in the warning polygon level, appropriate?

☆ **In some cases**

Further comments::

Almost all severe weather reports either in red or violet polygons. That said, purple polygons were awarded too often.

Q4

Are there systematic differences between NowCastMIX and KONRAD3D in terms of storm severity?

KONRAD3D typically shows higher intensity

Q5

Are the assessments of storm attributes (hail, heavy rain, gusts) appropriate in NowCastMIX?

Respondent skipped this question

#8

COMPLETE

Collector: Web Link 1 (Web Link)
Started: Saturday, October 07, 2023 11:51:25 AM
Last Modified: Saturday, October 07, 2023 12:00:52 PM
Time Spent: 00:09:27
IP Address: 160.45.74.5

Page 1: Evaluation

Q1

Are the storm cell tracking vectors in NowCastMIX appropriate with respect to the resulting warning polygons (for the next 60 minutes)?

☆

in most or all cases

Reasons::

26 August 2023, 13 UTC, S Germany: Pretty accurate prediction even for the southern cell.

Q2

Are there systematic differences between the tracking vectors in NowCastMIX and KONRAD3D?

About the same,

Further comments (too slow vs too fast motion, not accounting for propagation...):

26 August 2023, 13 UTC, S Germany: NowCastMIX is better for the southern cell, that turned right. Otherwise, similiar performance.

Q3

Are the assessments of storm severity in NowCastMIX, as seen in the warning polygon level, appropriate?

☆

In most or all cases

Further comments::

26 August 2023, 13 UTC, S Germany: Very good for the southern cell (intensification).

Q4

Are there systematic differences between NowCastMIX and KONRAD3D in terms of storm severity?

No or small differences,

Further comments::

26 August 2023, 13 UTC, S Germany: Better compared to Konrad3D for the southern cell (intensification much better indicated).

Q5

Are the assessments of storm attributes (hail, heavy rain, gusts) appropriate in NowCastMIX?

☆	**In some cases**
Are there some systematic differences to KONRAD3D?:	26 August 2023, 13 UTC, S Germany: Severe hail was more intense (extreme) compared to NowCastMIX (strong).

#9

COMPLETE

Collector:	Web Link 1 (Web Link)
Started:	Saturday, October 07, 2023 12:02:26 PM
Last Modified:	Saturday, October 07, 2023 12:05:59 PM
Time Spent:	00:03:32
IP Address:	160.45.74.5

Page 1: Evaluation

Q1

Are the storm cell tracking vectors in NowCastMIX appropriate with respect to the resulting warning polygons (for the next 60 minutes)?

☆ **in most or all cases**

Reasons::

26 August 2023, 15 UTC, S Germany: Pretty accurate prediction.

Q2

Are there systematic differences between the tracking vectors in NowCastMIX and KONRAD3D?

About the same,

Further comments (too slow vs too fast motion, not accounting for propagation...):

26 August 2023, 15 UTC, S Germany: Indeed, NowCastMIX was slightly better with respect to the movement speed of the bow echo (it was faster compared to Konrad3D predictions).

Q3

Are the assessments of storm severity in NowCastMIX, as seen in the warning polygon level, appropriate?

☆ **In most or all cases**

Further comments::

26 August 2023, 15 UTC, S Germany: Yes, in this case, it performed well.

Q4

Are there systematic differences between NowCastMIX and KONRAD3D in terms of storm severity?

No or small differences

Q5

Are the assessments of storm attributes (hail, heavy rain, gusts) appropriate in NowCastMIX?

☆ **In some cases**