Application and verification of ECMWF products 2017

Federal Office of Meteorology and Climatology MeteoSwiss

Christoph Spirig, Eugen Müller, Jonas Bhend, Samuel Monhart, Irina Mahlstein, Mark Liniger, Pascal Noti, Pirmin Kaufmann and Thomas Jordi

1. Summary of major highlights

2. Use and application of products

Use of ECMWF products within NinJo at MeteoSwiss

MeteoSwiss heavily relies on ECMWF products. While ensembles are increasingly used, the most important product in operational forecasting is still the IFS-HRES. Currently, we import around 26 GByte of raw data per run of the IFS-HRES into our visualization software NinJo (http://www.ninjo-workstation.com). Some additional parameters and aggregations are computed during import, so a final import of the IFS-HRES uses 68 GB of disk space. This data covers ‘only’ a larger European area and we still use the 0.125° resolution. We made some comparisons with the 0.1° dataset and the differences for the use by forecasters where negligible. Therefore the switch to the high resolution grid has not been as important for the forecaster desk as we thought it would be. Figure 1 shows exemplary IFS-HRES visualizations as provided in NinJo. A recently introduced new server framework running on Linux and using SSD technology improved the access speed to fields significantly, by a factor of two (single isobaric level) and 60 (hovmoeller request on all model levels), respectively.

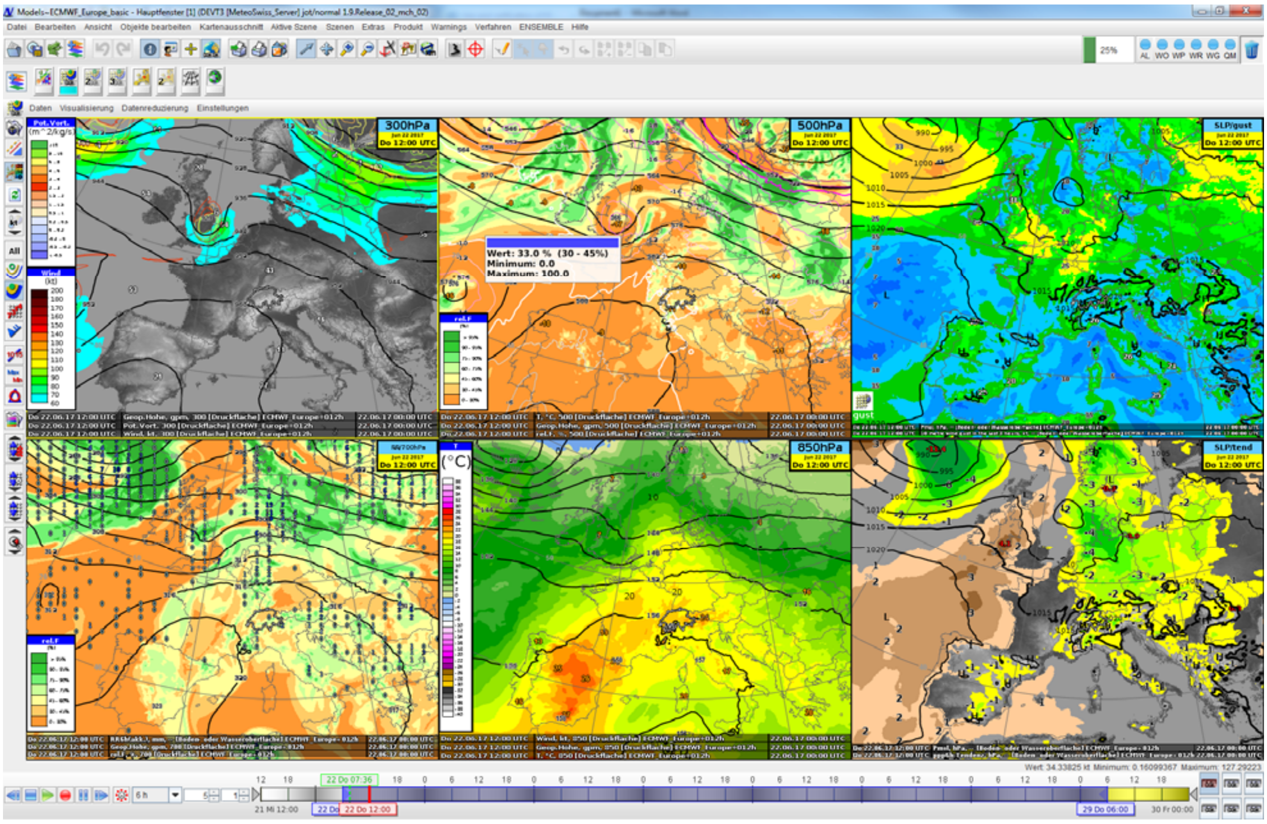


Figure 1: Overview of basic IFS-HRES visualization within NinJo. Parameters are displayed on surface or isobaric levels, full model coverage up to 100hPa is provided.

MeteoSwiss is downloading the full ENS output to the Swiss National Supercomputing Centre (CSCS) located at Lugano, where it is used to drive the COSMO Ensemble System of MeteoSwiss. Computations of different products are done there: To save bandwidth and disk space we only import ‘products’ into NinJo, such as percentiles and threshold probabilities. An overview on a basic ensemble visualization within NinJo is presented in Figure 2. Currently we do not import single members, but it is planned to import some key parameters of single members following the rollout of the next NinJo version in Winter 2017/18.

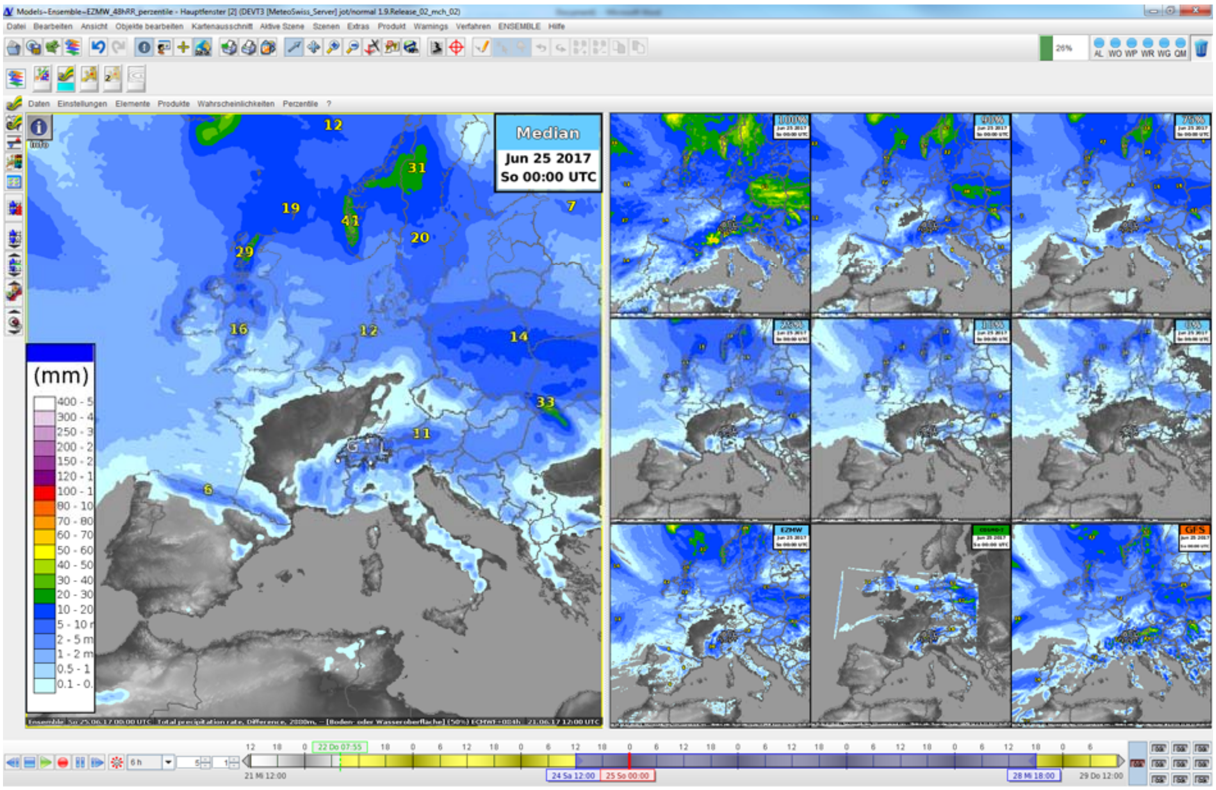


Figure 2: Visualization of ENS 24h precipitation totals in NinJo, showing the median (left) and 100/90/75/25/10/0 percentiles across Europe (top right), and IFS-HRES, COSMO-7, GFS (bottom right) for comparison.

2.1 Post-processing of ECMWF model output

2.1.1 Statistical adaptation

-/-

2.1.2 Physical adaptation

MeteoSwiss has modernized the suite of its limited-area models during the last year. The former COSMO-2 model has been replaced by the new deterministic model COSMO-1 and the new ensemble system COSMO-E. COSMO-1 has a grid spacing of 1.1 km and 80 levels and runs 8 times per day out to 33 hours (once per day to 45 hours). COSMO-E is a 20 member ensemble with 2.2 km grid spacing and also 80 levels, with a maximum forecast range of 120 hours. The regional model COSMO-7 with a grid spacing of 6.6 km, covering a larger domain than COSMO-1 and COSMO-E, is remaining for just a few more years. It is however planned to be phased out during 2019 and to be replaced by the ECMWF global model HRES. COSMO-1 and COSMO-7 are both directly one-way nested into the HRES, COSMO-E is nested into ENS. The domains of the COSMO models and their topography are illustrated in Fig. 3.

Figure 3: Domains of COSMO models at MeteoSwiss. Left: COSMO-1, middle: COSMO-E, right: COSMO-7

All COSMO models have their own assimilation cycle, which is updated in intervals of 3 hours. While COSMO-E uses an Ensemble Kalman Filter analysis, COSMO-1 and COSMO-7 both still use nudging.

2.1.3 Derived fields

-/-

2.2 ECMWF products

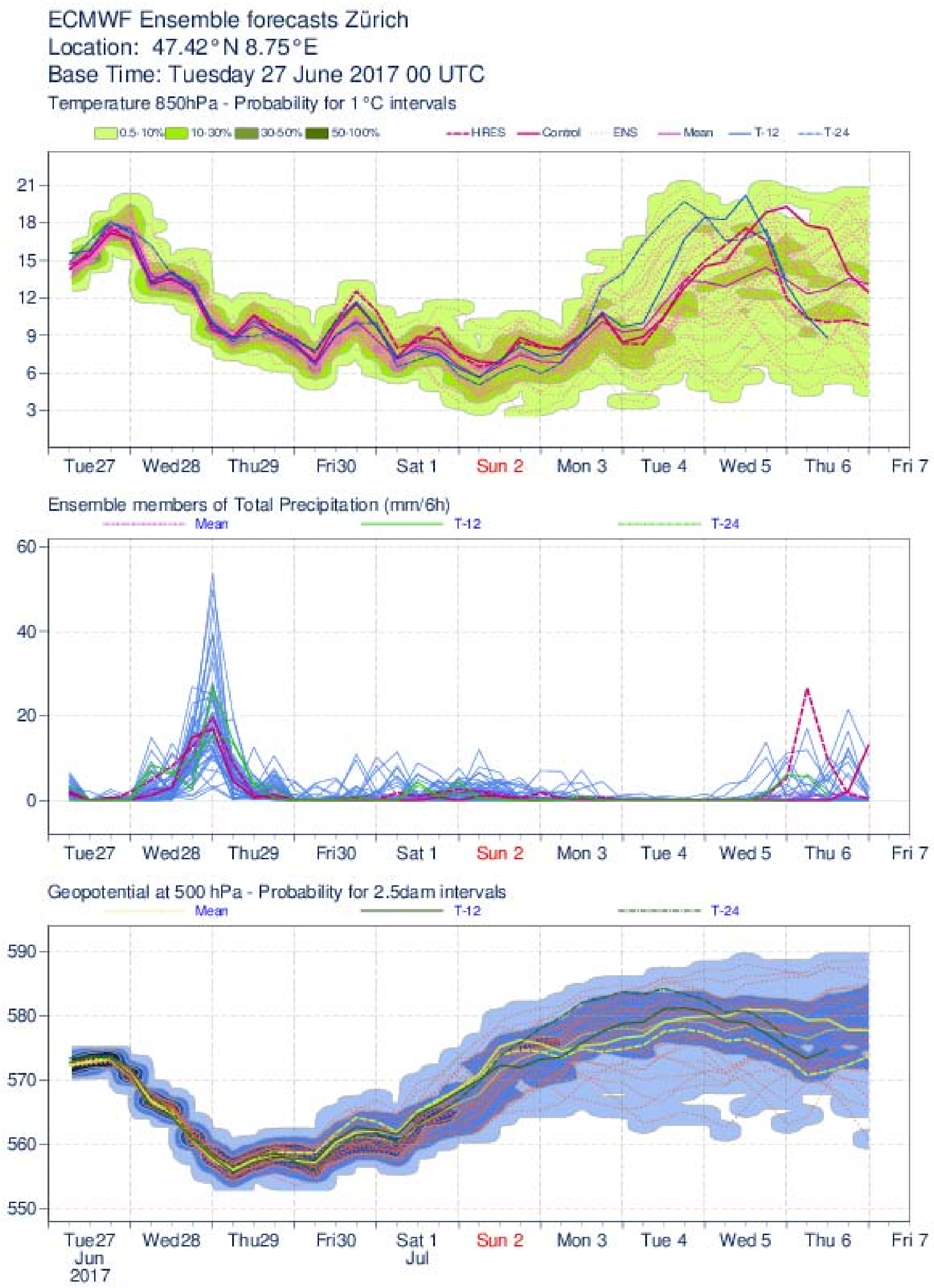
2.2.1 Use of Products

The forecast service of MeteoSwiss provides a daily medium-range forecast, based on the 12 UTC run of the HRES und ENS, and it’s updated with the following 00 UTC run. The medium-range forecast consists of a text forecast and a table whose input values come from the IFS and can be adapted by the forecaster. The table values then serve as basic data for automatically generated products, e.g. Website or App of MeteoSwiss. The predictions range up to D + 7, with the last two days (D + 6 and D + 7) being briefly and generally described as a trend. As already mentioned, the forecaster mainly based his work on the ECMWF IFS products when preparing the medium-range forecast. Depending on the predictability, the HRES has more weight or the ENS or both equivalent. The trend of the HRES and Ensemble Mean is compared with the overall ENS. If it runs parallel or opposite or even outside the ENS Spread. The recommendations in chapter 4.5 from the User Guide to ECMWF forecast products serve as a guideline for the interpretation by the forecaster.

In addition to various ENS products such as Ensemble Mean maps, probabilities or meteograms, the ENS plumes are an important basis for medium-range forecasting. In practice, the HRES run can be in good agreement with the ENS or not, and there may be big differences from run to run for the same valid time (jumpiness). Chapter 4.5 of the User Guide describes how to deal with the different prediction situations. The recommended way of interpretation is not only to use the current IFS run, but also the two previous HRES runs and the Ensemble Mean.

However, in the available standard ENS plumes provided by the ECMWF, together with the probabilities from the ENS predictions and the control run, only the current HRES run is displayed. The previous HRES runs and the Ensemble Mean are not in the plumes graphics. Since the ECWMF recommends in its User Guide to compare the ENS with the three latest HRES runs and the Ensemble Mean, MeteoSwiss's forecast service also wanted to implement this recommendation in the operational forecast process. apply. On request from MeteoSwiss, the ECMWF User Support Section extended the script for the generation of the ENS plumes accordingly, so that now the courses of the Ensemble Mean and the two previous HRES runs (H-12 and H-24) are displayed. The new plumes are produced operationally for Zurich, Geneva and Lugano.

The completed ENS plumes have proved their worth in the operational forecast service of MeteoSwiss, and the ECMWF recommendations for the interpretation of medium-range forecast products can be better adapted now. The new plumes help to estimate the predictability/reliability as well as effects such as jumpiness. As the meteorologist, who is responsible for the medium range, changes almost daily, the previous changes of forecasts can now better take into account than before. And with the additional previous HRES-runs it’s easier to form a small ensemble in certain situation of poor predictability, according to the ECMWF User Guide.



Extended ENS-plumes for Zurich with the ENS-Mean and the two previous HRES-runs.

2.2.2 Product requests

-/-

3. Verification of products

3.1 Objective verification

3.1.1 Direct ECMWF model output (both HRES and ENS)

The routine seasonal model verification at MeteoSwiss includes both HRES and ENS. The parameters verified are pressure (at station height and reduced to sea level), 2 m temperature, 2 m dewpoint, 10 m wind speed and direction, total cloud cover, precipitation (12-hourly and hourly), 10 m wind gusts, relative humidity, global radiation. These parameters are available at over 150 stations in Switzerland. Some of these parameters are not available on the larger verification domains of COSMO-7 (encompassing approximately 1400 SYNOP stations) and COSMO-1/COSMO-E (encompassing more than 560 SYNOP stations).

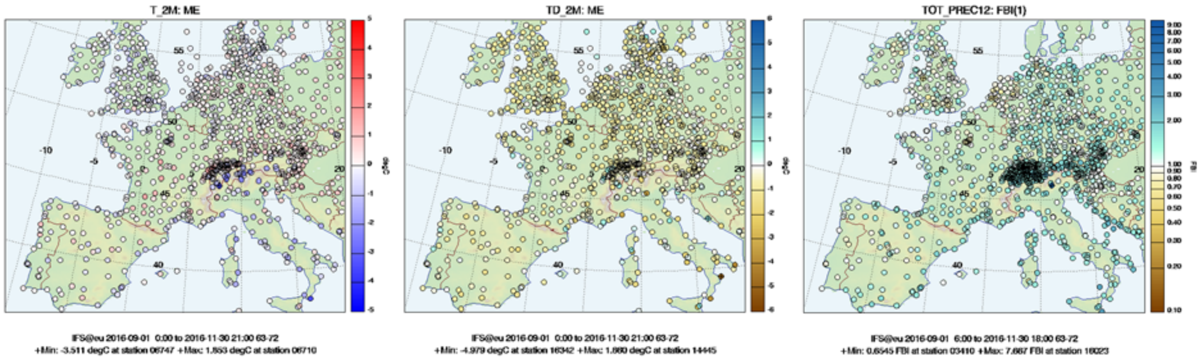


Figure 1: Verification of HRES surface parameters at the 63 – 72 hours lead-time range for Summer 2016 over the COSMO-7 domain. Left: bias of 2 m temperature, middle: bias of 2 m dewpoint, right: frequency bias of 12 hourly precipitation sums above the 1 mm/12h threshold.

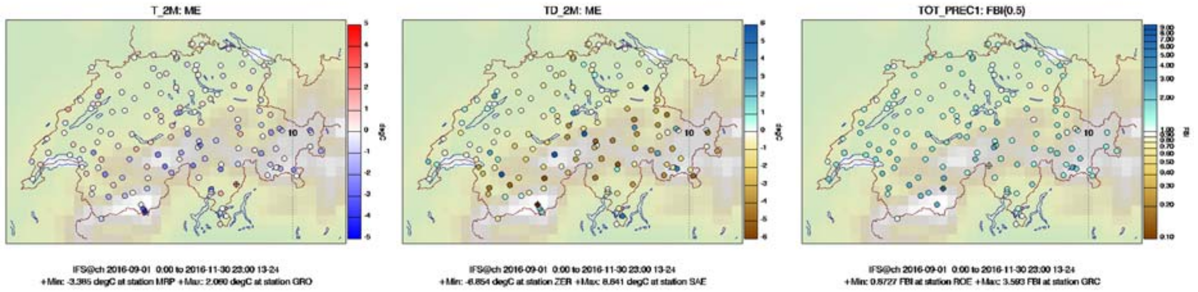
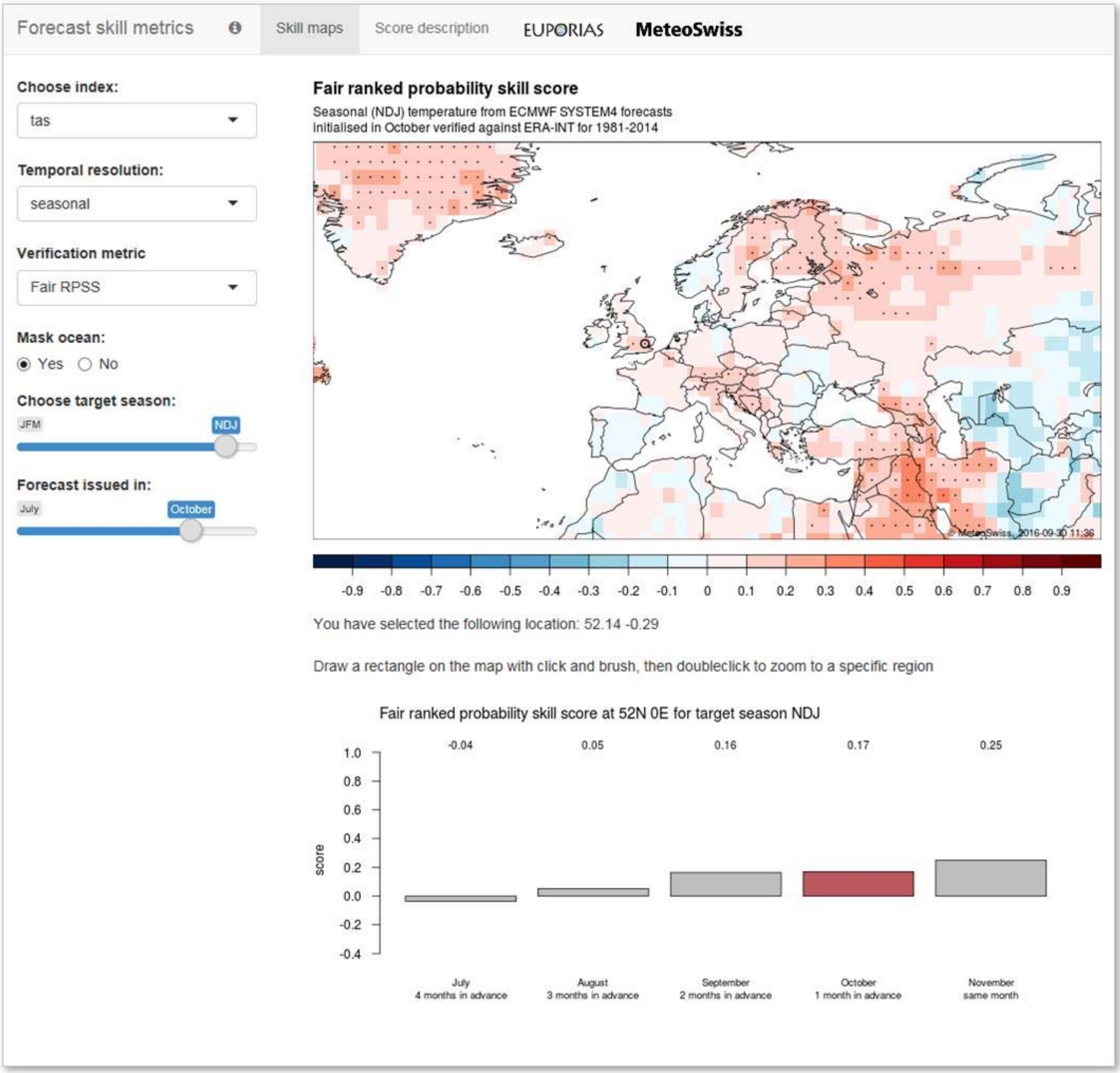


Figure 2: Verification of HRES surface parameters at the 13 – 24 hours lead-time range for Summer 2016 over Switzerland. Left: bias of 2 m temperature, middle: bias of 2 m dewpoint, right: frequency bias of hourly precipitation sums above the 0.5 mm/h threshold.

Some spatial plots of our operational surface verification for for Summer 2016 are shown in Figs. 1 and 2. Fig. 1 are the results for HRES on the COSMO-7 domain and day 3, Fig. 2 the results for HRES at the Swiss stations and day 1.

Seasonal Forecasts

As part of the FP7 project EUPORIAS, MeteoSwiss has compiled a comprehensive verification of the ECMWF System4. Monthly and 3-monthly mean near-surface temperature and precipitation forecasts have been verified against ERA Interim using a range of scores and measures to quantify various aspects of forecast quality. As part of this project, a public web platform https://meteoswiss-climate.shinyapps.io/skill_metrics has been set up that allows users to browse the verification results (Wehrli et al, 2017). The web platform allows users to zoom in regionally and to investigate the evolution of forecast quality for specific times of the year that are of relevance to the user's application (Figure 3).



Choose index: tas

Temporal resolution: seasonal

Verification metric Fair RPSS

Mask ocean: ● Yes ○ No

Choose target season: JFM NDJ

Forecast issued in: July October

Fair ranked probability skill score

Seasonal (NDJ) temperature from ECMWF SYSTEM4 forecasts initialised in October verified against ERA-INT for 1981-2014

You have selected the following location: 52.14 -0.29

Draw a rectangle on the map with click and brush, then doubleclick to zoom to a specific region

Fair ranked probability skill score at 52N 0E for target season NDJ

Figure 3: Screenshot of the web platform visualizing forecast quality measures for the ECMWF System4 seasonal forecasts. In the main panel, a map of ranked probability skill scores (RPSS) for November to January (NDJ) mean temperature from forecasts initialized in October. Stippling denotes RPSS that are significantly larger than zero (no skill). In the bottom panel, the evolution of RPSS for NDJ temperature forecasts at a grid box in the southwest of the UK is shown.

3.1.2 *ECMWF model output compared to other NWP models*

As part of the operational seasonal verification, HRES forecasts are regularly compared to the deterministic COSMO models COSMO-1 and COSMO-7, while ENS is compared to COSMO-E using probabilistic scores. For precipitation as one of the key parameters, an example of a verification result is shown in Fig. 4.

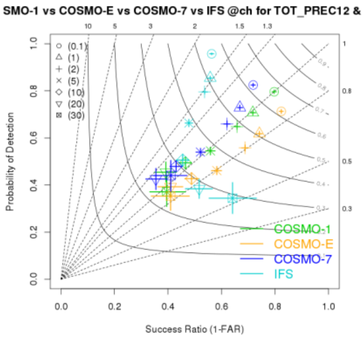


Figure 4: Comparison of the performance of the models COSMO-1, COSMO-E Control Run, COSMO-7, and IFS-HRES during Summer (JJA) 2016 in a performance diagram (for a detailed explanation see Roebber, 2009). The verified parameter is 12 hour accumulated precipitation (lead time +6 to +18 hours) over Switzerland.

Forecasts data from all 00 UTC and 12 UTC base times and observational data from all Swiss stations are used to produce the performance 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 the shape of the symbols. Compared to COSMO-1, COSMO-E Control Run, and COSMO-7, HRES shows a stronger overestimation of low precipitation amounts and also a stronger underestimation of high amounts. The crossover from over- to underestimation is approximately at 10 mm/12 hours. The strong overestimation of the low thresholds by HRES as indicated by the position of the respective symbol high above the diagonal is known as "drizzle problem". This deficit is most prominent in summer and, as can be seen in Fig. 2, is more prominent within topography and, according to Fig. 1, to the south of Europe than to the north, but these data outside Switzerland are not included in Fig. 4. The events for the high thresholds are rare and show correspondingly large confidence intervals as indicated by the large error bars (crosshairs).

Note that with the increase of the HRES resolution in Spring 2016, an averaging of 3 x 3 grid points has been introduced for the precipitation of HRES, like it is already applied to the COSMO models, to account for the low predictability at the grid resolution. This leads however to a slight over-emphasis of the aforementioned error characteristics and has therefore been removed for later seasons.

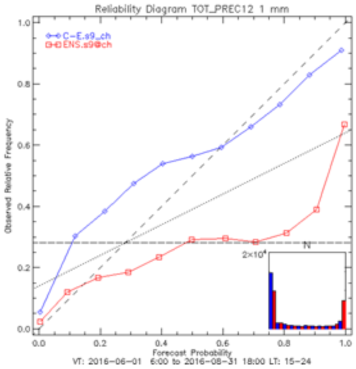


Figure 5: Reliability diagram for ENS (red) compared to COSMO-E (blue) for the 1 mm/12h precipitation threshold at Swiss stations during Summer 2016.

Fig. 5 shows the comparison of the two ensembles ENS and COSMO-E. ENS overestimates all but the lowest probabilities, while COSMO-E underestimates the low probabilities up to the 50% probability but is quite reliable for higher probabilities. This demonstrations the strong overestimation of weak precipitation occurrence over Switzerland in ENS. As in Fig. 4 for HRES, this effect is slightly over-emphasized in this plot through the 3 x 3 grid point averaging that has been applied to the model precipitation field.

3.1.3 Post-processed products

Monthly forecasts

In the framework of the Horizon 2020 project Heat-Shield (an initiative aiming to increase the thermal resilience of European workers in the context of global warming, https://www.heat-shield.eu/), MeteoSwiss is currently setting up an early warning system for heat stress based on ECMWF's monthly forecasts. As part of these developments, verification of the extended range ensembles against observations at 1800 European surface sites has been carried out. This extensive surface observation data set was obtained by combining the data sets ECA&D (European Climate Assessment & Dataset, http://www.ecad.eu), GSOD (Global Surface Summary of the Day, https://data.noaa.gov/dataset/global-surface-summary-of-the-day-gsod), and selected records of SMN (SwissMetNet, the national monitoring network of MeteoSwiss). At these locations, the performance of the 20 years hindcasts of cycle 41r2 initialized between begin of May and end of July has been analysed. The focus of the verification was on temperature and humidity as these are the dominating variables affecting physiological heat stress, which can be expressed by indices such as the wet bulb globe temperature (WBGT = *f*(T, humidity) for shaded conditions). Daily model output of temperature and dewpoint temperature has been bias-corrected with a quantile mapping approach (Rajczak et al., 2016), by using the hindcasts and corresponding observations in leave-one-out mode. WBGT hindcasts were then computed from these bias-corrected hindcasts of temperature and dewpoint temperatures. Figure 4 shows the resulting ensemble mean correlations for temperature, dewpoint, and WBGT at a lead time of 15 days. Overall, the correlation of the WBGT index combining humidity and temperature was slightly better than those of its underlying variables, and the spatial pattern of WBGT correlation seems to profit from the strengths in each of the underlying variables. On average, WBGT correlations remained marginally positive up to lead times of 30 days, skill in terms of the continuous ranked probability skill score (CRPSS) against climatology remained positive up to lead times of 20 days.

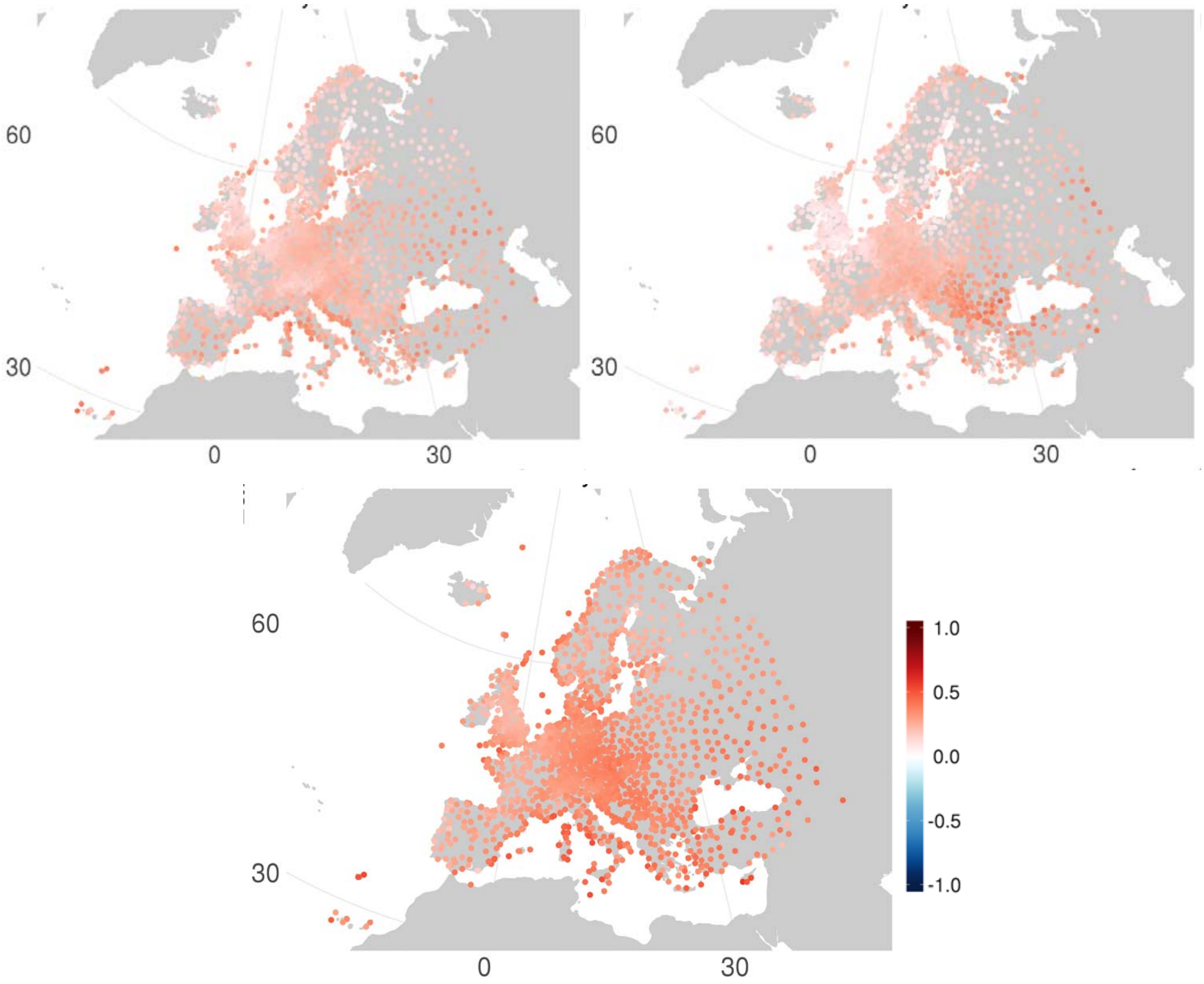


Figure 4: Ensemble mean correlation of surface temperature (top left), dewpoint temperature (top right), and WBGT (bottom) at a lead time of 15 days for 1996-2015 hindcasts initialized between May and July.

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

Rajczak J., S. Kotlarski, N. Salzmann, and **C. Schär** (2016)**.** Robust climate scenarios for sites with sparse observations: a two-step bias correction approach. International Journal of Climatology, **36** (3), 1226-1243.

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

Wehrli, K., J. Bhend and **M. A. Liniger** (2017). *Systematic quality assessment of an operational seasonal forecasting system.* Technical Report MeteoSwiss, 263, 52 pp. URL: http://www.meteoswiss.admin.ch/home/services-and-publications/publications.subpage.html/en/data/publications/2017/3/systematic-quality-assessment-of-an-operational-seasonal-forecasting-system.html