# Application and verification of ECMWF products 2018

Finnish Meteorological Institute – compiled by Weather and Safety Centre with help of several FMI's experts

## 1. Summary of major highlights

FMI's 0-10 days' forecasts are based on ECMWF but during the first few days we use other models together with ECMWF such as Harmonie and Hirlam. The weight of different models in our forecasts depends on the weather situation. We are generally very content with ECMWF model and products but still 2m temperature challenges, which have been reported during the past years', do exist: negative bias in spring especially in evenings and over-forecasting during cold spells.

## 2. Use and application of products

Different ECMWF products are used in our weather service, research side and by our customers. ECMWF deterministic and ENS products are utilised in our forecasts. ECMWF data is also used as boundary conditions for Hirlam/Harmonie. Usage and demand of ENS, monthly and seasonal forecasts is increasing by different customer sectors. However it is vital to understand the uncertainties related to these "longer term" forecast products so that customers know how those products should be used in their decision making.

### 2.1 Post-processing of ECMWF model output

The role of the forecaster in our production system is significant over the Scandinavian region. Forecasters are postprocessing the data provided by different weather prediction models, by making model selection(s) and necessary adjustments to model fields. European forecasts outside of the Scandinavian region are mainly based on raw ECMWF data except for the 2m temperature, which is MOS-calibrated (Model Output Statistics), height-corrected and land-sea interpolated. Also precipitation type is postprocessed from ECMWF fields and adjusted using MOS 2m temperature. Global forecasts outside of European region are mainly based on raw ECMWF data except for the 2m temperature, which is corrected using Kalman filter or calibrated by height correction and land-sea interpolation.

#### *2.1.1 Statistical adaptation*

FMI has substantially developed statistical calibration and post-processing systems in recent years. Our MOS approach is traditional: Station-specific forecasts are first generated, based on a linear regression with training period starting from December 2011. Station-specific MOS forecasts are gridded using Kriging method. MOS-corrected ECMWF 2m temperature forecasts have operationally been used in production since January 2017 for the European region. The results of MOS calibrated ECMWF 2m temperature have been good over Finland especially during spring and summer, whereas the improvement for our routinely used hit rate metric has not been as significant during autumn and winter (Figure 1). Forecasting wintertime 2m temperatures over Finland have long been very challenging for several reasons: Lack of predictable diurnal variation, very shallow and strong temperature inversions and infrequent variability posed by pronounced baroclinity all contribute to larger wintertime forecasting errors. For strong inversions, where raw ECMWF has large warm bias, the duty forecaster improves the forecasts but our MOS still behaves too conservatively. In particular MOS calibration has improved the 2m temperature forecasts in Finland during spring, when ECMWF has a systematic cold bias in evening temperatures. MOS calibration is also applied to dew point temperature, which is currently in validation (see Fig. 2). Even though T2 and D2 are separately calibrated using MOS technique, the derived RH values only very rarely (99,96% are below 100%) exceed 100% which is very encouraging. In general, MOS technique conservatively improves over raw ECMWF forecasts but preserves the general characteristic of its RH distribution.

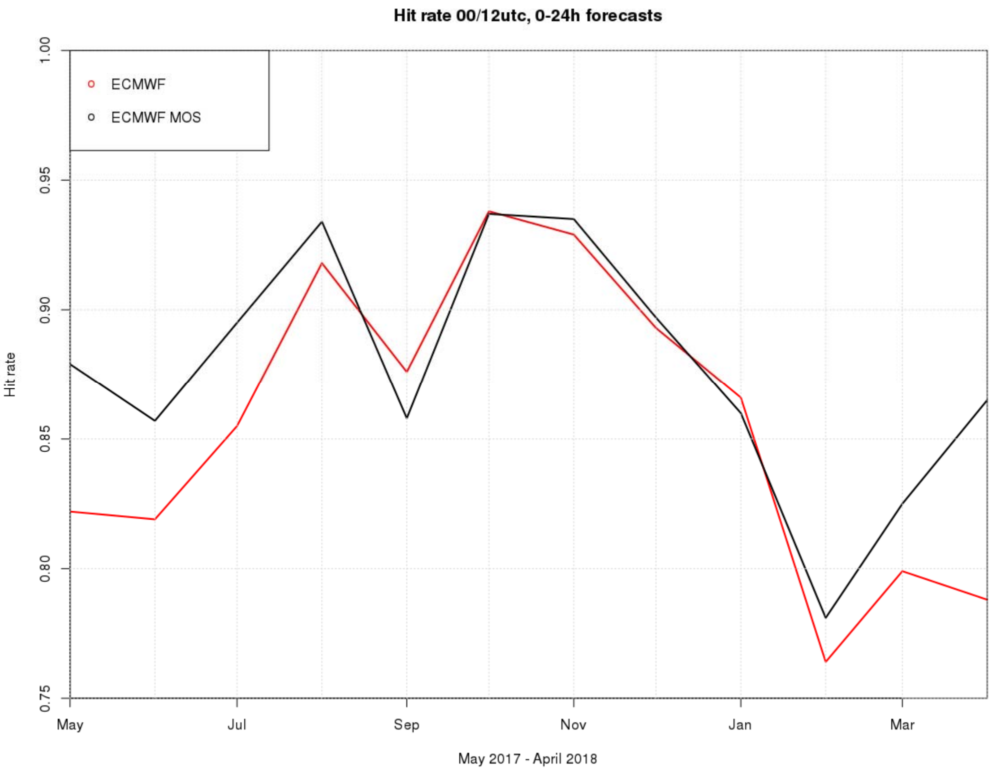

*Figure 1. Hit rates for MOS-calibrated and raw temperature forecasts for 30 inland stations in Finland. Mean value for forecast hours 3,6,9,12,15,18,21,24*

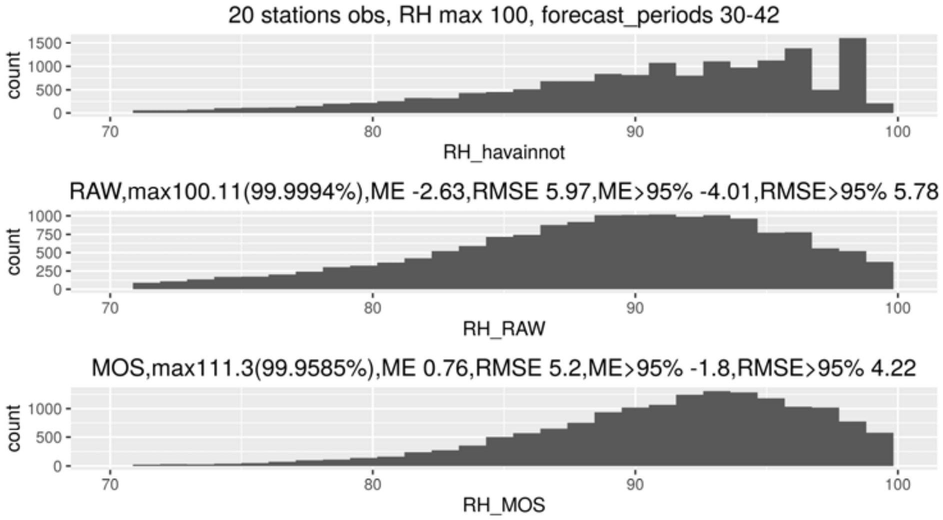

*Figure 2. RH verification results for 20 Finnish stations, DJF period over the season 2017-2018. In the figure, forecasts from the forecast lengths 30,33,36,39,42 are treated as one. In terms of both ME and RMSE, MOS-calibrated forecasts (bottom) show better agreement with the observations (top) as compared to direct model output (middle). In both forecasts, the general shape of the RH distribution is not well captured.*

### 2.1.2 Physical adaptation

ECMWF data is also utilised as boundary conditions for limited area models Hirlam and Harmonie (MetCoOp EPS) which is partly run at FMI (in collaboration with Sweden and Norway), dispersion and trajectory models, hydrological models (run by Finnish Environmental Institute), road condition models and wave models.

### 2.1.3 Derived fields

ECMWF data is used to calculate parameters that the model doesn’t provide. FMI has a dedicated group of forecasters who develop and implement functions in collaboration with our ICT and Data Production unit to calculate derived fields from model data e.g. probability of precipitation, probability of thunder and numerous parameters related to aviation weather.

## 2.2 ECMWF products

### *2.2.1 Use of Products*

Concerning severe weather situations various ECMWF products are used. Temperature, precipitation, wind gust and CAPE ENS distributions are among the most used ones. EFIs and cyclone tools are useful and becoming more utilized by our forecasters. The ENS precipitation type product has also proven to be useful, however it would be great to have it accessible also via clickable charts rather than only via ecCharts. In addition the new ECMWF sounding tool which indicates the spread of the ensemble is an excellent product and will probably be widely used by our forecasters.

FMI calculates several stability indices from the deterministic runs as a part of our post-processing routines. In addition forecasters can write various kinds of scripts within the FMI meteorological workstation which enables the generation of additional derived forecasting parameters such as precipitation type, wind gusts, probability of thunder, wind chill index, etc. These new ideas and parameters developed by forecaster can also be taken into post-processing routines and even into new customer products. This low-threshold development process efficiently supports innovative development with relatively low costs.

### *2.2.2 Product requests*

ECMWF has newly introduced several parameters for thunderstorm forecasting. Maximum CAPE, lightning density and sounding tool are all very good additions to the previous set of forecasting parameters. However because 9km global model cannot resolve convection, evolvement and advection of convective clouds and convective mode, in the future emphasis should be put even more to the ingredients of severe convective weather. Therefore we suggest adding the following parameters to the ENS production

- **Ingredient based approach**
    - Low-level moisture: 0-500m average mixing ratio of water vapour [g/kg]
    - Instability: 850hPa to 500hPa temperature lapse rate [C/km]
    - ”Fourth ingredient”, wind shear: (Maximum) Effective Wind shear (takes into account the depth of the cloud)
- **Potential for heavy rains** *(depends also on convective mode, namely deep layer wind shear)*
    - Vertical integrated water vapor
- **Movement of convective features**
    - Mean wind vector from LCL à 0.6*EL good approximation of convective cloud movement

# 3. Verification of products

## 3.1 Objective verification

### *3.1.1 Direct ECMWF model output (both HRES and ENS)*

3.1.1.1 Some examples of HRES challenges in Finland
In the pictures below there are some examples of ECMWF forecasting challenges for the winter season December 2017-February 2018

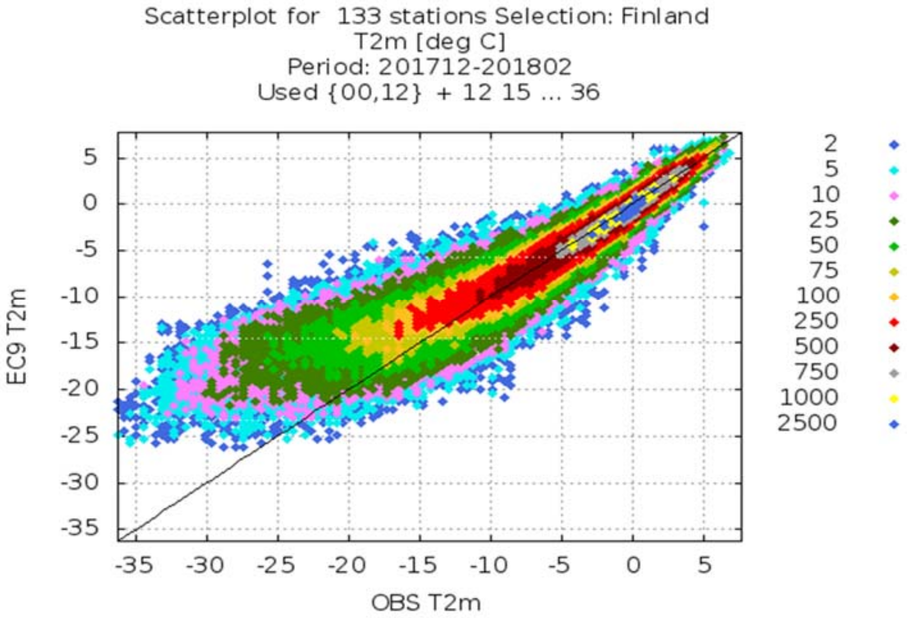

*Figure 3: Cold temperatures forecasted too high*

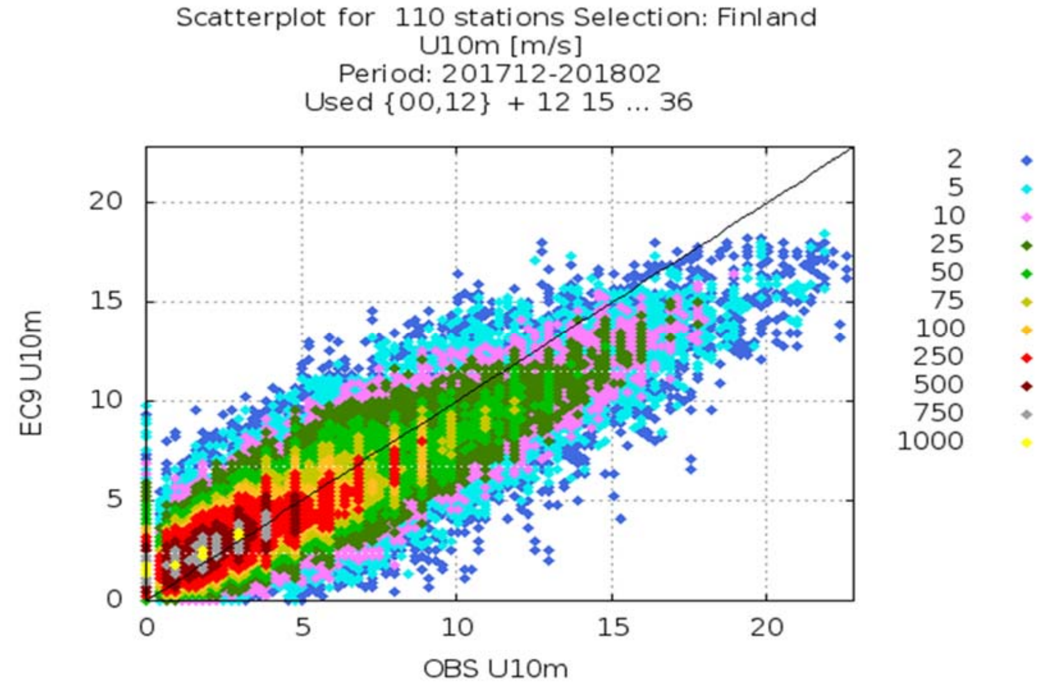

*Figure 4: High wind speeds are continuously underestimated.*

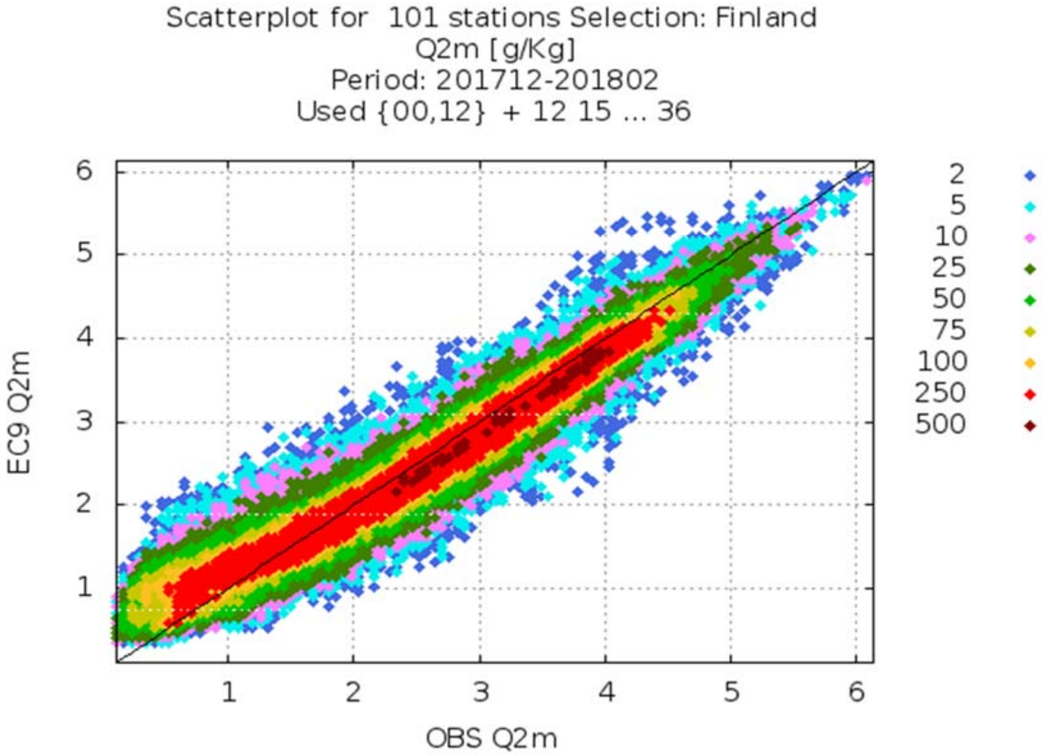

*Figure 5: Too low absolute humidity (despite very low amounts which are overestated)*

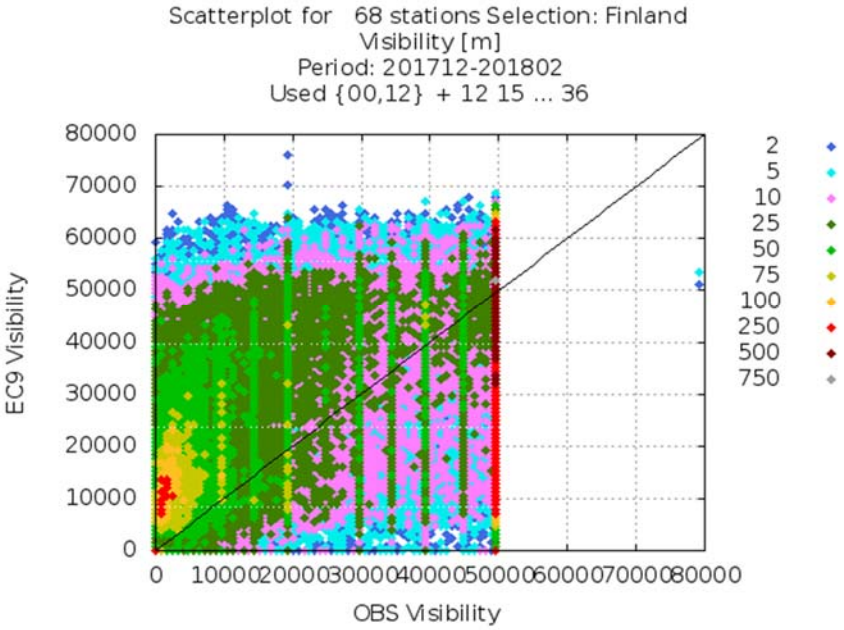

*Figure 6: Visibility is often too good, which goes hand in hand with too dry absolute humidity.*

### 3.1.1.2 First experiences of ECMWF seasonal forecasting system SEAS5

The present ECMWF seasonal forecasting system (SEAS5) became operational in November 2017. From this system, seasonal (three month) temperature anomaly forecasts for Helsinki, southern Finland, were briefly evaluated against observations from Helsinki, Kaisaniemi Park. The forecasts studied were issued between November 2017 and February 2018, for those forecasts the observed three month temperature anomalies were already available. The results are shown in Table 1. It appears that two of the four forecasts were correct, and two forecasts were too warm. February and March 2018 were cold in Finland, and thus the mean three month temperatures of periods JFM and FMA remained below the long-term average (1981-2010). This was not visible in the forecasted three month ensemble means of T2m temperature, there was a warm signal instead.

We have noticed that the previous seasonal system 4 didn't typically forecast a cold signal for three month temperatures in Finland. The present model version (5) has many improvements compared to the previous version (4), for example a better resolution and a prognostic sea-ice model. When creating products the present model version uses climatology from years 1993-2016, which is somewhat warmer than the 1981-2010 climatology, which was used in System 4. Thus, one would expect to see also negative signals in the temperature forecasts, but this was not the case concerning these forecasts studied here.

| Forecast issued | Forecast period | Forecasted T2m anomaly (degC) | Observed T2m anomaly (degC) | Comments |
|---|---|---|---|---|
| November 2017 | DJF 2017/2018 | +1…+2 | +1.4 | A correct forecast |
| December 2017 | JFM 2018 | +0.5…+1 | -0.6 | Forecast too warm |
| January 2018 | FMA 2018 | +1…+2 | -1.2 | Forecast too warm |
| February 2018 | MAM 2018 | +0.5…+1 | +0.9 | A correct forecast |

Table 1. A sample of three month temperature anomaly forecasts (ens. mean) from ECMWF seasonal forecasting system (SEAS5) for Helsinki, southern Finland, compared with observations from Helsinki Kaisaniemi observation site (baseline 1981-2010 climatology when calculating the observed three month anomalies). The forecast data were collected from the ECMWF website.

### 3.1.2 ECMWF model output compared to other NWP models

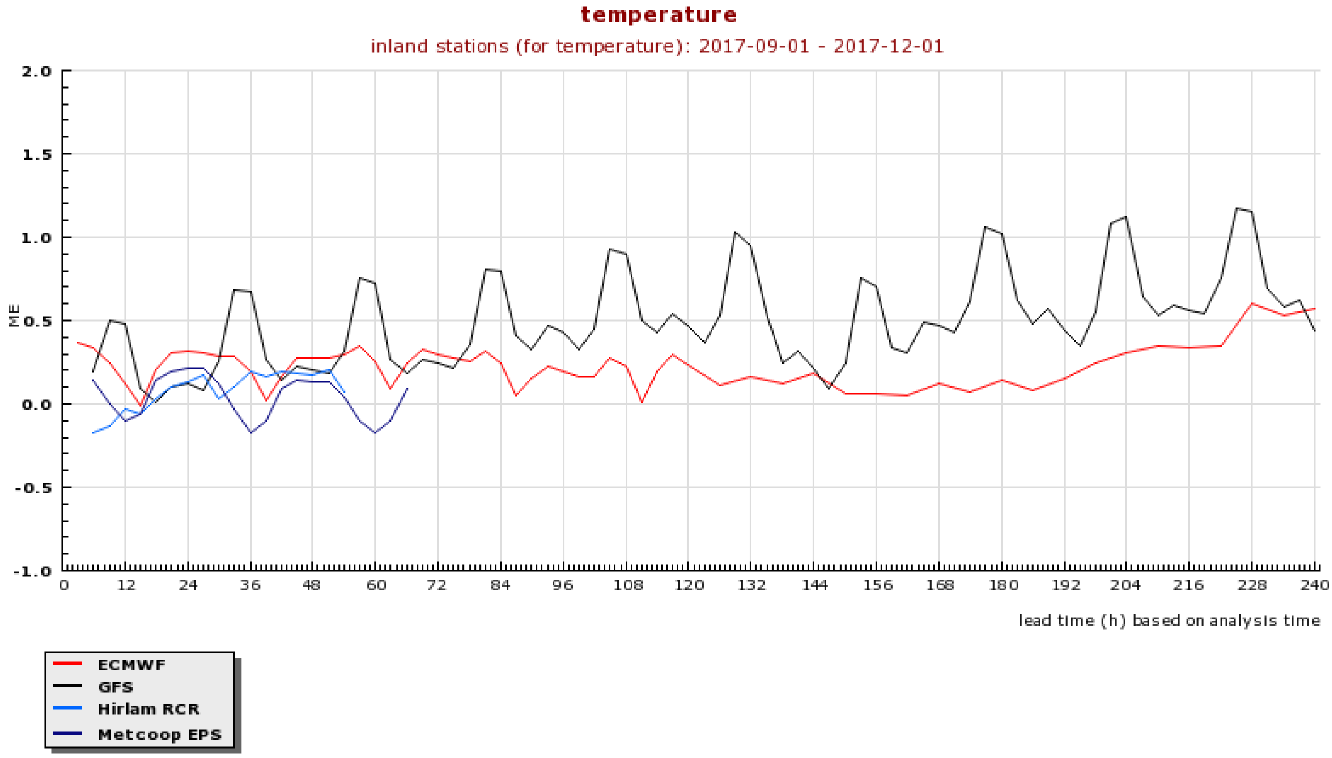

*Figure 7: Mean error (ME) of temperature. Data is from 30 Finnish surface weather stations during 1.9.2017-1.12.2017.*

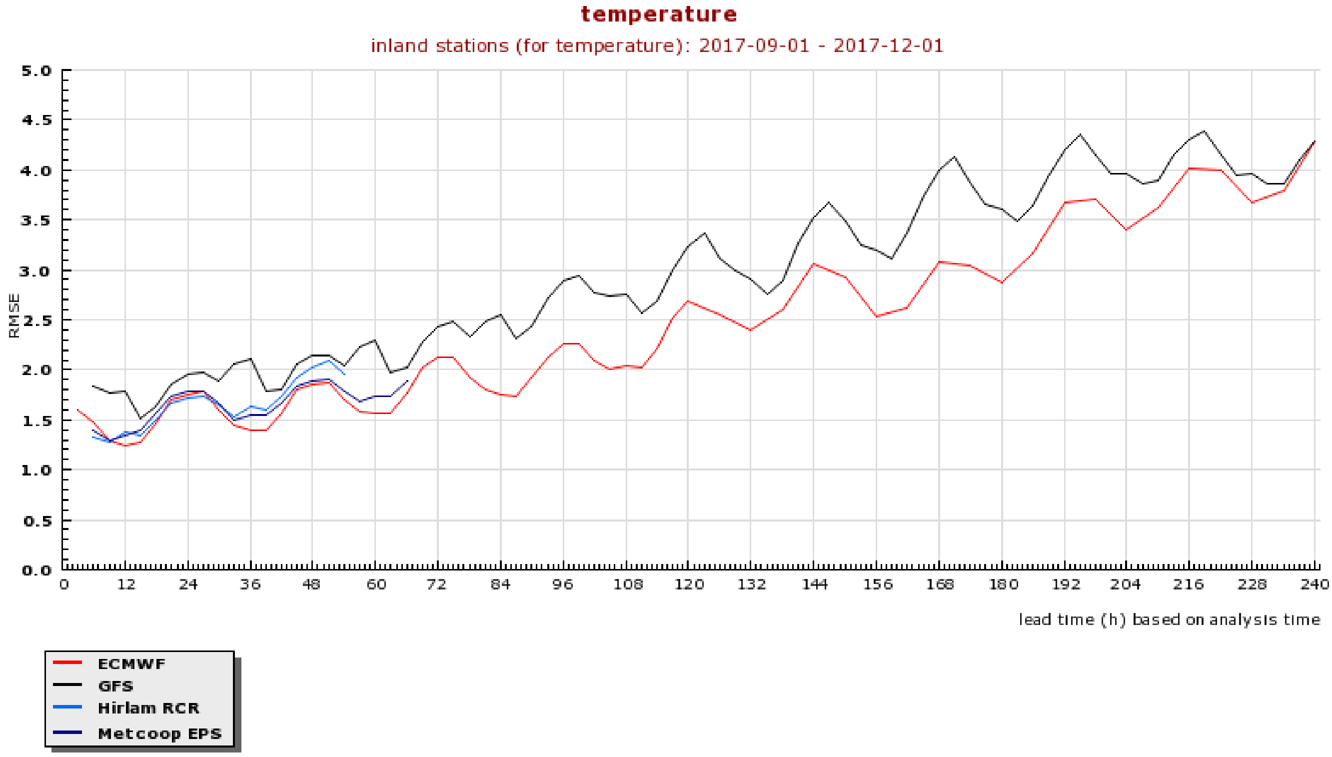

*Figure 8: Root mean square error (RMSE) of temperature. Data is from 30 Finnish surface weather stations during 1.9.2017-1.12.2017.*

During autumn and early winter ECMWF is usually best performing model. Daytime temperatures are generally quite well forecasted, as seen on figures above (time steps: 24, 48 and 72h). Problems are related to night time temperatures, which are usually too warm in ECMWF model, as seen on figures above (time steps: 12, 36 and 60h). Hirlam model and at times also Harmonie (MetCoOp EPS) model are performing better in night time temperatures.

ECMF –model is usually most reliable model when it's cloudy and moist, flow is westerly to southerly and we are experiencing low-pressure type weather. Problems are usually related to clear or mostly clear high-pressure type of weather, during spring and summer, daytime temperature forecasts are too cold. Night time temperature forecasts are too warm almost throughout the year. Wind gust forecasts have improved and are now more reliable. Under larger scale precipitation areas wind gusts are still over-forecasted.

### *3.1.3 Post-processed products*

We have investigated the ECMWF ensemble forecasts calibration on European area (Ylinen et al., 2018). At the moment the calibrated forecasts are tested and used in the EU-project I-REACT. The calibration methods that were used are Gaussian distribution for temperature and Box-Cox t-distribution for wind speed (and wind gust) forecasts. Calibration coefficients are calculated using data from all European stations from past 30 days, and coefficients are updated once a week.

Figures 9 and 10 shows the verification results for 2 meter temperature forecasts from May 2018. It can be seen from figure 9 that calibration corrects the underdispersion of ensemble forecasts at Finnish stations, especially for the lead times up to 180 hours. Also RMSE score is improved after calibration at day times (forecast analysis time was 00UTC). Figure 10 shows the Brier Skill Score (BSS) for probabilistic temperature forecasts with a threshold temperature of 25°C (all European stations included). Sample climatology is used as a reference forecast to calculate the BSS. BSS is better in the calibrated forecasts for all lead times which proves that tested calibration method is capable to improve ensemble forecasts also in the higher part of temperature distribution.

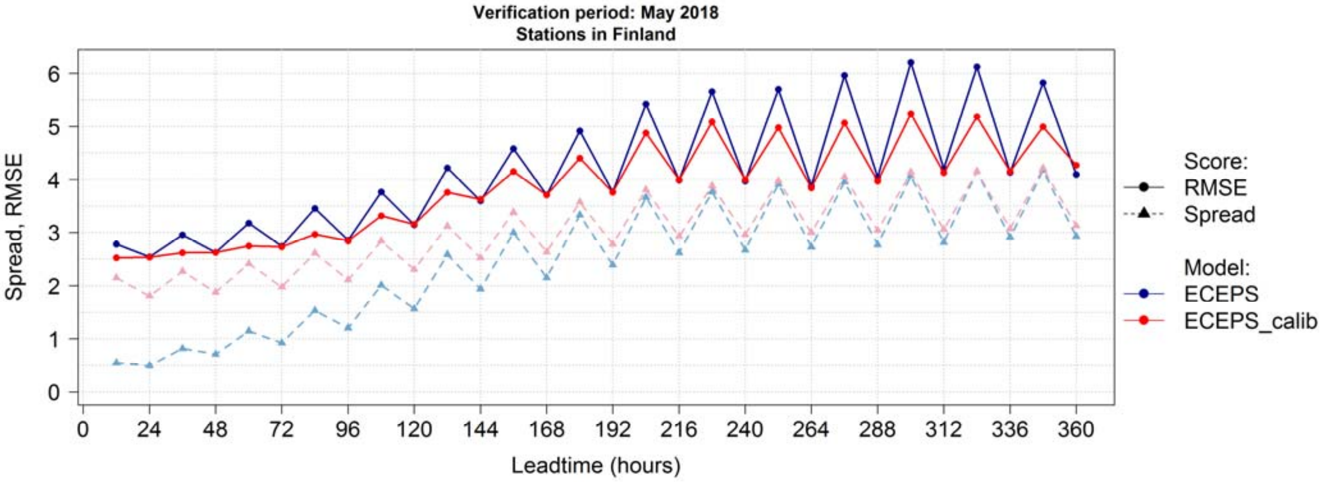

*Figure 9: Spread and skill (RMSE) for temperature forecasts at Finnish stations*

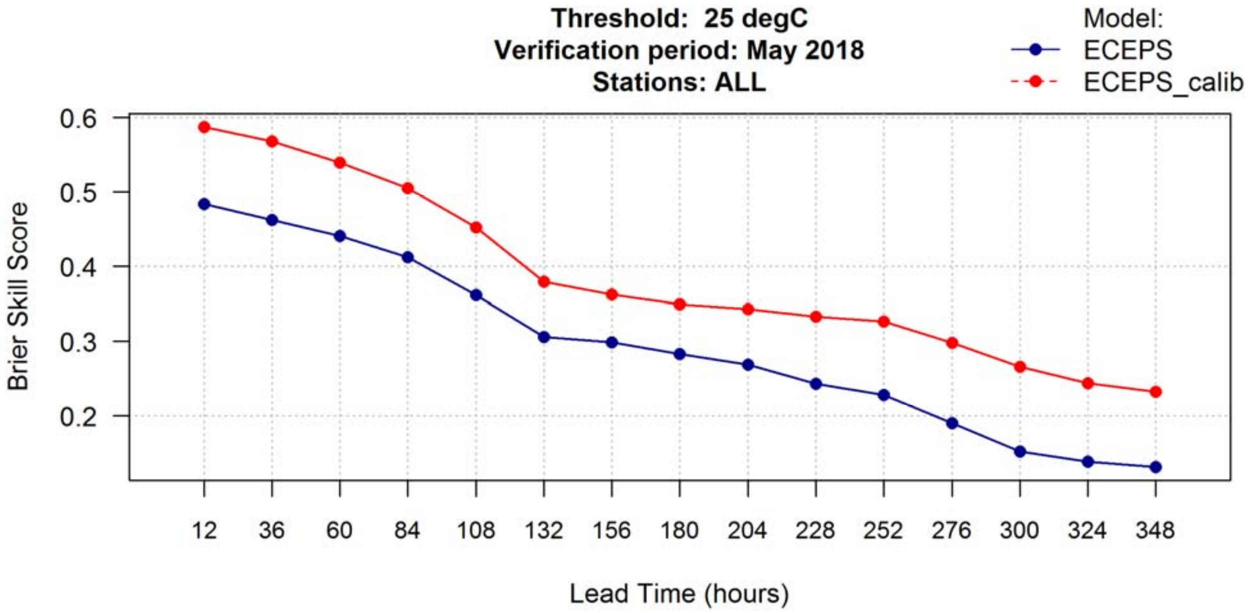

*Figure 10: The Brier Skill Score for the probabilistic temperature forecasts (threshold 25°C) at all European stations*

The intention is to implement HARP calibration to our production during this year and to test calibration first with ECMWF ENS temperatures.

### 3.1.4 End products delivered to users

In the project CLIPS (clips.fmi.fi) funded by Academy of Finland we developed and tested climate impact outlooks for the upcoming six weeks. The focus was to raise awareness of the Finnish public of the risk and benefits related to weather impacts. The new service prototypes were developed using the ECMWF extended range weather forecast data (up to 46 days). During June 2017-May 2018 users were engaged to piloting the operational climate service products with us. The products were tailored separately for each season. Summer season’s weekly forecasts included, e.g., beach and sport weather outlooks, thunderstorm outlook, the probability of a sultry heat spell and, blue algae conditions. Autumn season’s weekly forecasts included, e.g., weekly amount of frost days and road slipperiness outlooks. Winter season’s weekly forecasts included, e.g., heating degree days, weekly amount of snowfall and rainfall days outlooks. Spring season’s weekly forecasts included, e.g., soil frost outlooks, heating degree days outlooks, and amount of frost days outlooks.

The verification of the products has started using ERA-Interim (or ERA5 when available) reanalysis data from ECMWF. Since the services developed in CLIPS were based on ECMWF products, the CLIPS project team from FMI strongly collaborates with ECMWF.

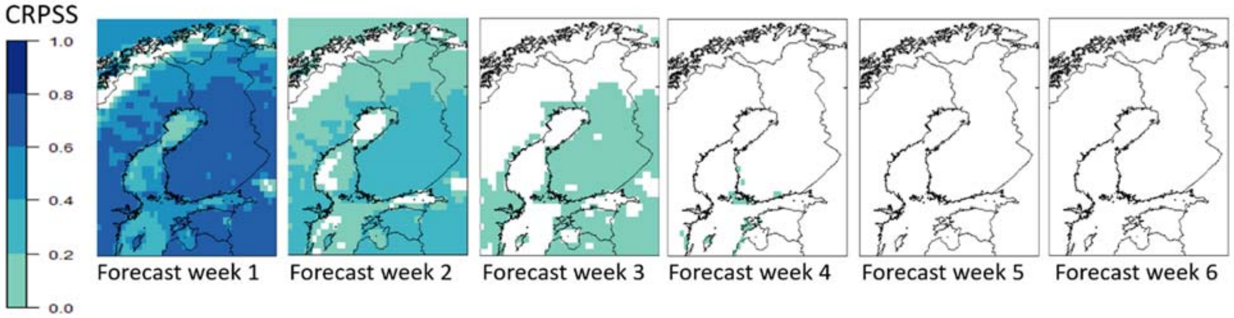

*Figure 11: First preliminary results of the verification of weekly mean temperature (t2m) for each grid point. The CRPSS of the reforecasts for years 1997-2016 using ERA-Interim climatology of 1996-2016 as the reference.*

## 3.2 Subjective verification

### 3.2.1 Subjective scores

We have already summarized the main challenges in the sections above but in addition we have gathered some comments from FMI´s aviation forecasters:

FMI's Rovaniemi's aviation office has been gathering feedback from the aviation forecasters on how well models have predicted the low-level clouds in northern Finland. ECMWF -model performance:

- June 2017: 24 markings (40% of shifts / model runs) and 17 cases good amount of low clouds, 1 case of too much low clouds, and 6 cases too low amount of low clouds.
- July 2017: 25 markings (40,3 %), 19 good, 0 too much, 6 too low.
- August 2017: 19 markings (30,6 %), 16 good, 2 too much, 1 too low.
- September 2017: 30 markings (50 %), 15 good, 5 too much, 10 too low.
- October 2017: 26 markings (41,9 %), 10 good, 6 too much, 10 too low.
- November 2017: 25 markings (41,7%), 10 good, 14 too much, 1 too low.
- December 2017: 19 markings (30,6%), 8 good, 11 too much, 0 too low.
- January 2018: 21 markings (33,9 %), 6 good, 11 too much, 4 too low.
- February 2018: 14 markings (25 %), 4 good, 9 too much, 1 too low.
- March 2018: 3 markings (4,8 %), 0 good, 3 too much, 0 too low.
- April 2018: 14 markings, (23,3%), 9 good, 2 too much, 3 too low.
- May 2018: 3 markings, (4,8 %), 3 good, 0 too much, 0 too low.

There is some "bias" in these markings, because they are more easily done when there is something wrong in the forecast. During good aviation weather forecasters tend to forget to do this reporting. As an overall result during the period of June 2017-May 2018 ECMWF -model had 117 times good amount of low clouds. Compared to other models, Harmonie model had 112 times good amount of low clouds and FMI's Hirlam 136 times good amount of low clouds.

Between November 2017 and early March 2018 ECMWF –model had several times too much low clouds and too much moisture in Northern Finland. This differs from previous years and previous model versions, when ECMWF –model was typically driest model. Model was especially wet during 22.11.2017, when it got stratus cloud over the whole area of Northern Finland, other models were much drier and better. Problematic periods were also 18. – 22.1.2018, and during late February – early March, especially 19.2.2018 and 22.2.2018.

Comments from FMI's Kuopio's aviation office:
ECMW-model had a quite good overall performance during the winter time. Most problems were related to moisture and cloud ceiling during April 2018. Problems began when snow started to melt. Model has been too dry when the snow melts and this has caused too optimistic cloud base and cloudiness forecasts in comparison to other models and observations. Visibility is so difficult parameter to forecast that is not followed so much but changes in visibility are somewhat usable (e.g is it under or over 7km) but e.g. 500m compared to 5km is not trustworthy.

Comments from FMI's Helsinki's aviation office:
During winter time in cold inversions, with very low temperatures, ECMWF –model generates stratus clouds over a large area. Typically weather is clear in these situations. Visibility doesn´t drop low enough when it is snowing. Usually boundary layer is too dry.

Office has been gathering feedback from the aviation forecasters on how well ECMWF-model has predicted the low level clouds and visibility in Finland during 6.3.2018 – 16.5.2018. Total number of days 54, time of the day 4 UTC:

- About one third of the cases the ECMWF-model forecasted low-level cloudiness well (well forecasted coverage & cloud base). One quarter of the cases the ECMWF –model forecasted too much low-level cloudiness including fog. One fifth of the cases the ECMWF-model forecasted too low low-level cloudiness. On some cases the ECMWF –model forecasted well the amount of low clouds, but the cloud base was forecasted either too low or high.
- Visibility was quite equally divided to three classes: well forecasted 16 times, forecasted too low 19 times, over-forecasted 18 times

Some remarks related to results above:

- Time of the day (4 UTC), affects the cloudiness/visibility (morning fogs).
- Only low clouds (especially stratus and fog) are taken into account. E.g. if there is cloud with base at 4000 ft and the model has not forecasted it, the forecast is still ranked well.
- There might be several comments for a one day, e.g. there might be different weather over southern and northern Finland.
- The results are strongly dependent on large-scale weather, on a consecutive days the weather was similar, e.g. too much low clouds on 30.4.-3.5.2018. And well forecasted low clouds during 10.5.-13.5.

## 4. Feedback on ECMWF "forecast user" initiatives

FMI will need to look in more detail to information which is available via Forecaster User Portal and Forecaster User Guide. Those seem to have a lot of useful information.

## 5. References to relevant publications

Ervasti, T., Gregow, H., Vajda, A., Laurila, T. K., and Mäkelä, A.: *Mapping users' expectations regarding extended-range forecasts*, Adv. Sci. Res., 15, 99-106, https://doi.org/10.5194/asr-15-99-2018, 2018.

Ylinen K. and Kilpinen J. (2018). *Calibrating Ensemble Forecasts to Produce More Reliable Probabilistic Extreme Weather Forecasts.* In Kees Boersma and Brian Tomaszewski, eds., Proceedings of the 15th ISCRAM Conference (pp. 1089-1097). Rochester, NY, USA May 2018.