Application and Verification of ECMWF Products 2019

Karl-Ivar Ivarsson, Swedish meteorological and hydrological institute (SMHI)

1. Summary of major highlights

The forecasts from ECMWF have a high quality, especially regarding near surface variables related to moisture such as 2-metre dew point temperature.

2. Use and application of products

2.1 Direct Use of ECMWF Products

The forecasts are used directly (DMO data) by duty forecasters for determining the risk of severe weather, such as strong winds, including wind gusts, extreme temperatures and large precipitation amounts.

2.2 Other uses of ECMWF output

The use is similar to previous years and is summarized below:

2.2.1 Post-processing

A Kalman filter is used for adjusting 2m- temperature and 10m-wind. The ensemble mean (computed locally at the institute) is used for products such forecast chart etc.

2.2.2 Derived fields

A smoothing technique is used for all meteorological model outputs of cloud cover and precipitation, including ECMWF deterministic forecasts. The grid-point information from an area of 20 km radius is used to provide a mean value, a median value, a 90% percentile value and a 10% percentile value. Those values are calculated for all grid-points in an area covering north west Europe, basically that same area as AROME, but with a slightly different grid. It is a rotated lat-lon grid of 0.025 degrees (2.75 km). The smoothing technique is not applied for ECMWF ensemble mean.

2.2.3 Modelling

Visibility is calculated by using an algorithm based on relative humidity, precipitation and latitude.

ECMWF provides model data for lateral conditions and other input data such as 'large scale mixing', (LSM) and blending. Thus, the larger scale structures of the analysis or short forecasts are used as input for the first guess field, but the finer ones are retrieved from the first guess of the high resolution limited area models. This technique is used for AROME.

ECMWF is also used for longer (up to ten days) oceanographical forecasts. (the NEMO model) Here, ECMWF meteorological input is used as upper boundary condition.

MEPS is an ensemble system based on ten members with AROME physics at 2.5 km resolution. It is a collaboration between Sweden, Norway, Finland and Estonia. The perturbations are based on the difference between older and newer lateral boundaries from ECMWF. (Often refereed to as SLAF, Scaled Lagged Average Forecasting.) The model domain covers north-western Europe. All four countries use the MEPS forecasts. An ensemble system based on lateral boundaries from IFS-ENS (CMEPS) is pre-operational since the beginning of this year. It is planned to replace MEPS during the autumn.

3. Verification of ECMWF products

HRES, ENS, monthly and seasonal forecasts are all within scope. ECMWF does extensive verification of its products in the free atmosphere. However, verification of surface parameters is in general limited to using synoptic observations. More detailed verification of these weather parameters by national Services is particularly valuable.

3.1 Objective verification

3.1.1 Direct ECMWF model output (both HRES and ENS), and other NWP models

The general performance of ECMWF and AROME (deterministic runs) for some commonly used near surface parameters is seen in table 1:

Table 1:

Verification results for different models and seasons: AROME is the control run in the MEPS ensemble system. '10M wind' is 10 metre wind speed, 't2m' is 2 metre temperature and 'td2m' is 2 metre dew point temperature. The area for verification is north-western Europe and the forecast length ranges from 3 hours up to 48 hours.

Summer (June – August 2018)

parameter	Systematic error or bias		Mean absolute error	
model	ECMWF	AROME	ECMWF	AROME
10m wind	0.27	0.17	1.35	1.26
t2m	-0.20	-0.11	1.39	1.26
td2m	-0.13	-0.19	1.26	1.28

Autumn (September – November 2018)

parameter	Systematic error or bias		Mean absolute error	
model	ECMWF	AROME	ECMWF	AROME
10m wind	0.40	0.29	1.55	1.37
t2m	-0.02	-0.02	1.26	1.04
td2m	-0.08	-0.19	1.07	1.04

winter (December 2018– February 2019)

parameter	Systematic error or bias		Mean absolute error	
model	ECMWF	AROME	ECMWF	AROME
10m wind	0.34	0.20	1.56	1.38
t2m	0.00	-0.11	1.61	1.46
td2m	-0.08	-0.35	1.60	1.54

spring(March – May 2019)

parameter	Systematic error or bias		Mean absolute error	
model	ECMWF	AROME	ECMWF	AROME
10m wind	0.04	0.13	1.44	1.31
t2m	-0.67	-0.46	1.53	1.35
td2m	0.01	0.37	1.41	1.54

AROME has generally a lower absolute error for the near surface variables, mainly due to higher horizontal resolution.

ECMWF is still somewhat too cold during spring. ECMWF does not capture the lowest t2m below -30 C very well. Also AROME has a this problem, but it is less apparent. The frequency bias (FB) in winter (December to February) for t2m below -30 was near zero for ECMWF and 0.3 for AROME. In summer, ECMWF under-predicts high temperatures. For example, in June this year the FB was 0.5 for t2m above 30C. AROME had the opposite systematic error, a FB of around 1.4.

AROME is too dry (negative 2 metre dew point bias) in winter, but the opposite is seen in early spring. ECMWF forecasts of 2 metre dew point have small systematic errors during the whole year and thus a valuable guidance for 2m forecast variables regarding moisture. (This means also the 2m relative humidity etc.)

The ECMWF forecasts of low clouds are generally somewhat better than those from AROME during spring, probably related too an excess of surface latent heat flux in AROME. Thus, AROME over-predicts low clouds and fog especially during April. Also here, ECMWF becomes a valuable guidance, since the model has less- or no such systematic errors. But in case of moderate convection, ECMWF tends to decrease the amount of low clouds too much in the evening. (During April to August)

The precipitation forecasts from ECMWF have a high quality, but one weak point is that the summertime convective precipitation over land areas starts too early and that the diurnal cycle is over-amplified.

24 hour precipitation is regularly verified against a dense network of climate stations, mainly over Sweden and some parts of northern Norway. Only the short time forecasts (the 24 hour period starting at six- and ending at 30 hour forecast length) are verified. Fractions skill score (FBSS) is used with 'sample climate' as reference forecast. The period for verification is July 2018 to April 2019. The result is seen in in figure 1.

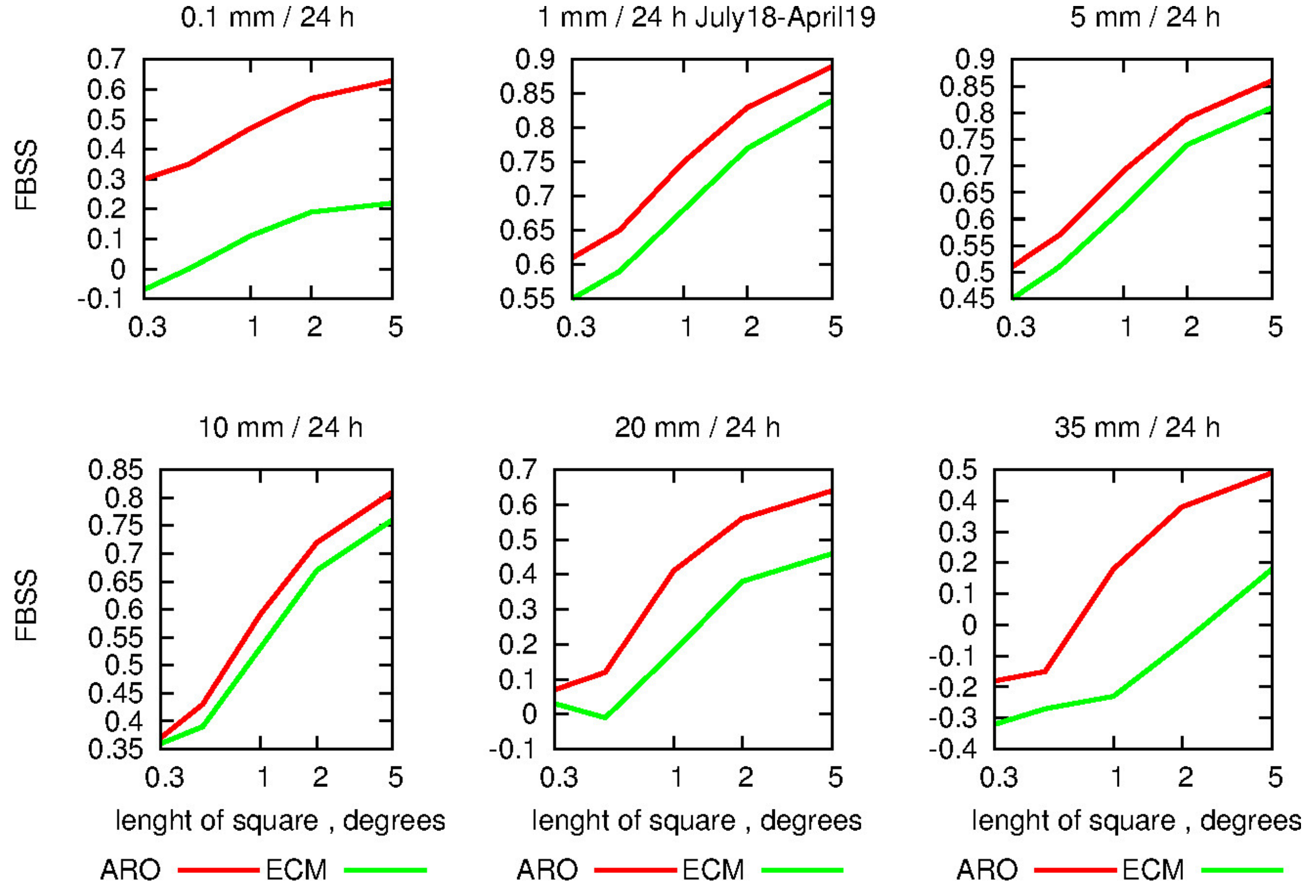

Figure 1: FBSS for different thresholds of precipitation. FBSS at the vertical axis and the size of different squares in degrees. There should be at least three observations in a square for being used in the verification. One degree is about 111 km. ECM (green) is ECMWF forecasts and ARO (red) is AROME.

AROME has the highest FBSS for all scales and all thresholds, but the result may be a coincidence for the highest thresholds, since there are few cases with large precipitation. The low skill for ECMWF for 0.1mm threshold may partly be caused by that interpolated ECMWF fields are used.

The forecasts of **maximum wind gusts** was of particular interest last winter due to some severe storms. ECMWF forecasts had a positive bias during last winter period (December – February) 1.70 m/s, compared to 0.54 for AROME. (Mean absolute errors where 2.39 and 1.78 respectively) The main reason for the general over-prediction was too excessively forecasts of 10-13 m/s. But the skill scores where quite good regarding higher thresholds, figure 2.

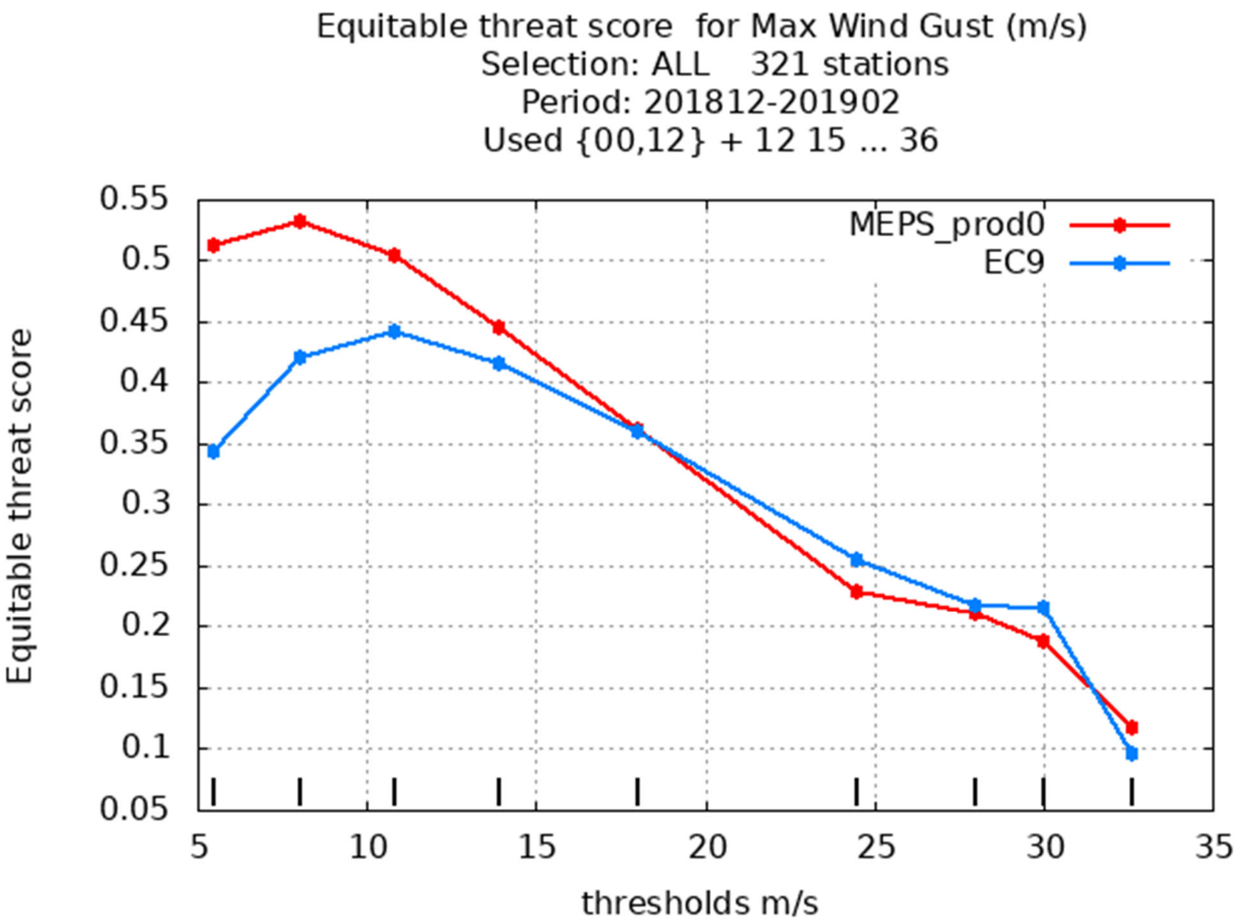

Figure 2: Equitable treat score (ETS) for one hourly maximum of wind gust during 20181201 – 20190228 for different thresholds. Red is the MEPS control run (AROME), blue is ECMWF. ETS at vertical axis, the thresholds (marked as vertical bars) at the horizontal axis.

A verification system for evaluation of the ensemble system MEPS is under development. Although not finished, some results may be mentioned. Most scores such as Brier skill score (BSS), continues rank-probability score (CRPS) etc. are better with MEPS than with IFS-ENS for the shortest lead times. For forecast lengths larger than about 36 hours, the result is mixed. Regarding 12 hour precipitation in the summertime, there is also a clear difference between precipitation during daytime and night time, figure 3.

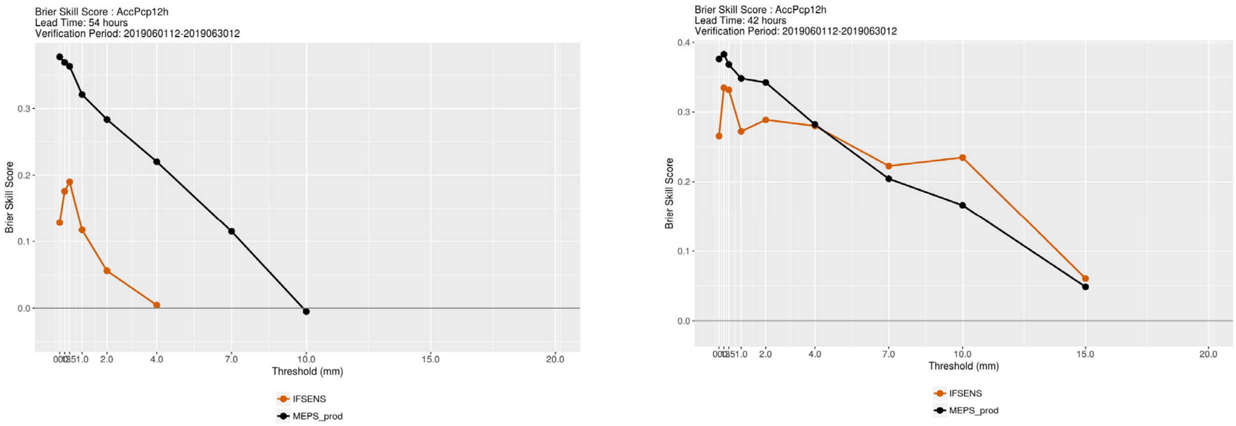

Figure 3: To the left: BSS for 12-hour precipitation June 2019. Forecast length 54 hours (42-54 hour period) and forecasts issued at 12 UTC, so the forecasts are valid at daytime, 06-18 UTC. The area is north-western Europe. Black is MEPS and light red is IFS-ENS. BSS at vertical axis, the thresholds at the horizontal axis. Negative values are disregarded. To the right: As in the left figure, but the forecast length is 42 hours (30-42 hour period) so the forecasts are valid at night-time, 18-06 UTC.

Apparently, MEPS performs better at daytime, especially because of better reliability, not shown. At night-time IFS-ENS is better than MEPS for higher precipitation thresholds.

3.1.2 Post-processed products and end products delivered to users

3.2 Subjective verification

3.2.1 Subjective scores (including evaluation of confidence indices when available)

3.2.2 Case studies

Severe weather events/non-events are of particular interest. Include an evaluation of the behaviour of the model(s). Reference to major forecast errors, even if they are not in a “severe weather” category, are also very welcome.

4. Requests for additional output

Low-,middle- and high clouds closer to WHO definitions: 0-2000m, 2000-5000,m and >5000m above ground respectively.

5. Feedback on ECMWF “forecast user” initiatives

We invite comments on how useful you find the information provided on ECMWF’s “Forecast User Portal”, see: (https://software.ecmwf.int/wiki/display/FCST/Forecast+User+Home), and on any changes you would like to see. The web-based “Forecast User Guide” was introduced in May 2018 (https://confluence.ecmwf.int/display/FUG/Forecast+User+Guide) and we would particularly welcome feedback on that.

6. References to relevant publications

(7. Structure of these Reports)

A deadline for the report in the middle of the summer (July 31) is not optimal. Most staff are on vacation, which makes it harder to collect information about the model performance and to receive help with retrieving some relevant verification results. One possibility is then to do this work earlier, but that means that the verification results will be older than necessary. A suggestion is to have a deadline in late winter (E.g. in the end of February) and let the report reflect the previous year.

Sweden