# Application and Verification of ECMWF Products 2019

Deutscher Wetterdienst (DWD)

Reinhold Hess, Andreas Lambert, Guido Schröder, Sebastian Trepte (Forecast Application Development)

## 1. Summary of major highlights

ECMWF products are used as input for the NWP post-processing procedures at DWD.

## 2. Use and application of products

### 2.1 Direct Use of ECMWF Products

### 2.2 Other uses of ECMWF output

#### *2.2.1 Post-processing*

HRES 00 UTC und 12 UTC are used for the DWD point forecast product MOSMIX. The DWD MOSMIX product provides statistically optimized point forecast for more than 5400 points worldwide based on multiple linear regression techniques. For detailed information, see

https://www.dwd.de/EN/ourservices/met_application_mosmix/met_application_mosmix.html.

The IFS ensemble is used to generate daily warning charts for high impact weather in the daily model guidance (DMG). The was presented at UEF2019

(https://events.ecmwf.int/event/119/contributions/576/attachments/113/198/UEF2019-Schroeder.pdf).

MOS system on HRES for aviation: TAF Guidance. Calibrated point forecasts (airports) for lead times up to 41 hrs. Lagged average of 00 and 12 UTC HRES model runs (weighting based on MOS). Deterministic and probabilistic forecast parameters and automated TAF encoding (AutoTAF).

ECMWF-ENS output is postprocessed by EnsembleMOS of DWD as part of ModelMIX in order to support warning management for severe weather. Statistical products of ECMWF-ENS like mean and spread are used as predictors for the MOS-System in order to calibrate probabilistic forecasts by multiple linear and logistic regression.

#### *2.2.2 Derived fields*

- Extreme Weather Index (EWI). A short description can be found in: *https://www.ecmwf.int/sites/default/files/elibrary/2019/18944-global-icon-eps-contribution-tigge.pdf*

#### *2.2.3 Modelling*

## 3. Verification of ECMWF products

### 3.1 Objective verification

ECMWF HRES and ENS forecast serve as a verification benchmark for the global DWD models ICON and ICON-EPS. An extensive verification of the latest forecasts against observations from different observation systems (mainly SYNOP and TEMP but also AMV, scatterometer, GPSRO, satellite radiation, aircrafts etc.) is performed on a daily basis. In order to achieve fair and comparable scores, the same verification suite and set of observations is used for all models. For the deterministic model the widespread set of continuous verification scores is computed (RMSE, MAE, BIAS, STDEV,R2). The categorical verification of threshold exceedances is performed using contingency table based scores for the deterministic models. The ensemble system is verified with measures like CPRS, reliability diagram, rank histogram, values score or ROC are available. Scores are available as domain- and verification period average, as time series of domain averages and period average station based.

The following plots should give an indication of bias and quality differences between the global DWD and ECMWF models with an emphasis on seasonality and quality trends. Note that the verification activities at DWD model developing group are not focused on finding structural errors of the ECMWF forecast system. ECMF forecast are rather used as a benchmark.

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

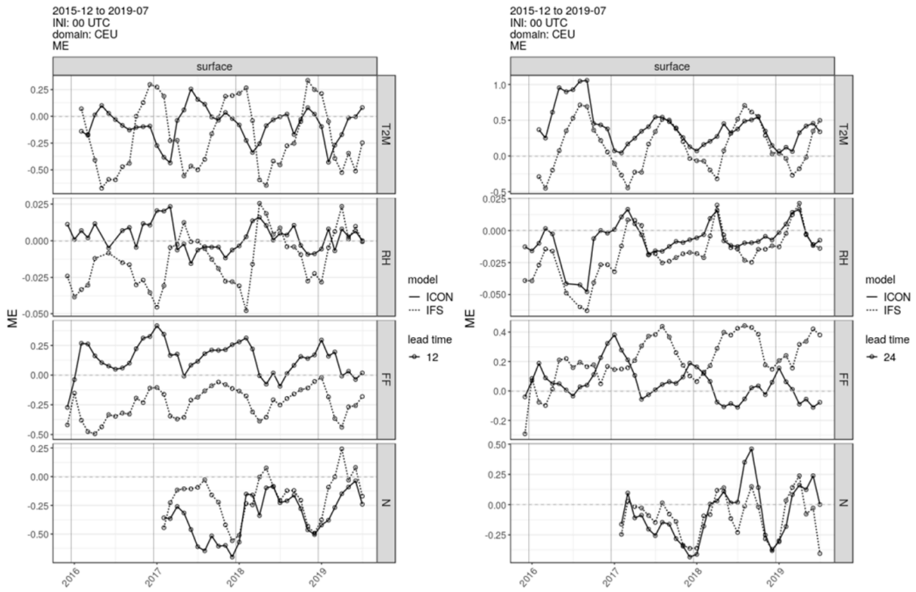


Fig.1 Mean Error (ME) of deterministic IFS (HRES) and ICON surface observations 2m temperature (T2M [K]), relative humidity (RH [0..1]), wind speed (FF [m/s]), and total cloud cover (N [oct.]). The scores are for Europe (old COSMO-EU domain) calculated with the same method on the same set of observations. The 12h (left) and 24h (right) forecasts of the 00UTC runs are evaluated. Mayor differences between the models can be seen in the seasonality of T2M and RH at noon. Generally the IFS-HRES seasonality of T2M is overestimated ant night and underestimated at noon. For wind speed the bias differences between day and night are more pronounced for IFS-HRES forecasts. Could cover biases happen to be comparable between both models and independent from the daytime, showing too little cloud cover in winter months.

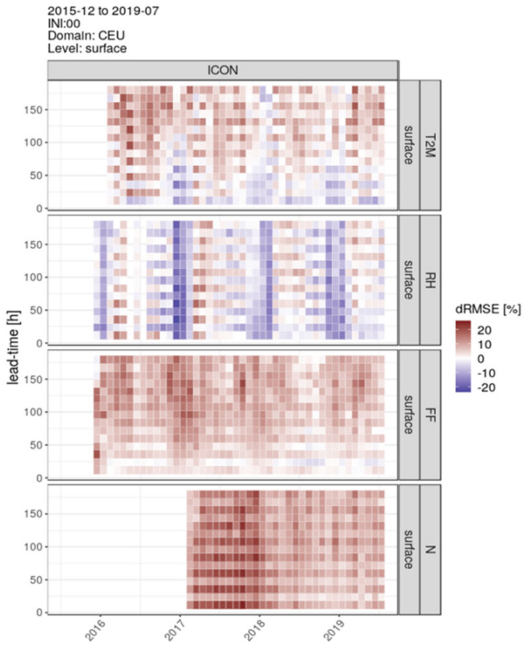


Fig.2 Relative difference in root mean squared error (dRMSE [%]) of deterministic ICON-IFS for surface observations (SYNOP), domain and runs as in Fig 1. Blue (red) colours show better scores for ICON (IFS). For shorter lead-times ICON quality is comparable to that of IFS, RH and T2M in winter show advantages for ICON. At longer lead-times IFS forecast become increasingly better compared to ICON. IFS total cloud cover forecasts are superior in all seasons at all lead-times, though ICON seems to be approaching in the recent months.

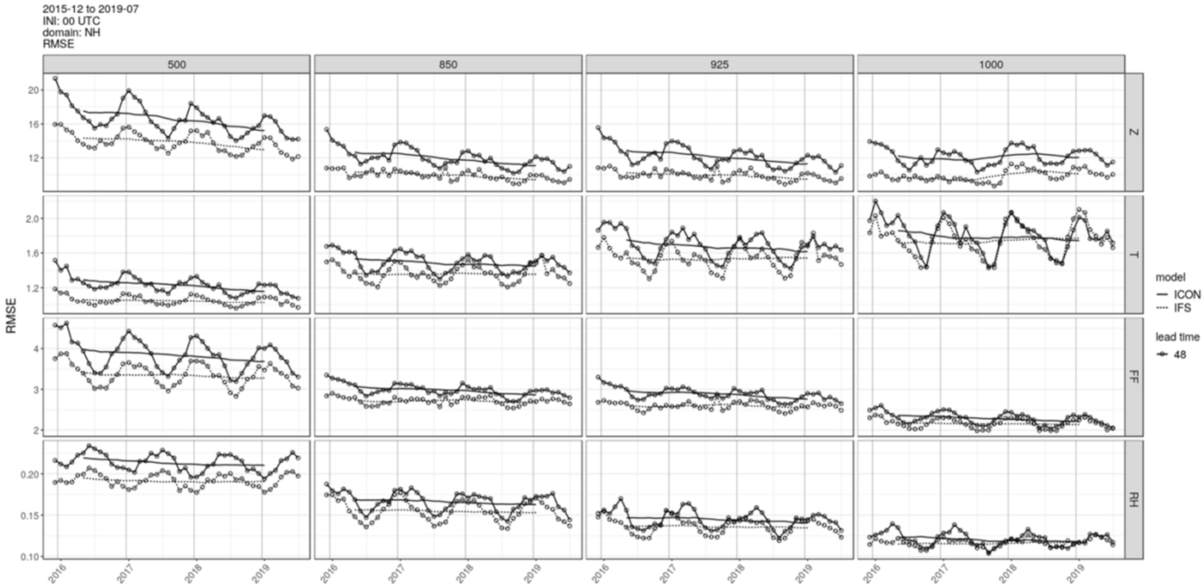


Fig.3 Time-series of monthly RMSE scores against TEMP observations at different levels [hPa] for 00UTC runs in the northern hemisphere. The 12 months running mean is shown in addition. IOCN and IFS-HRES show a very similar seasonality, scores seem to have improved for both models over the last years with a slightly faster improvement of the ICON in most cases. Generally the differences between the scores are smaller in the lower levels and also for shorter lead-times (not shown).

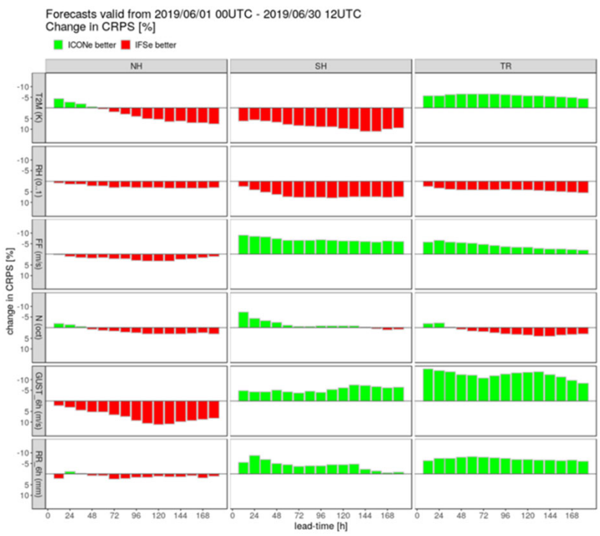


Fig.4 Relative difference in CRPS between ICON-EPS and IFS-EPS in June 2019. 00UTC and 12UTC runs are aggregated and scores are shown for northern-, southern hemisphere and tropics (NH, SH, TR). Verification is performed against SYNOP observations. For short lead-times ICON-EPS and IFS-EPS show comparable scores in the NH, larger quality differences occur in TR and SH. Especially for wind speed and gusts the ICON-EPS performs better in the SH and TR.

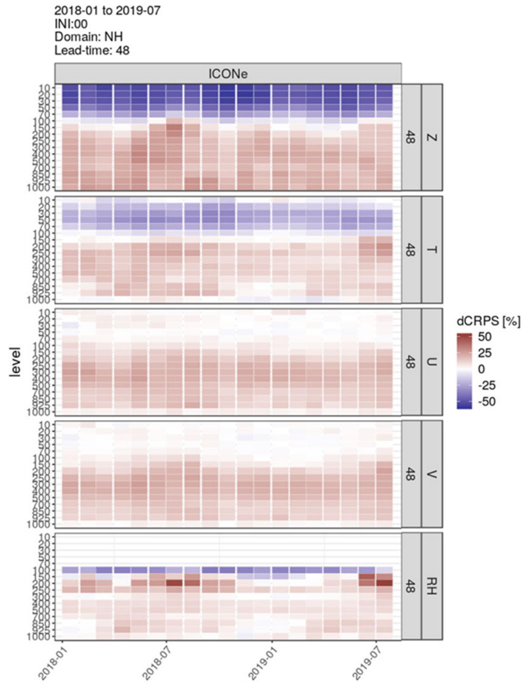


Fig.5 Monthly relative differences in CRPS between ICON-EPS and IFS-EPS 00UTC runs in the northern hemisphere for different levels [hPa]. Scores are relatively close near the surface, again with advantages for the IFS-EPS with growing lead-time (not shown). The much better ICON-EPS scores in the stratosphere (geopotential (Z) and temperature (T)) can be attributed to large biases of the IFS-EPS. The improvement in ICON scores compared to IFS, as in case of the deterministic forecast, is less pronounced.

#### *3.1.2 Post-processed products and end products delivered to users*

External verification of TAF Guidance and AutoTAF products for locations by Austro Control. Based on contingency table of probabilistic parameters against METAR. Related score KPI=(HSS+PSS)/2 and Performance Diagram. For details see attached EGU talk (Trepte_EGU2018.pdf).

#### *3.1.3 Monthly and Seasonal forecasts*

### 3.2 Subjective verification

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

- Subjective verification of the ICON-EPS versus the ECMWF-EPS focusing on events relevant to the forecasters of DWD when issuing warnings.
  https://www.ecmwf.int/sites/default/files/elibrary/2019/18944-global-icon-eps-contribution-tigge.pdf

#### *3.2.2 Case studies*

## 4. Requests for additional output

## 5. Feedback on ECMWF “forecast user” initiatives

## 6. References to relevant publications

**Mahringer, G.**, 2008: Terminal aerodrome forecast verification in Austro Control using time windows and ranges of forecast conditions. *Meteorol. Appl.* 15, 113–123.

**Trepte, S., M. Eckert** and **G. Mahringer**, 2018: Using Model Output Statistics for Aerodrome Weather Forecasts. *Geophysical Research Abstracts*, Vol. 20**,** EGU2018-16112-1, EGU General Assembly 2018

https://www.dwd.de/DE/leistungen/lf_17_taf_guidance/taf_guidance.html (german only)