Application and verification of ECMWF products 2016

Federal Office of Meteorology and Climatology MeteoSwiss

Christoph Spirig, Jonas Bhend, Samuel Monhart, Irina Mahlstein, Mark Liniger and Pirmin Kaufmann

1. Summary of major highlights

2. Use and application of products

2.1 Post-processing of ECMWF model output

2.1.1 Statistical adaptation

Extended and long-range forecasts

In the framework of a national research program on renewable energies, MeteoSwiss investigates the use of extended range ensemble forecasts for hydrological modelling to be employed for management of hydropower reservoirs. Daily model output of the extended range ensembles has been processed using a mean de-biasing technique (Mahlstein et al., 2015) and a quantile mapping approach. Hindcasts of the ECMWF cycle 40r1 were bias-corrected using surface observations of about 1'000 sites of the ECA&D data set (www.ecad.eu). Figure 1 shows that forecasts benefit from both mean bias-correction techniques, whereas quantile mapping resulted in somewhat better results as measured by the continuous ranked probability skill score. Skilful temperature forecasts (verified against climatological forecasts as a reference) up to week 3 could be obtained, with forecasts initialized in winter and fall performing better than those in summer and spring. For precipitation, positive skill score values were limited to lead times of two weeks, at most.

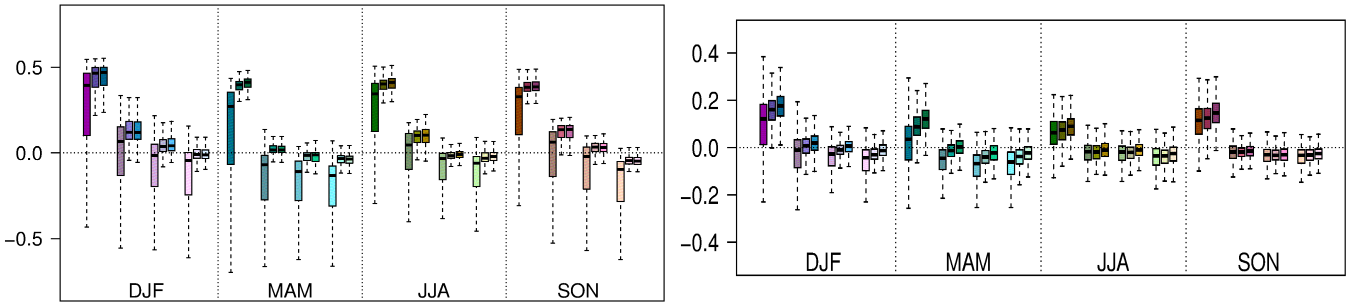
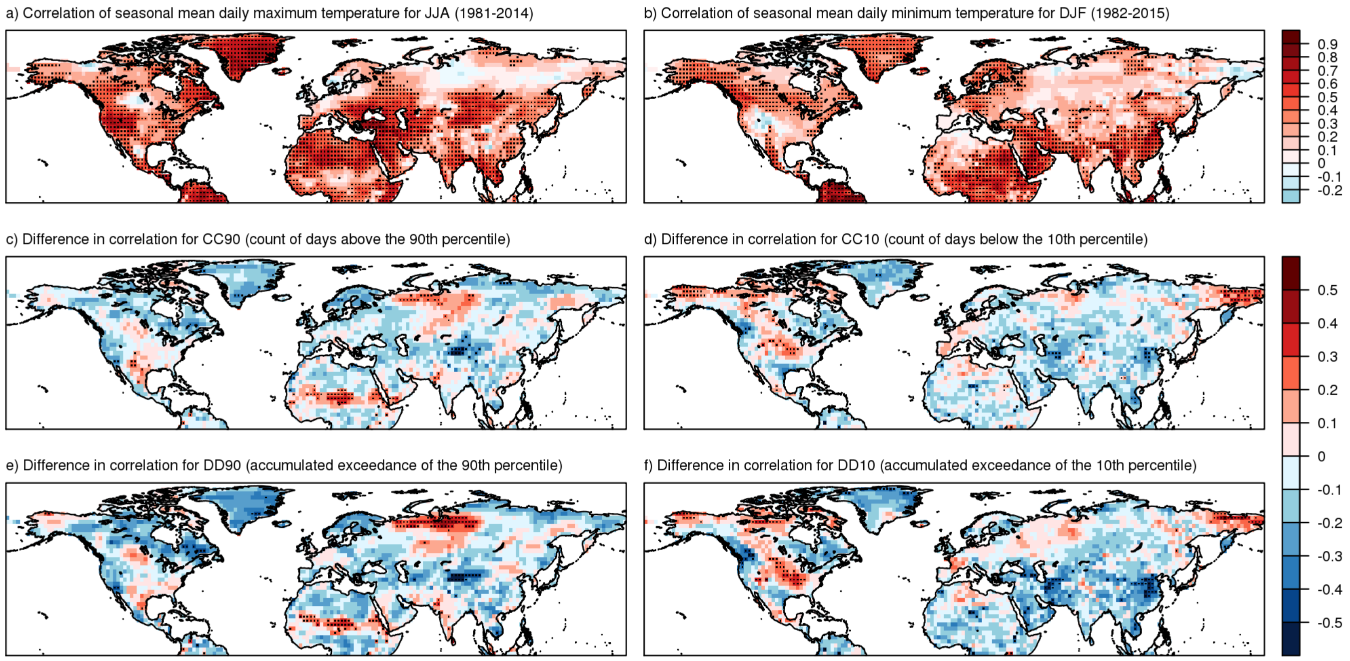


Figure 1: Effect of two bias correction techniques on the continuous ranked probability skill score of weekly average temperature (left) and precipitation sum predictions (right). For each season, the four box groups indicate the skill of week one (day 5-11) through week four (day 26-32), whereas the triple boxes represent raw model output, post-processing using mean de-biasing, and post-processing with quantile-mapping (from left to right).

MeteoSwiss serves recalibrated seasonal forecasts to customers. The recalibration is carried out using the climate conserving recalibration (CCR) proposed by Weigel *et al.* (2009). Also MeteoSwiss is exploring the potential for skilful forecasts of application-relevant climate indices. Such forecasts require post-processing of daily model output. In a recent publication (Bhend, Mahlstein, and Liniger, 2016) we investigated the relationship of predictive skill of forecasts of climate indices derived from daily values with the predictive skill of seasonal averages of said meteorological quantities. We find no indication of enhanced or reduced skill in forecasts of indices beyond what is expected due to the predictive skill in seasonal mean and the systematic reduction of skill when computing indices that describe events that are relatively rare (*Figure 2*).



a) Correlation of seasonal mean daily maximum temperature for JJA (1981-2014)

b) Correlation of seasonal mean daily minimum temperature for DJF (1982-2015)

c) Difference in correlation for CC90 (count of days above the 90th percentile)

d) Difference in correlation for CC10 (count of days below the 10th percentile)

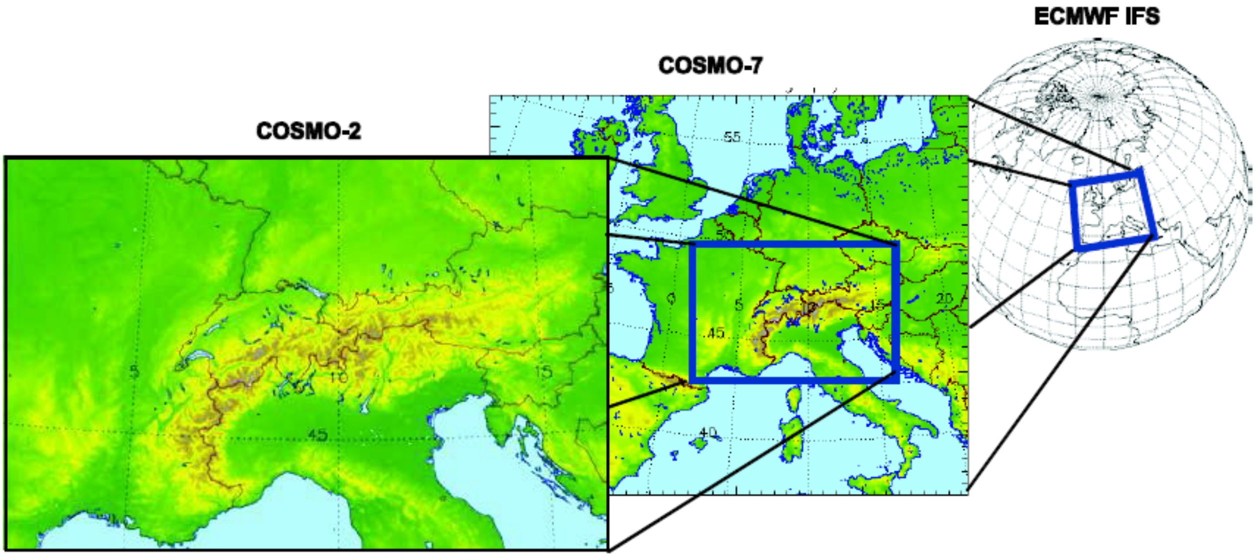
e) Difference in correlation for DD90 (accumulated exceedance of the 90th percentile)

f) Difference in correlation for DD10 (accumulated exceedance of the 10th percentile)

Figure 2: Correlation of seasonal mean daily maximum (a) and minimum (b) temperature and difference in correlation of counts of the exceedance of the 90th percentile (c) and below the 10th percentile (d) and of accumulated threshold exceedances respectively (e, f). Blue areas in c-f indicate that the correlation of the index forecast is lower than the correlation of the seasonal mean quantity. Stippling in panels a and b indicates correlations significantly (at the 5% level) larger than zero. Stippling in panels c to f indicates correlations for the index that are significantly different from what is expected due to the correlation in the seasonal mean and predictability of the seasonal mean only. None of the differences in c-f are field significant after controlling the False Discovery Rate (Bhend, Mahlstein and Liniger, 2016).

2.1.2 Physical adaptation

MeteoSwiss runs its own short-range forecasting system. The core of this system is the non-hydrostatic model COSMO (www.cosmo-model.org). It is running operationally at two spatial scales: The regional model COSMO-7 with a horizontal resolution of about 6.6 km is driven by the ECMWF global model IFS. The local model COSMO-2, having a horizontal grid spacing of about 2.2 km, is nested in COSMO-7. The nesting of NWP models is illustrated in Fig. 3.



Fiugure 3: NWP System of MeteoSwiss

Both COSMO-7 and COSMO-2 have their own assimilation cycle, which is updated in intervals of 3 hours. Three daily 72 hours COSMO-7 forecasts are calculated, based on the 00, 06 and the 12 UTC IFS (main or boundary conditions) runs. One COSMO-2 forecast is computed every 3 hours just after the computation of the necessary COSMO-7 boundary conditions. The lead time of the COSMO-2 forecast starting at 03 UTC is 45 hours, and 33 hours otherwise. The cut-off time for all forecasts is 45 minutes. At the end of March 2016, the new model COSMO-1 with 1.1 km grid spacing was becoming operational. It is directly nested into IFS-HRES. COSMO-2 will be phased out by the end of September.

A new ensemble prediction system COSMO-E with 2.2 km grid spacing was introduced in May. It is a 20 members ensemble with an additional control run, nested into IFS-ENS.

A sophisticated set of scripts controls the whole operational suite, and allows for a very high reliability of the system, with less than 2% of the forecasts requiring manual intervention. This same environment is also used to run parallel suites to validate proposed modifications to the system and to facilitate experimentation by the modelling group.

The Lagrangian particle dispersion model Flexpart is applied with IFS-HRES data to calculate the dispersion of air contaminants in the case of an accident worldwide.

2.1.3 Derived fields

2.2 Use of ECMWF products

3. Verification of products

3.1 Objective verification

3.1.1 Direct ECMWF model output (both HRES and ENS)

3.1.2 ECMWF model output compared to other NWP models

As part of the operational seasonal verification with SYNOP observations, IFS forecasts are regularly compared to the COSMO models operated at MeteoSwiss. For this, parameters such as pressure, 2 m temperature, 2 m dewpoint, 10 m wind speed and direction, cloud cover, precipitation (12-hourly and hourly), and 10 m wind gusts are compared to associated observations. For precipitation as one of the most important parameters, an example of a verification summary is shown in Fig. 4. Compared to COSMO-2 and COSMO-7, IFS shows a stronger overestimation of low precipitation amounts and also a stronger underestimation of high amounts. For thresholds of 10 mm/12 hours the values of all three models are quite similar.

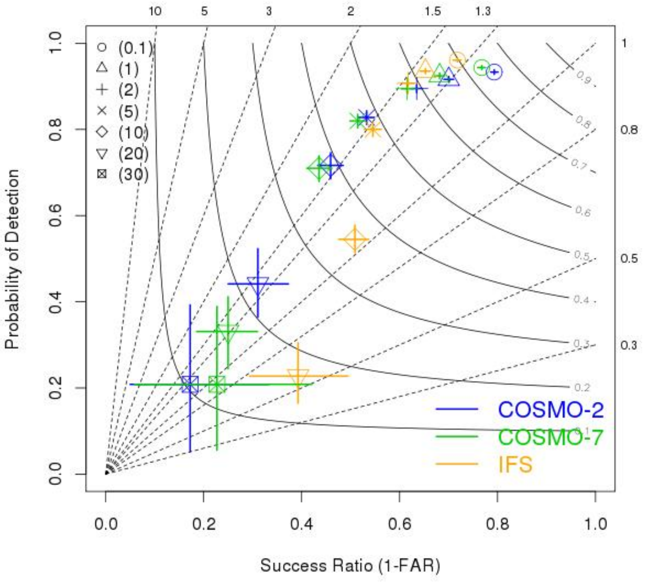


Figure 4: Comparison of the performance during Winter 2015/16 of COSMO-2, COSMO-7 and IFS 12 for 12 hour accumulated precipitation forecasts over Switzerland with the lead time +6 to +18 hours (for a detailed explanation of the diagram see Roebber, 2009).

Forecasts data from all 00 and 12 UTC base times and observational data from 146 Swiss stations are used to produce the diagram shown in Fig. 4. The results for 7 different thresholds ranging from 0.1 mm/12h (i.e. rain/no rain) to 30 mm/12 h (very strong precipitation) are differentiated by symbols. The low thresholds are overestimated by all models, and for the two lowest thresholds, the most by IFS-HRES. The high thresholds are about right for the COSMO models but underestimated by IFS-

HRES. The false alarm ratio however is lower for IFS. These two effects balance each other for IFS so that the resulting equitable threat score (curved lines) is similar to those of the COSMO models.

3.1.3 Post-processed products

A Kalman Filter is used to produce 3-hourly forecasts of temperature, dew point temperature and wind speed for station locations in Switzerland.

3.1.4 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

4. Feedback on ECMWF "forecast user" initiatives

5. References to relevant publications

Bhend, J., I. Mahlstein and **M.A. Liniger**, 2016. Predictive Skill of Climate Indices Compared to Mean Quantities in Seasonal Forecasts. *Quarterly Journal of the Royal Meteorological Society,* accepted for publication.

Mahlstein, I., C. Spirig, M.A. Liniger, and **C. Appenzeller**, 2015: Estimating daily climatologies for climate indices derived from climate model data and observations. J Geophys Res-Atmos, 120, 2808-2818

Roebber, **P. J.**, 2009: Visualizing multiple measures of forecast quality. Wea. Forecasting, 24, 601-608.

Weigel, A.P., M.A. Liniger and **C. Appenzeller**, 2009. Seasonal Ensemble Forecasts: Are Recalibrated Single Models Better than Multimodels? *Monthly Weather Review* **137,** 1460–1479. DOI: 10.1175/2008MWR2773.1.