Application and Verification of ECMWF Products 2021

Croatian Meteorological and Hydrological Service, DHMZ

Authors: Tomislav Kozaric, Kristian Horvath, Zoran Vakula, Lovro Kalin, Ivan Guettler, Tatjana Vujnovic and Darijo Brzoja

1. Summary of major highlights

At Croatian Meteorological and Hydrological Service (DHMZ), ECMWF products are considered as the main source in the operational forecast, particularly for the medium-range and long-range. For short-range forecast other models are included also, especially ALADIN operated by our Service. ECMWF model is also widely used as input for running other meteorological and hydrological forecast models, such as ALADIN, RegCM4, MIKE11, WRF and LOTOS-EUROS.

Since 2019 a more comprehensive real-time verification has been performed by the grid-to-point method, where gridded values are interpolated to SYNOP data points and verified. Some results and scores of the verification of 2m-temperature will be shown compared to other models and our forecasters. A verification of precipitation against SYNOP measurements was done for a one year period and some results will be presented in Objective verification part as well. Long-range forecasts are also occasionally verified.

For subjective verification, results of the updated forecasters' survey are presented, raising several comments/issues of the model, particularly its 46r1 and 47r1 cycles available after June 2019. Here are the highlights: improvement in precipitation forecast (unlike temperature), better forecast of temperature inversions with low stratus and fog, very good cloudiness forecast in general, huge popularity of Precipitation Type products among forecasters as well as EFI and 15-days meteograms, more interest in ENS Vertical Profiles product etc.

2. Use and application of products

A variety of ECMWF products is used and as already mentioned, ECMWF products are considered as the main source in the operational forecast, particularly for the medium- and long-range where we benefit from ENS.

2.1 Direct Use of ECMWF Products

Direct use of ECMWF products includes GRIB fields from MARS database disseminated via RMDCN and visualized in house using Metview and our Integrated operational environment for weather analysis and forecast called Visual Weather. However, not much of ENS data are processed by Visual Weather due to big data volumes.

After the installation of Visual Weather which allows us visualizations of almost any kind, EcCharts are less often but still regularly used with focus on ENS products. Mainly we use it for specific tailored prognostic products and access it via Dashboard app. For instance, EcCharts products include set of charts for thunderstorm forecasting (CAPE, different shears, convective indices, lightning, EFI, vertical profiles etc.), set of charts for winter weather forecasting (snow and temperature combinations, precipitation type, precipitation type probabilities, accumulated freezing rain, vertical profiles etc.) and set of charts for fog forecasting (visibility, probability of low visibility, cloud base height, vertical profiles etc.).

A Permalink option is widely used for receiving EFI products and ENS meteograms from the old charts catalogue, which is however, planned to be closed. Similar option is still not available in the new Open charts catalogue.

Some long-range forecasts products are visualized at the Service, but most are now monitored at the ECMWF new web layout.

Open Charts product is used at our Service as well and becomes more and more popular. It is undoubtedly a big step forward in public availability of ECMWF forecast and we welcome this initiative a lot, as well as the open data policy of this kind. Huge variety of charts are available, from medium to long range, these are directly accessed and even interactive to some extent, e.g. meteograms on Click are one nice feature. The new Open charts catalogue itself is very intuitive and easy to search. However, very often when browsing charts, we experience problems with loading of charts and get error message "Sorry, it seems that image could not be loaded" or "We're sorry. Package has not been found".

2.2 Other uses of ECMWF output

2.2.1 Post-processing

In order to provide an unified pointwise direct model prognostic information for public at Service's website, instead showing two different model forecasts for cities, blending of ALADIN 3-day and ECMWF 7-day forecasts is used. Mesoscale model 0-

72h forecasts are merged with global model 75-132h forecasts near the forecast range of the merge (72h). Joint 7-day ALADIN and ECMWF forecast is presented in meteograms and published on https://meteo.hr/prognoze_e.php?section=prognoze_model&param=7d.

2.2.2 Derived fields

As mentioned before Visual Weather allows easy combining of different meteorological parameters into derived fields. Although not many new fields have been derived since the last report, some of them are still worth mentioning: probability of road icing (as a combination of temperature, humidity, precipitation and wind), frost alert (combination of temperature and wind), guidance for fog occurrence (combination of vertical temperature gradient, humidity, surface wind), guidance for deep moist convection (CAPE, CIN, vertical wind shear, sources of lift etc.), heat wave indicator (geopotential heights and temperature). There is also a nice tool in Visual Weather, called Field Diagnostics, which is specially designed to derive new fields out of raw ones and export them as GRIB, e.g. we derive thunderstorm indices, advection and divergence parameters, turbulence, vorticity etc.

2.2.3 Modelling

ECMWF lateral boundary conditions (LBC) are used for running a 72hr forecast with ALADIN-HR4 (4km grid spacing) operational runs. Due to time constraints for availability of ALADIN 72hr forecast ECMWF LBCs in production are used in so-called lagged mode (ECMWF forecast initialized 6h earlier than ALADIN initial time is used as LBC). As no SST data assimilation is performed locally, the initial SST is taken as a copy from global model analysis. Since 6th Feb 2018, hourly LBCs are provided by ECMWF but still in operations 3-hourly coupling is used. Also, from June 2020 model configuration 903 is used for generation of LBC with same horizontal grid spacing as before (possible upgrade should be considered). As response to earthquake that damaged DHMZ headquarters in 2020, backup NWP operational suite was set up on ECMWF’s High-Performance Computing Facility (HPCF). So additionally, one more copy of ECMWF LBC files is regularly being sent to agreed location at ECMWF HPCF where it is used in the backup suite. To provide detailed forecast of height and direction of wind waves, Wind Wave Model (WWM) was set up at DHMZ. WWM is run once per day with boundary data at the Otranto strait obtained from the global WAM model.

First DHMZ-RegCM4 nonhydrostatic simulations performed using the ECMWF HPC for several cases studies of the heavy precipitation events over the Alpine region are published as a part of the Coppola et al. (2020) study. In this reporting period, we completed ERA-Interim forced RegCM4 simulations at the 12.5 km horizontal resolution, and nested 4 km simulation using the non-hydrostatic version of the model in both experiments for the period 1999-2012. The model domain covers greater Alpine region, and modelling efforts are done under the CORDEX FPS Convection protocol. The results of these experiments are the part of the study Ban et al (2021). In 2021 we have initialized experiments using CMIP5 global climate model EC-EARTH as boundary conditions for the 12.5 km RegCM4 setup. By the end of this year, we plan to complete the 5-year period of the historical climate at both 12.5 km and 4 km resolution. Based on the available SBU quota for these modelling activities, additional 5-year period will be covered too.

The Operational hydrological forecasting system within MIKE11 software is based on the real-time data received from available online stations in Croatia and Bosnia and Herzegovina, relevant online data received from Slovenia and temperature and precipitation forecast from the NWP models ALADIN (4 km resolution, lead time 0-71 h, 4 runs per day - 00, 06, 12 and 18 UTC) and ECMWF (10 km resolution, lead time 72-120 h, 2 runs per day - 00, 12 UTC). The hydrological forecasting model is mainly driven by the precipitation input. The model has to deal with uncertainty in rainfall, which is usually the largest source of uncertainty in hydrological modelling.

Air Quality Modelling is currently making use of ECMWF supercomputer resources for running LOTOS-EUROS chemical transport model, for which IFS meteorological fields (F1280 grid) and variables are used as a driver:

- **Operational archive (class: od), Atmospheric model (stream: oper), Forecast (type: fc), Surface: (sfc), 1h-F1280, param:** 'blh' -> boundary layer height, '2t' -> 2 metre temperature, '2d' -> 2 metre dewpoint temperature, '10u' -> 10 metre U wind component, '10v' -> 10 metre V wind component, 'sstk' -> sea surface temperature, 'ci' -> sea ice area fraction, 'swvl1' -> volumetric soil water layer 1, 'swvl2' -> volumetric soil water layer 2, 'swvl3' -> volumetric soil water layer 3, 'swvl4' -> volumetric soil water layer 4, 'sd' -> snow depth, 'ssr' -> surface net solar radiation, 'ssrd' -> surface solar radiation downwards, 'sshf' -> surface sensible heat flux, ‘slhf' -> surface latent heat flux, 'zust' -> friction velocity, 'lsp' -> large scale precipitation, 'cp' -> convective precipitation, 'sf' -> snowfall, 'tcc' -> total cloud cover
- **Model levels: (ml), 3h-F1280, param:** 'lnsp' -> logarithm of surface pressure, 't' -> temperature, 'q' -> specific humidity, 'u' -> U component of wind, 'v' -> V component of wind, 'cc' -> fraction of cloud cover, 'clwc' -> specific cloud liquid water content
- **Fire emissions:** stream=gfas, type=gsd

In addition, IFS is used for WRF meteorological model on fine-scale resolution. Furthermore, emission data from CAMS are used for all numerical air quality models (LOTOS-EUROS, EMEP), and in the near future, CAMS ensemble pollutant concentrations will be used as boundary conditions for CTMs. ERA 5 reanalysis data are used for the development of statistical models and 'training' of air quality models based on neural networks. So far, ECMWF meteorological data are validated implicitly.

3. Verification of ECMWF products

3.1 Objective verification

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

As already introduced in the Summary here we present some results of the real-time verification of 2-m temperature, the results of one year verification of precipitation and the verification in the scope of SEE-MHEWS-A project in which Croatian Meteorological and Hydrological Service takes part.

Real-time 2-m temperature verification

The emphasis of the verification is put on the forecast of minimum and maximum temperature for the next day which is one of the everyday tasks of our forecasters. The min and max temperature values for many places (most matching synoptic stations) are forecast one day in advance (for tomorrow). The main goal of this verification is to evaluate particularly forecasters' performance but also to compare it to available NWP models.

A method used is grid-to-point, where gridded values of direct model outputs are interpolated to SYNOP observation points and verified. Visual Weather system normally interpolates gridded data of the numerical models, so this method was natural and simple choice. Models included in the verification are: ALADIN 4 km (00 UTC run), DWD ICON 6.5 km (00 UTC run) and ECMWF-IFS HRES (00 UTC run, always operational suite), referred as Ala4, DWD and EC, respectively. DHMZ's forecasters are referred as D1. Continuous verification scores were calculated for each synoptic station but also aggregated for all stations (32 stations, including three mountainous) and include Bias, MAE and RMSE.

In the year 2020 aggregated scores for the maximum and minimum 2-m temperature forecast from yesterday (Tmax and Tmin) are shown in Figure 1.

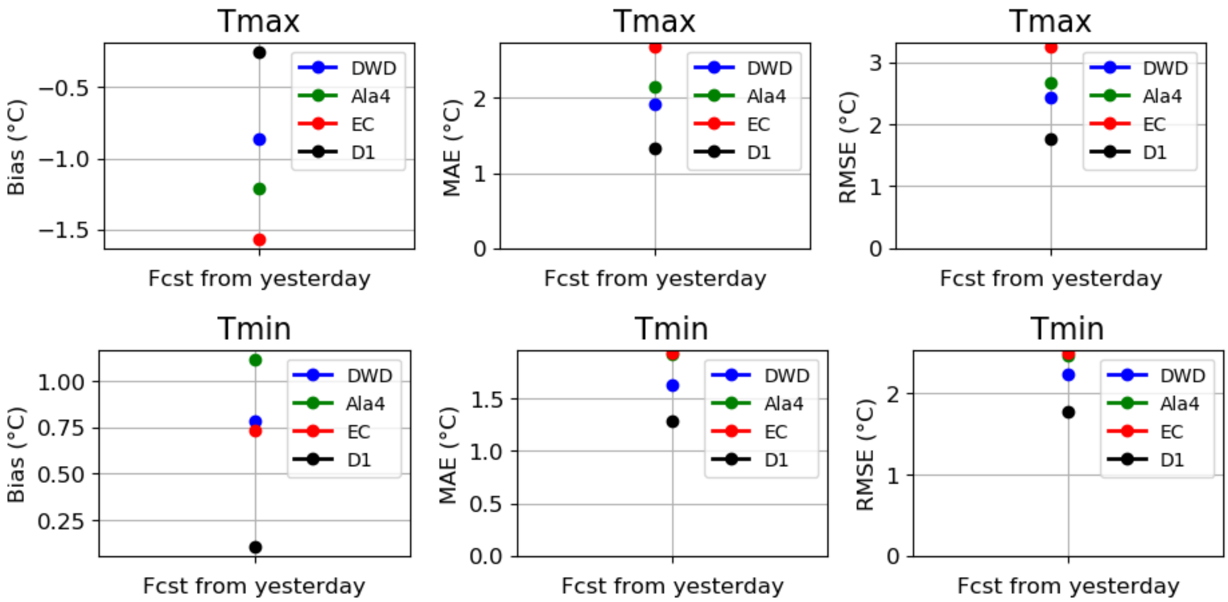

Figure 1. Bias, MAE and RMSE for maximum (up) and minimum (down) 2-m temperature forecast one day in advance for all available synoptic stations in 2020.

Biases for Tmax were overall negative in 2020 but smallest for D1 and biggest for EC (-1.5 °C). As usual, this underestimation came from the stations at the Adriatic Sea (not shown) and was less pronounced in winter. Similarly, MAE was biggest for EC, also RMSE (both around 3 °C) whilst DHMZ's forecasters (D1) produce smallest errors. It is obvious that coarse resolution of ECMWF model, in comparison to other two models, affects Tmax forecast the most near the sea. However, Tmin Bias of the ECMWF forecast was slightly positive in 2020, as was the case for all models compared to forecasters. Mean absolute error was similar for all forecasts, around 1.5 °C, also was RMSE but around 2 °C. ECMWF usually a bit overestimates minimum temperature on land, while bias near sea is mixed (not shown).

As already mentioned, RMSE of the ECMWF's Tmax forecast was biggest among all models. The error was larger for stations at Adriatic Sea or near it as well as for mountainous stations (Figure 2). Some stations at the seashore and on small islands which are not distinguished by the model (i.e. model point is located in the sea) show RMSE of 4 °C or more. A simple post-processing method like the bias removal could make these EC temperature forecasts usable (something similar we have developed in our Visual Weather system).

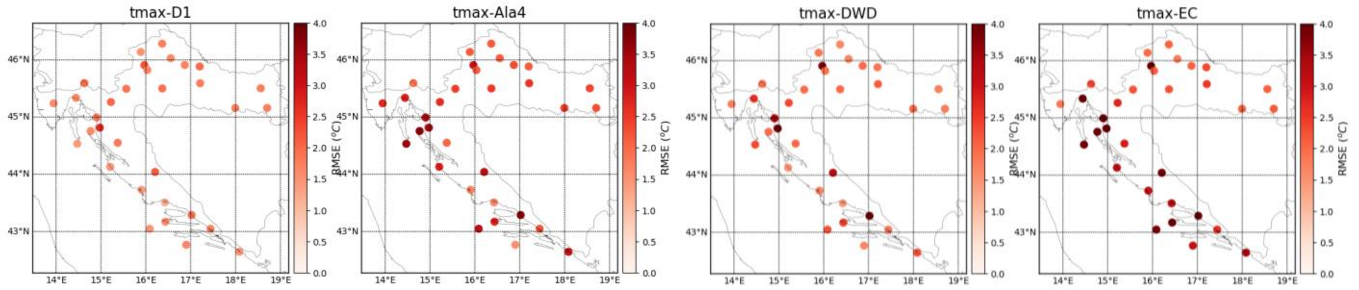

Figure 2. Spatial maps of different models RMSE of the maximum 2-m temperature forecast one day in advance in 2020. From left to right: D1, Ala4, DWD and EC.

Another interesting example of the verification of maximum and minimum temperature forecast is presented in the Figure 3. If focus is put just on ECMWF forecast, it is clearly visible that model in general underestimated Tmax (for one day in advance) during the last two years, particularly in summer months when differences between model and observations are clearly biggest. For Tmin, ECMWF forecast is better with smaller positive bias and smaller error throughout the year.

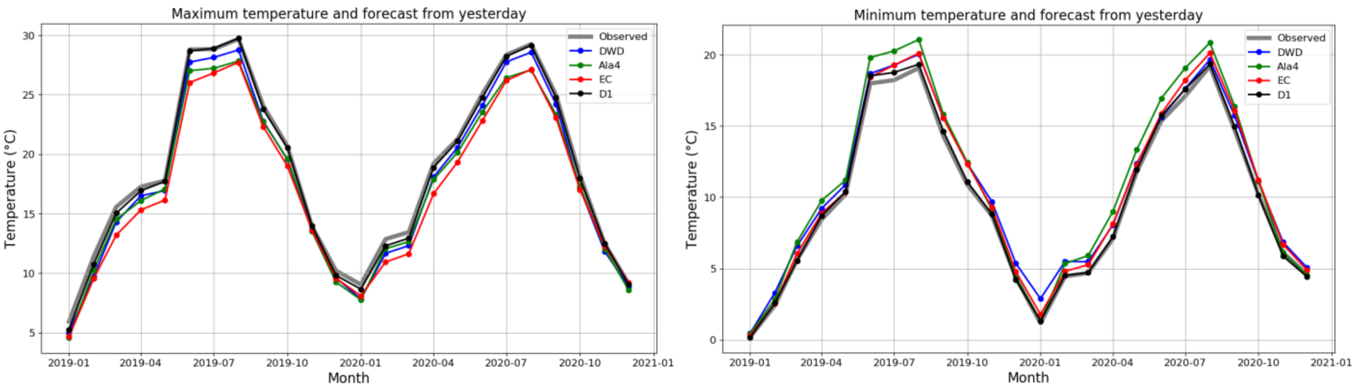

Figure 3. Observed and forecast mean monthly maximum (left) and minimum (right) 2-m temperatures for all available synoptic stations in 2019 and 2020.

Categorical verification of 24-hourly precipitation 54 hours ("two" days) in advance

A method used here was also grid-to-point, where gridded values of direct model outputs are interpolated to location of observations and verified. Models included in the verification are: ALADIN 4 km (00 UTC run), DWD ICON 6.5 km (00 UTC run), ECMWF-IFS HRES (00 UTC run, always operational suite) and GFS (00 UTC run) referred as ALA4, DWD, EC00 and GFS, respectively. Categorical verification of 24-h cumulative precipitation was performed using 73 observation-forecast pairs per day in one year period (September 2018 – August 2019) put into a multi-category contingency table for which scores were calculated (Figure 4). Categories were "dry" [0, 0.5 mm/24h>, light [0.5, 10 mm/24h>, moderate [10, 30 mm/24h> and heavy precipitation [30 mm/24, inf>.

The overestimation in all models is obvious in light precipitation category (Bias more than 1) and is more pronounced in the coarse resolution global models, i.e., ECMWF and especially GFS which tend to spread precipitation. "Dry" category is a little bit underestimated so it can be stated that the models are in general too wet, with ALADIN being the least and GFS the most. As "dry" weather in Croatia is climatologically the most frequent it is normal to see higher probabilities of detection or Hit Rates (H, values change depending on a definition of "dry" threshold) and low False Alarm Ratios (FAR) in all model forecasts. However, the forecast of larger amounts of precipitation is more interesting and valuable than this of a "dry". In light and moderate categories H of ECMWF is higher than in other models, particularly the finer resolution ones. Given the equal FAR between models ECMWF can be considered performing the best in this range of precipitation amounts. Finer resolution models, i.e., ALADIN and DWD have higher detection of heavy precipitation but with higher FARs at the same time meaning that locations of heavy precipitation in not always accurately forecast ("double penalty" problem). Proportion of correct forecasts (PC) and skill, both Hansen-Kuiper's Skill Score (KSS) and Heidke Skill Score (HSS), are more or less the same for ECMWF, ALADIN and DWD, presenting values about 0.75 for PC and 0.45-0.5 for skill, while values for GFS model are lower.

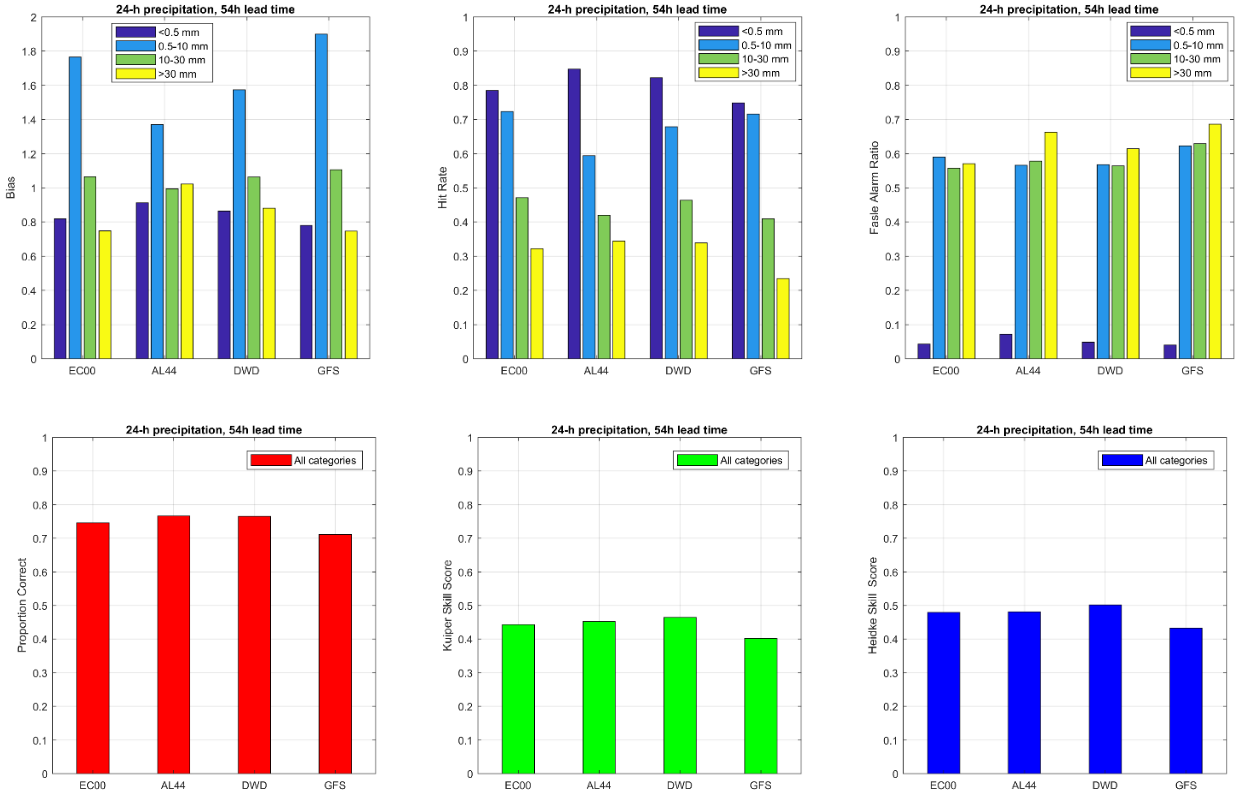

Figure 4. The resulting scores (Bias, Hit Rate, False Alarm Ratio, Proportion Correct) and skill scores (Hansen-Kuiper's Skill Score, Heidke Skill Score) of the verification of 24-h cumulative precipitation for "two" days in advance in the period from September 2018 to August 2019 for all available observing stations.

Verification in the scope of South-East European Multi-Hazard Early Warning Advisory System

The main goal of the South-East European Multi-Hazard Early Warning Advisory System (SEE-MHEWS-A) is to provide support for the National Meteorological and Hydrological Services in Southeast Europe to produce timely and accurate warnings of hazardous weather and hydrological events. To establish a multi-hazard early warning advisory system, it is necessary to perform verification of numerical predictions, which is within SEE-MHEWS-A project evaluated for five numerical weather prediction systems:

- **ALADIN-ALARO** (Aire Limitée Adaptation dynamique Développement InterNational)
- **COSMO** (Consortium for Small-scale Modelling)
- **ECMWF-IFS** (Integrated Forecast System)
- **ICON** (Icosahedral Nonhydrostatic)
- **NMM-B** (Nonhydrostatic Multi-scale Model)

The result of applying the verification methodology to the South-East Europe region is shown in country-specific verification report. The countries participants of the SEE-MHEWS-A project are Albania, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Greece, Hungary, Israel, Jordan, Lebanon, Moldova, Montenegro, North Macedonia, Romania, Serbia, Slovenia, Turkey, and Ukraine. The verification methodology, described in the SEE-MHEWS-A project deliverable (https://confluence.ecmwf.int/display/SEEM/Numerical+Weather+Prediction+Component), is applied on 24-h cumulative precipitation data gathered from 33 observing stations in Croatia from September 22nd to November 22nd 2020. The analysis included missing data and climatology analysis as well as verification of 24-hour precipitation as continuous and categorical predictands and single observation-neighborhood forecast approach (SO-NF).

Continuous verification

All models have a correlation coefficient above 0.50, which means there is a strong linear correlation between each model's forecasts and observations (Table 1). Model ECMWF exhibits the highest (0.74) coefficient value, and COSMO the lowest (0.62). Bias is overall positive for the ALADIN model (0.08 mm/24h), indicating observation overestimation, and negative for the other models.

Table 1. Conventional verification results for ALADIN, COSMO, ECMWF, ICON and NMM-B model for 24-h cumulative precipitation data in the period September 22nd – November 22nd 2020 at 33 observing stations in Croatia. The values significantly different from the ECMWF forecast (0.05 sig. level) are marked with an asterisk.

Model	r	Bias (mm/24h)	MAE (mm/24h)	RMSE (mm/24h)
ALADIN	0.72	0.08	3.71	9.85
COSMO	0.62*	-0.92*	3.86	11.11
ECMWF	0.74	-0.32	3.50	9.44
ICON	0.68*	-0.26	3.71	10.76
NMM-B	0.64*	-1.06*	3.90	10.98

The ECMWF model overestimates the precipitation at the coastal complex terrain, whereas the underestimation can be noticed in the mountainous area and towards more flat continental mainlands (Figure 6). The smallest error is exhibited by the ECMWF model, measured both by MAE (3.50 mm/24h) and RMSE (9.44 mm/24h). Model NMM-B exhibits the largest MAE and model COSMO the largest RMSE. The error is most pronounced in the southwestern part of central Croatia (Figure 5). It increases from the dryer and flatter continental mainlands towards the complex terrain near the sea. There is also conspicuous MAE in the south for all models. The increased error is probably due to the noticeable influence of the relief on the moist air masses coming from the sea.

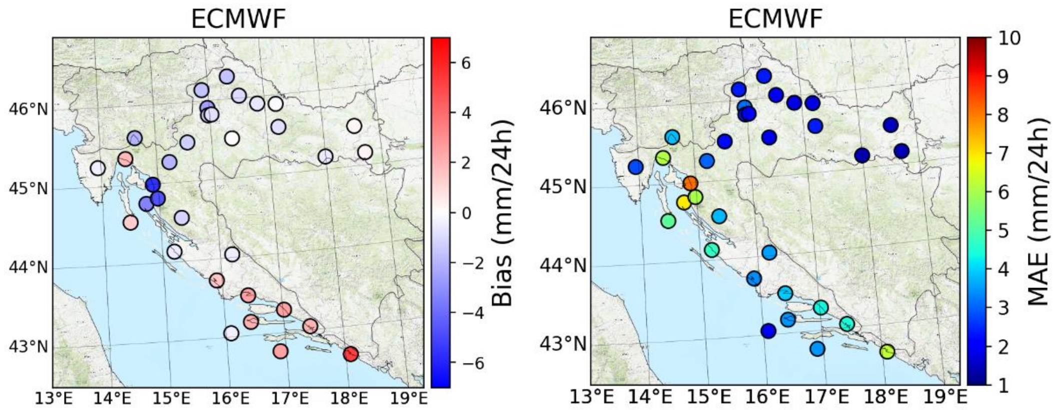

Figure 5. Bias and MAE spatial maps for ECMWF model results for 24-h cumulative precipitation data in the period September 22nd – November 22nd 2020 at 33 observing stations in Croatia.

Categorical verification

The 24-hour precipitation is therefore considered as a categorical variable, using 7 different precipitation categories, defined by threshold values. The dry day is the most frequently observed event, whereas the event frequency reduces towards the higher thresholds (Figure 6). Measured by equitable threat score (ETS), all models produce the best results for the dry day comparing to the other event categories, especially ALADIN. The ETS results lower than 0.20 for [0.3 mm/24h, 1 mm/24h> and [6 mm/24h, 12 mm/24h> precipitation events indicate low predictability of such events and poor quality for all models, whereas the models generally perform better for the higher threshold events. The ETS values generally range from 0.09 up to 0.20 for precipitation events above 60 mm/24h, for which ECMWF performs the best. The latter event assessment is very important since the warning system is a goal of the SEE-MHEWS-A project.

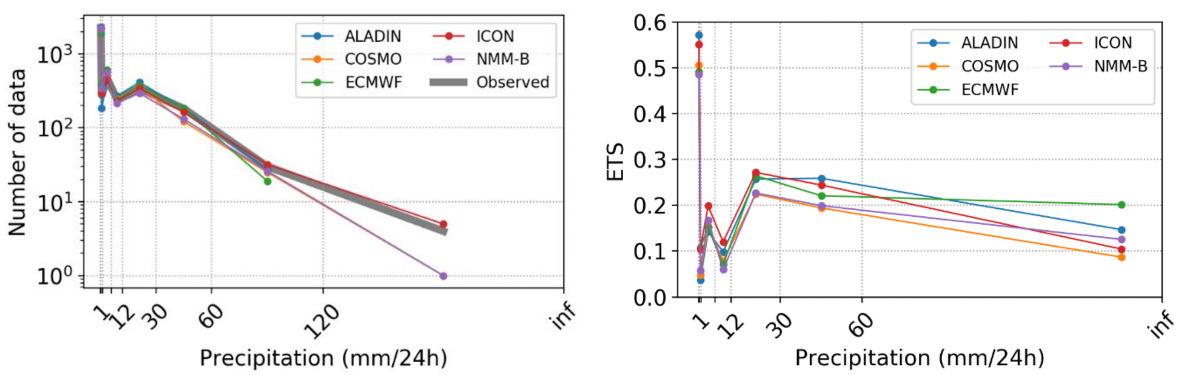

Figure 6. The number of the observed and forecasted 24-h cumulative precipitation data (left) and equitable threat score ETS (right) in the period September 22nd – November 22nd 2020 at 33 observing stations in Croatia. An event is defined as 24-h cumulative precipitation being within two adjacent threshold values presented with the vertical lines on the x-axis.

Single observation-neighborhood forecast approach (SO-NF)

The neighborhood methods overcome the limitations of the gridpoint-based error measures by relaxing the requirement for exact matches between forecasts and observations. The results in Figure 7 show that there are even more useful dry day forecasts in the proximity of the observing stations, improving the ETS value with the neighborhood size and reaching an optimal value at the 27 km scale. Except for a very light precipitation event (< 1mm/24h), a similar improvement can be noticed for the precipitation event forecasts. A better assessment of forecast quality for events above 60 mm/24h shows the benefits of the SO-NF approach in terms of better forecast quality for a rare event. The ECMWF achieves the best performance at the 27 km spatial scale in this category, due to a noticeable reduction of false alarms ratio at larger spatial scales.

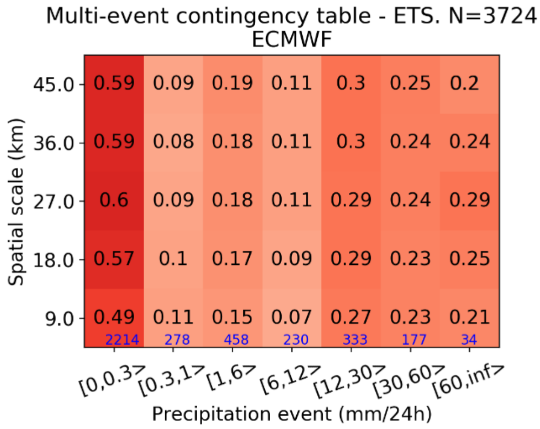

Figure 7. The SO-NF verification result (ETS) for 24-h cumulative precipitation data in the period September 22nd – November 22nd 2020 at 33 observing stations in Croatia. The darker shades indicate good performance, while the lighter shades indicate poor quality.

3.1.2 Post-processed products and end products delivered to users

Analysis of blending of ALADIN 3-day and ECMWF 7-day forecast was performed in order to assess potential inconsistencies when merging mesoscale model 0-72h forecasts with global model 75-132h forecasts near the forecast range of the merge (72h). Tests included several combinations of blending and covered both temperature and temperature tendency which is more sensitive to merging. Comparison was performed for 30 meteorological stations in Croatia. The period of analysis was from Nov 2017 to Oct 2018, and model data was extracted with 3-hourly frequency. The analysis included both models forecast 1-year climatology and verification scores. The analysis suggests that merge does not result in strong discontinuities in forecast climatology or scores (Figure 8). Larger and notable differences, however, can be expected in regions where there is a difference in land-sea mask or orographic heights, but then differences between models are not specific to the forecast range of merge, but are found for model forecasts as a whole.

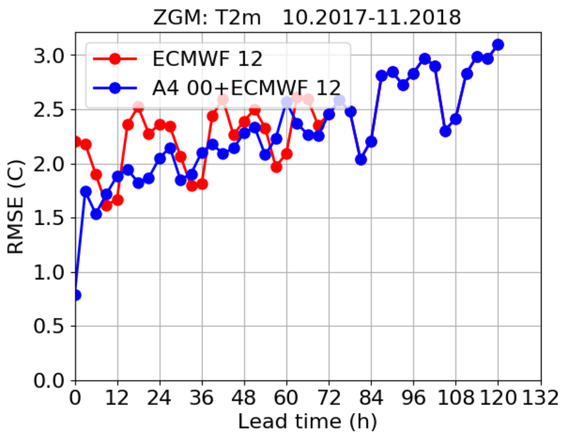

Figure 8. Comparison of RMSE for temperature for ALADIN forecast at 00 UTC and ECMWF forecast at 12 UTC of the preceding day, simulating available forecast data for joint merged forecast after the ALADIN forecast at 00 UTC is available.

3.1.3 Monthly and Seasonal forecasts

Monthly forecast temperature verification

The changes of the forecast system made in March 2016 mostly reduced errors and the skill score of the average weekly temperature forecast (mwt) was improved more than anomaly (wta) from the average (Figure 9). Reference forecast is climatology 1981-2010. A negative skill that existed for weeks 3 and 4 has disappeared. The model underestimates the measured temperature, mean forecast errors are negative. Model runs on Mondays are better than Thursdays, especially for weeks 1 and 2 and it is better to monitor the forecast of the weekly anomaly from the average weekly temperature than the forecast of the average weekly temperature itself. Since March 2016, compared to before, the frequency of forecasts that differ more than 2 °C for 1 week in advance has significantly decreased (from 16% to 8%). For 2 and 4 weeks in advance the frequency has remained around 20% and 1%, respectively. Interestingly, for the 3rd week this has even increased from 2% to 7%, revealing a decrease in model performance. Since 2016, for consecutive Monday runs frequency of weekly anomalies greater than 2 ° C has decreased from 29% to 22% when the 2nd week forecast becomes 1st week; for consecutive Thursday runs decrease was from 37% to 33%. There has been no major change when 4th week becomes 3rd week for both runs, as well as no change in frequency of greater anomalies when 3rd week becomes 2nd week for Thursday runs; however, for Monday runs this has increased from 24% to 28%.

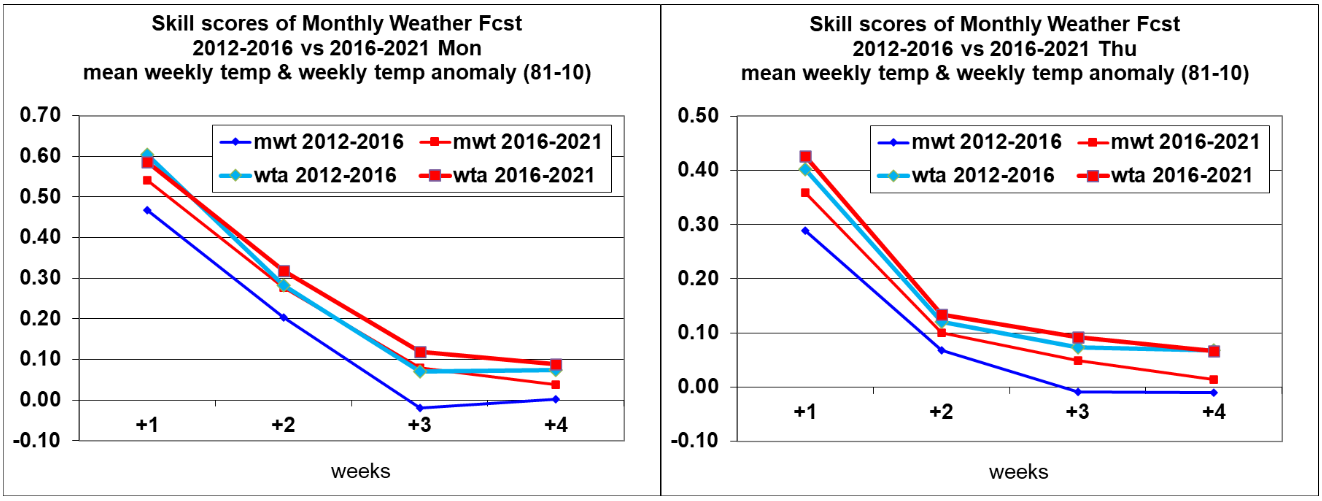

Figure 9. Skill scores of the monthly weather forecast before March 2016 and after, for Monday runs (left) and Thursday runs (right). Skill scores are calculated for Zagreb, the Capital, representing continental part of Croatia.

Seasonal forecast temperature verification

The results by months are relatively poor - the mean absolute error in system SEAS5 is higher than in system SEAS4. However, it must be noted that the data sets are not the same (the sample is much smaller in SEAS5) and the monthly temperatures after the introduction of SEAS5 have become more extreme compared to climate due to climate change (0.4 °C warmer). The monthly skill (reference forecast is climatology 1981-2010) of SEAS5 compared to SEAS4 (Figure 10, left) is better for all 6 months in advance, except for 2nd and 6th month, with even a significant improvement in skill for the 4th and especially 5th month. The results by seasons are however more promising – the mean absolute forecast error (MAE) of the anomaly from the seasonal average is less in comparison to months and skill scores are better. Of course, there are also higher MAEs in SEAS5 than in SEAS4, due to the already mentioned more extreme months and a smaller data set. But skills are better in SEAS5 (Figure 9, right), particularly for the period from 3rd to 5th month in advance. As before, the forecasts in relation to the ensemble MEAN are mostly a little better than in relation to the ensemble MEDIAN, so use of the forecast anomalies from the seasonal mean is recommended for Croatia.

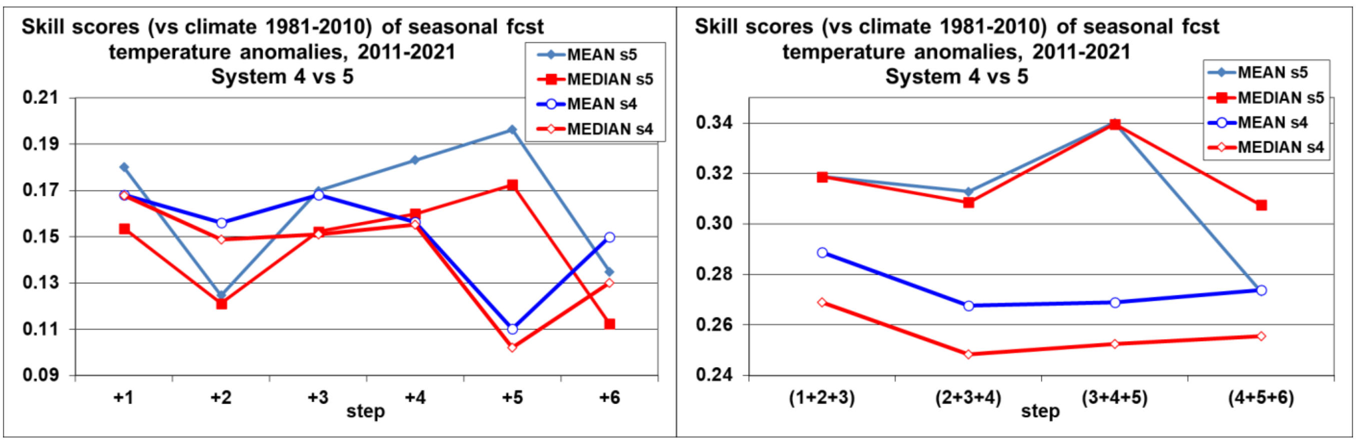

Figure 10. Skill scores of seasonal forecast systems SEAS4 and SEAS5 by months (left) and seasons (right). Skill scores are calculated for Zagreb, the Capital, representing continental part of Croatia.

3.2 Subjective verification

3.2.1 Subjective scores

For subjective evaluation purpose of ECMWF model fields and products in Croatia, a survey is being carried out yearly in the Forecasting Department. Regular topics, also new one, are proposed to the forecasters to confirm/reject and others can join the discussion/voting. Of particular interest were changes brought by IFS cycles 46r1 and 47r1 available after June 2019. Twenty-eight forecasters participated in the survey this year.

The results of the subjective evaluation of ECMWF model fields and products:

- Contrary to results presented in the last report, two thirds of forecasters now think that HRES precipitation forecast has improved in the last two years. Other are not sure about that fact and some of them state that it still spreads out the convective precipitation in warm part of the year with convective processes in the model starting much earlier in the day. It is interesting there is no single negative feedback on the improvements in HRES precipitation forecast. Still, many agree that HRES model occasionally forecasts extreme precipitation by the end of forecasting period but have improved a bit. Point rainfall probability product which is supposed to be more accurate in heavy precipitation than total precipitation probability is unfortunately not monitored much.

- On the other hand, about two thirds of forecasters disagree that HRES 2-m temperature forecast has improved in the last two years. Other are not sure, while there is no positive feedback on the improvements in HRES temperature forecast. According to this results, subjective impression among forecasters is that HRES temperature forecast in Croatia did not change much in the last two years or could be even worse. Exact values might not be accurate enough, particularly near the sea, however daily trends are fine and useful. There are still forecasters agreed with the negative bias in minimum temperature in case of clear nights with snow cover on the ground; however, many of them didn’t vote because there were just a few days with snow and snow cover in the last two winters and springs.

- All agree that HRES is very good in differentiating between high, middle and low cloud cover and in cloudiness forecasting in general. Accordingly, in most forecaster’s opinion HRES is reliable in depicting low stratus cloud or fog during temperature inversion situations both on land and sea, even more reliable than some finer resolution models. However, there are some complaints about daily duration and spatial distribution of low stratus and fog, given stratus and fog disappear too early, particularly in shallow cases. A sort of a complement to cloudiness, new product, sunshine duration in the last 24 hours seems to be followed very rarely and has not arisen much interest between forecasters.

- Interestingly, about improvements in HRES wind forecast two thirds of forecasters are neutral i.e., could not agree or disagree (others are mixed). Maybe the reason lies in the fact that for wind forecasters rely more on models with fine resolution, particularly in short-range, and therefore could not provide a concrete answer to the question.

- Latest survey results show opinion that CAPE forecast becomes more realistic, even over the sea, but there are still those who find it unrealistic. In the survey before, there was criticism against CIN calculation, but now, although there are no evident critics, only few can state that CIN forecast has improved in the last two years. CAPESHEAR parameter is still not used a lot and just half of the forecasters use it for convection forecast, especially type of organised convection.

- About two thirds of the forecasters now report less “jumpiness” of HRES forecasts in longer lead times than it used to be. But in cases when HRES is “jumpy”, “jumpiness” in often evident in ENS forecasts also, phenomenon that forecaster dislike in general. Although, not all of the forecasters are experts in monthly and seasonal forecasts, those who are point out that monthly forecast is now even more “jumpy” (tend to change more and follow the actual anomaly). Seasonal forecast still has large errors in monthly anomalies but most forecasters do not monitor seasonal forecast regularly.

- Although, there is a big praise among forecasters about EFI products for precipitation, wind and temperature (almost all agree EFI/CDF are good and reliable indicator of weather extremes), there is just some positive feedback on reliability of CAPE/CAPESHEAR EFI products. Still, most forecasters seem not to use the latter two very often.

- HRES visibility product became more popular between forecasters as a tool for fog forecasting. It has proved itself as relatively reliable indicator of fog development over some broader area even for longer lead times, although for point forecast can be misleading. In that sense it is useful in general or public weather forecasts and not very useful in airport forecasts for instance. Some aviation weather forecaster claim probability of visibility (e.g. <1.5 km) is more useful.

- Lightning density product from HRES is accepted well by forecasters who state its reliability in potential for convection development, but opinion is still divided as some state it could be misleading in case of weak signals. It is most useful when used together with other convection parameters.

- Forecaster's interest in ENS vertical profiles product has grown a lot since the last report in 2019. It is now regularly monitored by many forecasters who state its great value in many situations, e.g. fog, severe weather, winter precipitation type.
- ECMWF's 15-days extended ENS meteograms, particularly those with climatology included, seem to be a great success, almost all forecasters regularly use them as they are now easily available via permalink.
- Despite having mild winters in Croatia in the last years, there were several cases of dangerous freezing rain, allowing forecasters to evaluate freezing rain products - a feedback is rather positive. Probability of freezing rain is quite useful and showed itself satisfactorily accurate even days ahead. Accumulated rain as well, giving forecasters idea of possible severity of freezing rain when issuing warnings. But accumulations in some situations were significantly overestimated (see Section 3.2.2).
- Although the products of the wave model have proven useful in the Adriatic Sea, particularly the significant wave height, it is difficult to assess how reliable it is since there are almost no measurements at the open sea. Along the coast, the model often overestimates the height of the waves.
- Extended range forecasting products: Early warning for cold spells over Europe, Stratospheric Sudden Warming and Weather regimes extended range forecasts, are inspected regularly by our monthly forecast experts who like those a lot and find them very useful and reliable.
- Gold medal of subjective forecaster's evaluation again goes to Precipitation Type product with hundred percent agreement for its reliability, both HRES and probabilities from ENS. Of course, it is reliable as much as precipitation forecast itself allows, but its huge everyday value in the cold part of the year cannot be disputed. Forecaster like these products a lot.

3.2.2 Case studies

As it can be concluded from objective and subjective verification, ECMWF is generally very good and reliable model for Croatian area and with no doubt the most popular among forecasters. There were rarely severe weather situations it missed or in which it had big errors, so here we list also non-severe cases, not strictly linked to short and medium-range forecasts:

- In the period from 15th to 18th March 2021 HRES had some serious problems in low cloudiness forecast. Widespread cumulus clouds streets in the northern Croatia and Slovenia were not forecast as reported by aviation meteorologists.
- Seasonal forecast missed the Spring 2021 anomaly. There was a rather big confidence in seasonal forecasts for positive anomaly, however in turned out that the Spring was colder than normal.
- On the 27th of February 2021 widespread and dense fog event in the southern Adriatic was completely missed by the model (both HRES and ENS). The event affected air and marine traffic a lot, as fog is not very frequent there.
- There was a big and sudden change from warm to very cold in medium-range forecasts in the beginning of February 2021; for the same reason, a big error occurred in the monthly forecasts (older runs).
- Although freezing rain event on the 3rd December 2020 was captured well in advance, there was relatively large error in the intensity of the event (even for short lead times). A model forecast much larger area affected by freezing rain with much bigger accumulations in some areas. At the end, the impact of the event was minimal.
- Major heavy rain event in norther-western part of the country in the evening of the 24th July 2020 (e.g. Zagreb, the Capital, was heavily flooded) was forecast very well by the model, but locations of the precipitation maxima were not precise enough and missed bigger cities like Zagreb and Varazdin. Nevertheless, the forecast of that event can be considered as a hit.

4. Requests for additional output

As we are Mediterranean country very prone to wildland fires and wildland-urban interface fires, it is a huge interest in ECMWF's fire weather forecasting products. Unfortunately, most of them are under licence and thus not available, even not in the EcCharts. We are particularly interested in Fire Weather Index HRES (daily and max. hourly) and ENS (anomaly, ranking, EFI), preferably as GRIB data. A Permalink to new Open charts catalogue is welcome as well.

5. References to relevant publications

Ban, N., Caillaud, C., Coppola, E. *et al.* The first multi-model ensemble of regional climate simulations at kilometre-scale resolution, part I: evaluation of precipitation. *Clim Dyn* (2021), https://doi.org/10.1007/s00382-021-05708-w

Coppola, E., Sobolowski, S., Pichelli, E. *et al.* A first-of-its-kind multi-model convection permitting ensemble for investigating convective phenomena over Europe and the Mediterranean. *Clim Dyn* **55,** 3–34 (2020)