Verification of ECMWF products at the Deutscher Wetterdienst (DWD)

DWD – Dr. Marcus Paulat, Dr. Ulrich Damrath, Gabriele Schweigert

1 Summary of major highlights

The usage of a combined GME-MOS and ECMWF-MOS at DWD continues to form the basis for the production of local short and medium range forecasts. It has been augmented in the short range by forecasts from the regional model COSMO-EU in the best available guidance called Objectively Optimised Guidance (OOG). ECMWF high resolution forecasts in conjunction with GME forecasts are also being used for the production of a probabilistic warning guidance based on the MOS technology. Furthermore, a growing number of ECMWF Ensemble data are adapted to the Ensemble Layer within DWD's visualisation software NinJo.
The high resolution ECMWF model is one of four driving models for the high resolution COSMO-DE-EPS which is operational since May 2012 and accepted in forecasting deep convection fairly good compared to available deterministic information.

2 Use and application of products

2.1 Post-processing of model output

2.1.1 Statistical adaptation

The high resolution ECMWF model (both 12 and 00 UTC run) and DWD's model GME are statistically interpreted up to 7 days in terms of near surface weather elements by means of a perfect prog scheme (AFREG) as well as by MOS and subsequent weighted averaging of the two interpretations to form „AFREG/MIX" and "MOS/MIX". Because of the upcoming change from GRIB1 to GRIB 2 format and the good skills from MOS technique AFREG/MIX will be replaced by "MOSGEB" (forecasts produced for certain areas) in the end of 2013.
Since 2008 ECMWF high resolution forecasts in conjunction with GME forecasts have been used for the production of a probabilistic warning guidance based on the MOS technology.

2.1.2 Physical adaptation

2.1.3 Derived fields

2.2 Use of products

The high resolution ECMWF model forms together with DWD's model GME the general operational data base. ECMWF's high resolution model is always used together with other models in short- and medium-range forecasting. For medium range forecasting the EPS is used additionally and products are presented in the weekly weather discussion; in the short range COSMO-LEPS (Local model nested into EPS clusters) provides ensemble information. EPS products are used intensively in order to create a daily simple confidence number and describe alternative solutions. Furthermore, they are used to estimate the prospect for extreme weather events. Here, extensive use of the Extreme Forecast Index (EFI) is made. There is growing usage of the new products as presented on the ECWMF website. To make some of these products more easily usable in the context of DWD environment (layer technique for comparison to other meteorological data), ECMWF-EPS, LEPS and COSMO-DE-EPS products are displayed within NinJo with growing number of available products. As an example, figure 1 illustrates the visualisation of ECMWF ensemble data in meteogram format for a certain time and location (available at each model grid point). The data can be viewed in 6 hourly time steps up to 240 hours for (top to bottom) total cloud cover, precipitation, CAPE, wind gusts and 2m temperature (Tmax in red, Tmin in blue). We also added deterministic information from ECMWF's high resolution model as symbols (total cloud cover), columns (precipitation) and lines (other parameters). The very helpful visualisation of wind direction shown in ECMWF ensemble meteograms is planned. Note, that for winter season CAPE is exchanged by snowfall.

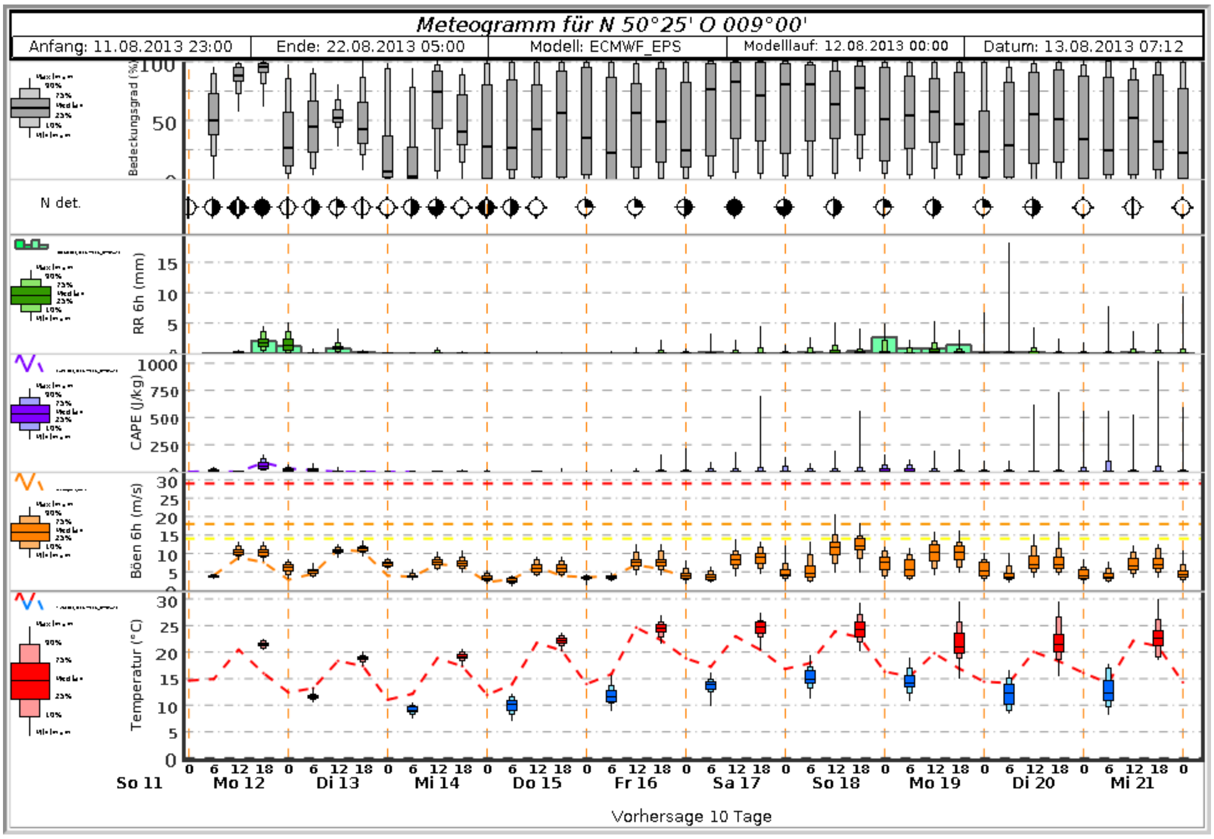


Fig. 1: Ensemble meteogram from NinJo for a certain ensemble run and grid point in summer season. ECMWF-EPS data are visualised in 6 hourly time steps up to 240 hours. DWD also added deterministic model information.

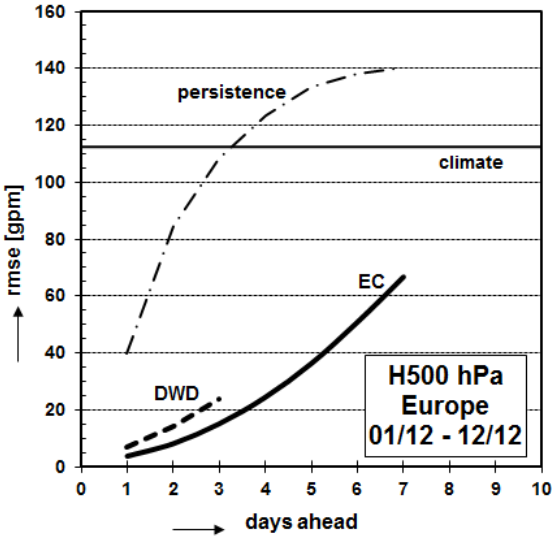
3 Verification of products

3.1. Objective verification

3.1.1 Direct ECMWF model output (both deterministic and EPS)

3.1.2 ECMWF model output compared to other NWP models

Again in 2012, upper air forecasts from ECMWF continued to exhibit smaller errors than DWD-GME forecasts (Fig. 2). The RMSE of the ECMWF model for 500hPa geopotential height has not significantly improved in the short range from 2009 to 2012. Since we could find an improvement in the medium range from 2009 to 2010, there is no further improvement compared now. ECMWF MSLP error growth with forecast range remains about one day better than for DWD-GME in the short range (fig. 3).



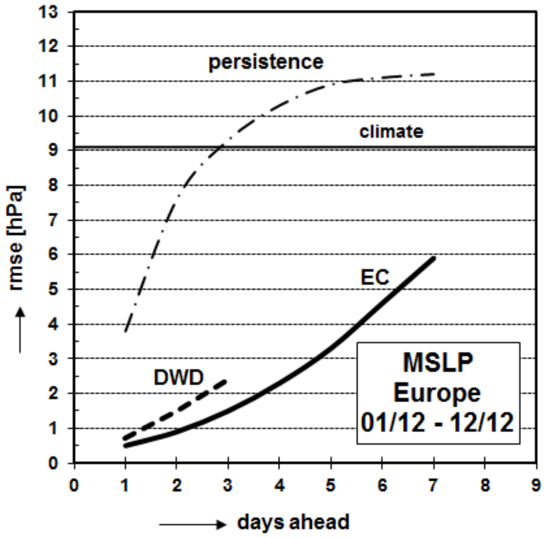


Fig. 2: RMSE 500hPa geopotential over Europe. DWD (Numerical Weather Prediction model GME), EC (high resolution ECMWF model), persistence (analysis from the initial state is used as a forecast for all following days), climate (long term mean of the predictand (H500, MSLP) serves as a constant forecast.

Fig. 3: Same as fig. 2, but for RMSE of mean sea level pressure.

3.1.3 Post-processed products

Here, various statistically post-processed model forecasts are compared for the following:

Predictands

MIN = daily minimum temperature (°C)
MAX = daily maximum temperature (°C)
SD = daily relative sunshine duration (%)
dd = surface wind direction (°) 12 UTC. Only verified, if ff(obs) ≥ 3 m/s
ff = surface wind speed (m/s) 12 UTC
PoP = Probability of Precipitation > 0 mm/d
PET = potential evapotranspiration (mm/d)
RR = a binary predictand: precipitation amount > 0 mm/d: Yes/No;

Forecast Types

AFREG/MIX = Perfect prog product AFREG(MIX) = AFREG(EC)+AFREG(DWD)/2
EC = high res. ECMWF model, DWD = operational DWD Global Model "GME" (initial time: 00 UTC). AFREG is generated for several *areas* of whole Germany, but verified against *point* observations at 6 stations.

MOS/MIX = post processed product, a weighted average of Model Output Statistics of MOS/GME and MOS/EC

Verification measures

rmse is used for both categorical and probabilistic forecasts (equals square root of the Brier Score)
RV = Reduction of Variance against reference, $1-(rmse/rmse^*)^2$, here: mean value for day 2-7
$rmse^*$ = smoothed climate as the best reference forecast to evaluate forecast skill
HSS = Heidke Skill Score, only for binary predictands
$\overline{\textbf{HSS}}$ = mean value for day 2-7

rmse		day							rmse*	
		+2	+3	+4	+5	+6	+7	+8	(climate)	RV [%]
MIN	AFREG/MIX	2,34	2,42	2,63	2,83	3,04	3,31	3,72	4,45	64
	MOS/MIX	1,50	1,76	2,09	2,50	2,88				76
MAX	AFREG/MIX	2,45	2,57	2,83	3,18	3,49	3,85	4,19	4,76	63
	MOS/MIX	1,68	1,97	2,37	2,88	3,32				72
SD	AFREG/MIX	26,0	26,6	27,5	28,2	28,7	29,6	30,3	30,7	18
dd[1)]	AFREG/MIX	43,9	47,8	50,9	54,5	61,5	71,0	74,4	88,5	66
	MOS/MIX	31,6	39,4	45,9	51,6	61,7				72
ff	AFREG/MIX	1,73	1,79	1,88	1,96	2,03	2,13	2,16	2,12	21
	MOS/MIX	1,44	1,60	1,74	1,86	1,93				34
PoP	AFREG/MIX	38,6	39,6	40,5	41,5	43,0	44,4	45,2	46,1	25
	MOS/MIX	34,0	36,5	39,0	40,6					33
PET	AFREG/MIX	0,722	0,745	0,764	0,794	0,822	0,865	0,881	0,889	22
HSS%										$\overline{\text{HSS}}$
RR	AFREG/MIX	45	39	41	35	30	22	22		40
	MOS/MIX	63	54	44	39					50

Table 1: Verification of operational medium range forecasts for 6 stations in Germany (Hamburg, Potsdam, Düsseldorf, Leipzig, Frankfurt/M., München); 01/2012 - 12/2012; rmse and HSS, respectively. Day of issue = day +0 = today at noon. [1)] Here, persistence is used as a 'reference forecast'.

While the skill (RV) of minimum temperature was 6 % higher in 2012 than in 2011, wind speed forecasts reached a significant worse performance (9 % lower in 2012 than in 2011). Also the Probability of Precipitation (3 %) and RR (2%) slightly lost skill.
MOS/MIX forecasts have substantially smaller errors than AFREG/MIX, which is only partly due to the lower (and thus less realistic) variability of MOS forecasts. The lower variability of MOS, especially in the medium range, is an obstacle for the use of it for forecasts of more severe weather. Here, the more variable solutions of the ECMWF-EPS serve as an important additional guidance.
The application of post-processing lead to largely reliable probability of YES/NO precipitation (PoP) forecasts (fig. 4) showing one big exception at forecasted probabilities of 0% (where precipitation occurs in 50% of all cases). This could be seen in 2011, too, but the 'effect' is much more obvious in 2012. That could be due to changes in the MOS system which might lead to high frequencies of drizzle rain events. This task is under investigation.
Here, AFREG/MIX is nearly perfect and performs better than MOS/MIX.

Figs. 5-6a, b show two things:

i) the MOS technology performs better than a perfect prog technology (AFREG);
ii) mixing post-processed products from both models lead to a very moderate improvement of the forecast. However, in the medium range the gain in skill due to mixing is about half a day.

Figure 6a shows very small differences in quality comparing results in 2012 to that of 2011.

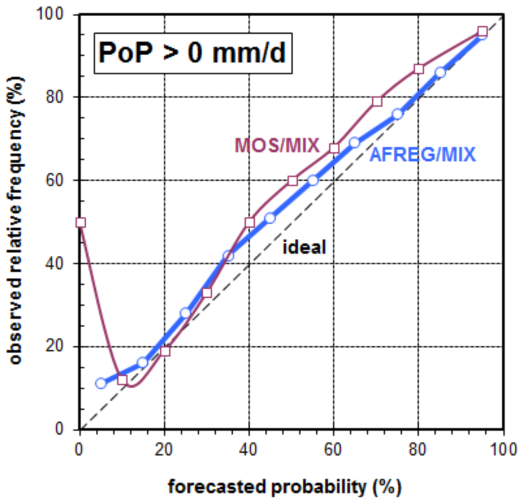


Fig. 4: Reliability diagram (6 stations, 01/12 – 12/12, day+2 ... day+7; only up to day+5 for MOS/MIX))

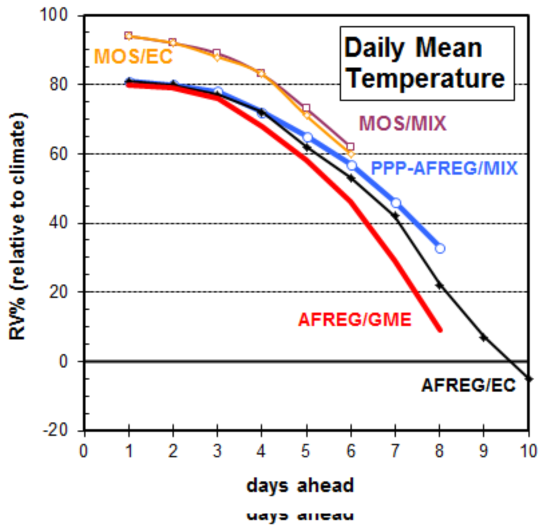


Fig. 5: Forecast skill RV for Daily Mean Temperature (DWD, 6 stations, 01/12 – 12/12

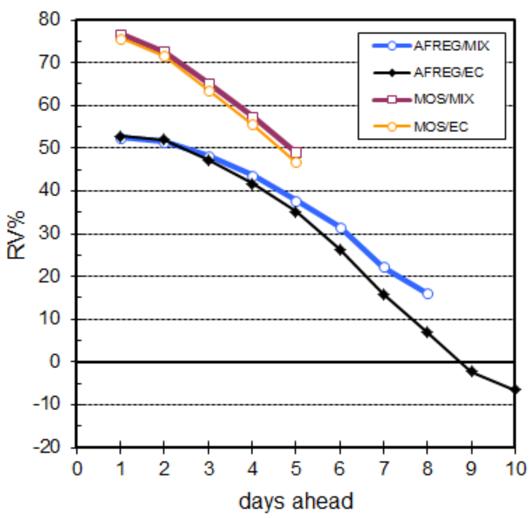


Fig. 6a: Forecast skill RV as a function of range, averaged for all predictands taken in table 1 (without PET and RR)

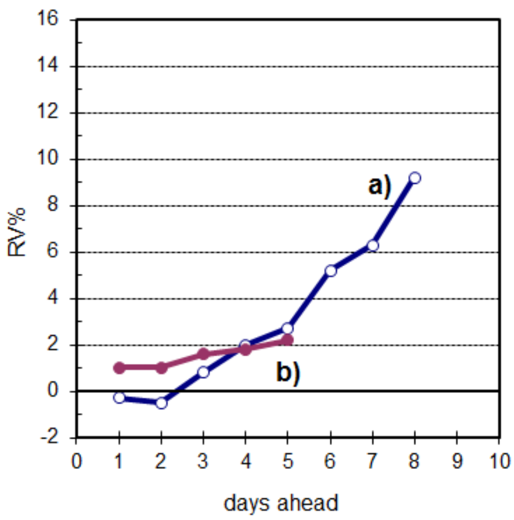


Fig. 6b: follows from fig. 5a:
a) Blue line:
RV(AFREG/MIX) - RV(AFREG/EC)
b) Claret red line:
RV(MOS/MIX) - RV(MOS/EC)