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 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, ECMWF Ensemble data are now adapted to the Ensemble Layer within DWD's visualisation software NinJo.

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 AFREG/MIX will be replaced by "MOSGEB" in 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. Furthermore, the ECMWF model is one of four driving models for the high resolution COSMO-DE-EPS which is operational since May 2012.

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 the 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 now displayed within NinJo.

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

In 2011, upper air forecasts from ECMWF continued to exhibit smaller errors than DWD-GME forecasts (Fig. 1). The RMSE of the ECMWF model for 500hPa geopotential height has not improved in the short range from 2009 to 2011. Since we could find an improvement in the medium range from 2009 to 2010, there is no further improvement compared to 2011 now. ECMWF MSLP error growth with forecast range remains about one day better than for DWD-GME in the short range (fig. 2).

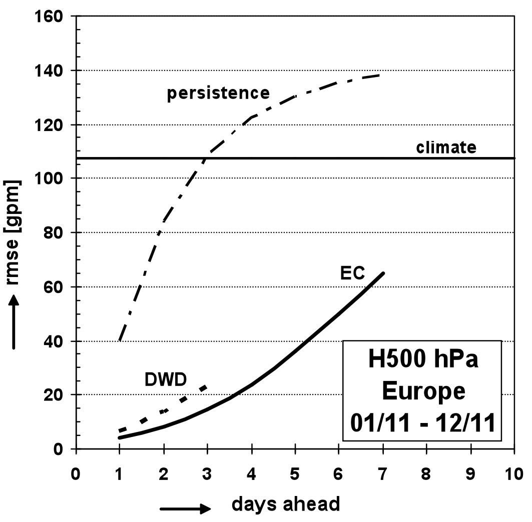

Fig. 1: RMSE 500hPa geopotential over Europe. DWD (Numerical Weather Prediction model GME), EC (high resolution ECMWF model), persistence (analysis from the initial state is used as a forecast for all following days), climate (long term mean of the predictand (H500, MSLP) serves as a constant forecast.

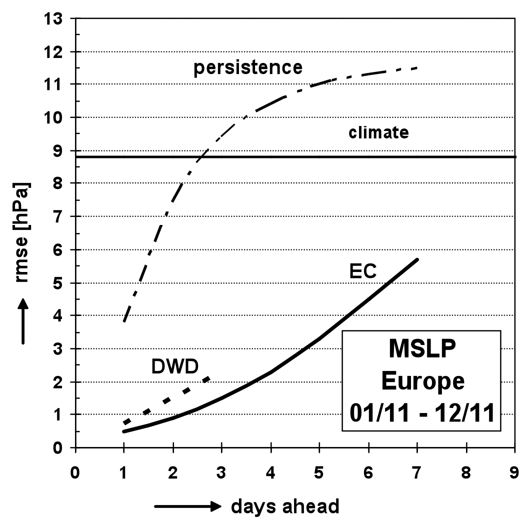

Fig. 2: Same as fig. 1, but for RMSE of mean sea level pressure.

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 and **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	2,44	2,53	2,62	2,76	3,01	3,26	3,50	4,07	57
	MOS/MIX	1,63	1,87	2,14	2,49	2,87				70
MAX	AFREG/MIX	2,53	2,61	2,80	3,07	3,41	3,77	3,97	4,77	63
	MOS/MIX	1,79	2,05	2,39	2,85	3,25				72
SD	AFREG/MIX	26,7	26,8	27,4	28,6	29,5	30,3	31,4	32,9	26
dd[1)]	AFREG/MIX	42,1	45,1	49,8	54,9	62,5	69,0	75,6	86,1	65
	MOS/MIX	31,8	37,1	45,5	52,3	59,7				72
ff	AFREG/MIX	1,63	1,69	1,85	1,93	2,07	2,15	2,25	2,26	34
	MOS/MIX	1,36	1,53	1,73	1,87	1,97				43
PoP	AFREG/MIX	39,5	39,5	41,0	42,4	44,7	46,2	48,4	48,7	30
	MOS/MIX	35,9	38,1	39,6	42,7					36
PET	AFREG/MIX	0,764	0,750	0,774	0,807	0,842	0,860	0,890	0,956	30
HSS%										$\overline{\text{HSS}}$
RR	AFREG/MIX	52	53	48	40	32	29	20		48
	MOS/MIX	62	56	51	40					52

Table 1: Verification of operational medium range forecasts for 6 stations in Germany (Hamburg, Potsdam, Düsseldorf, Leipzig, Frankfurt/M., München); 01/2011 - 12/2011; 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 and wind direction forecasts was slightly lower in 2011 than in 2010, all other parameters reached a more or less significant better performance.
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 EPS serve as an important additional guidance.
The application of post-processing lead to largely reliable probability of YES/NO precipitation (PoP) forecasts (fig. 3). Here, AFREG/MIX is nearly perfect and performs better than MOS/MIX. Compared to 2010, MOS/MIX shows an underforecasting of probabilities around 0 % now. Figs. 4-5a, 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 5a shows differences in quality comparing results in 2011 to that of 2010. As concluded for the RV's in table 1 the RV's dependent on the lead time are higher (2-10%) than in 2010.

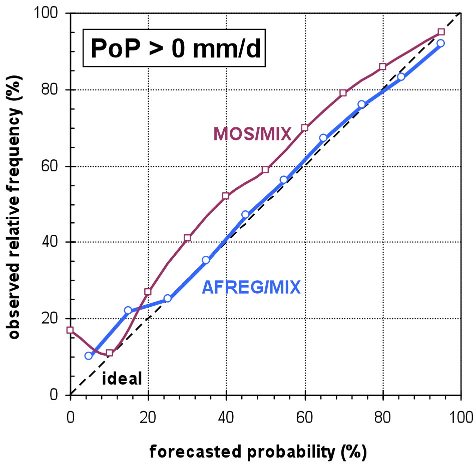

Fig. 3: Reliability diagram (6 stations, 01/11 – 12/11, day+2 ... day+7; only up to day+5 for MOS/MIX))

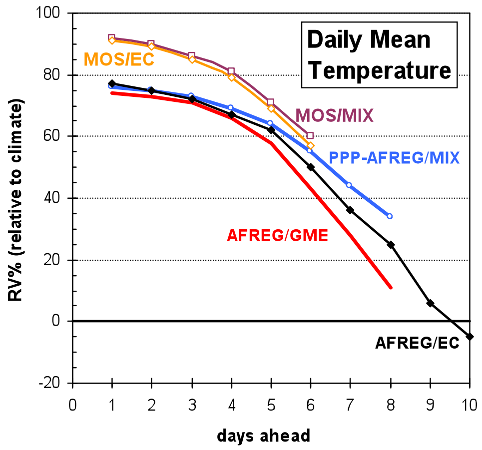

Fig. 4: Forecast skill RV for Daily Mean Temperature (DWD, 6 stations, 01/11 – 12/11

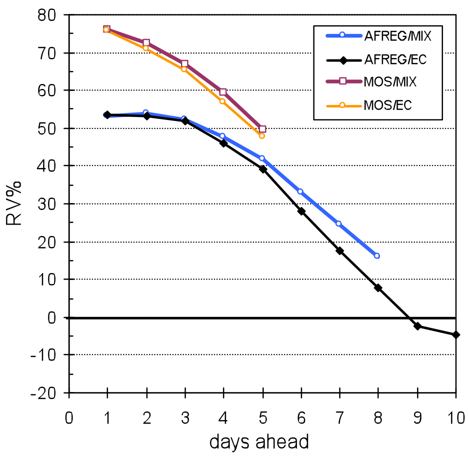

Fig. 5a: Forecast skill RV as a function of range, averaged for all predictands taken in table 1 (without PET and RR)

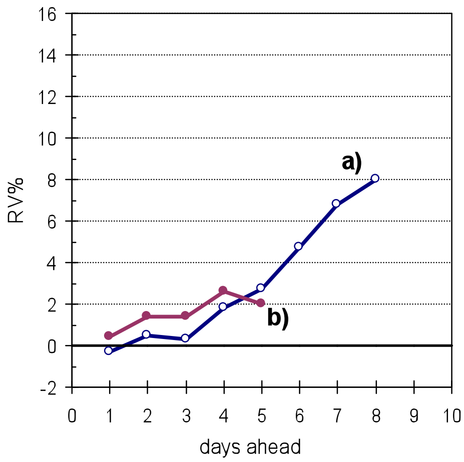

Fig. 5b: follows from fig. 5a:

a) Blue line:
RV(AFREG/MIX) - RV(AFREG/EC)

b) Claret red line:
RV(MOS/MIX) - RV(MOS/EC)