Common Verification Activity during JJA 2018 - MAM 2019: Main Results

Kirsanov A., Gofa F., Boucouvala D., Batignani F., Bogdan M., Linkowska J., Pflüger U., Tesini M.S., and A. Bundel

1 Introduction

As decided during COSMO General Meeting in Lugano 2012, the performance of COSMO models is demonstrated through the Common Plot activity. Verification results of statistical indices for main weather parameters are derived using the operational COSMO model implementations in each service. The domain, resolution, statistical scores/methods, frequency and graphical representation, are decided on an annual basis from WG5 and listed in the guidelines document (http://www.cosmo-model.org/view/repository/wg5/commonPlots/reports).

The main findings of this organized analysis are presented during the GM plenary session together with the long-term trend of the scores, providing a basis to track the performance of COSMO model as well its systematic errors. As a Common Verification Software (CVS) is used by all services, this allows for a homogeneous, standardized and objective way to apply, calculate and present the verification scores.

The common geographical areas for the coarser and the higher resolution models that are used in this analysis are shown in Figure 1. ECMWF (IFS), ICON-EU and ICON results are also included and compared to COSMO models. The common area includes part of Northern Italy, Austria, Slovenia, Croatia, Germany, Bosnia and Herzegovina, Hungary, Slovakia, and Check Republic.

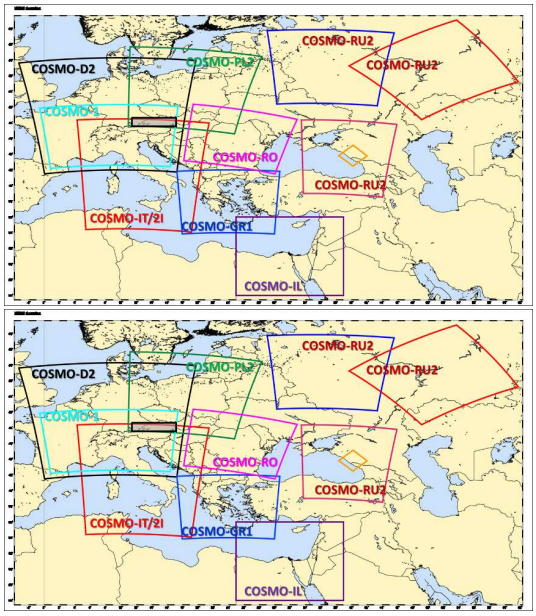

Figure 1: Common Area 1 (upper) and 2 (lower) domains as located within each country simulation area.

doi:10.5676/dwd_pub/nwv/cosmo-nl_20_03

While the transition from COSMO to ICON model is ongoing, the focus of Common Verification activity for the period Summer 2018 – Spring 2019 was also on the comparison of COSMO and ICON models as this would provide a perspective on the existing and the resolved biases that models exhibit.

2 Methodology and Data

The models participating in verification of Common Area 1 (coarse resolution) are COSMO-GR4, COSMO-5M, ICON-EU, COSMO-ME, COSMO-PL, COSMO-RU7 and the global models IFS (ECMWF), ICON(DWD). The fine resolution models participating in verification of Common Area 2 are COSMO-D2, COSMO-IT, COSMO-P. Comparison with the larger scale models IFS, ICON in this domain is also performed. The forecast parameters were interpolated to the observation point by using the 3D method height optimized, except for TCC (15km radius method) and Precipitation (8km radius method). The forecast quality was estimated by applying continuous and dichotomic scores on a 3-hourly time step, depending on the parameter type as follows:

ME and RMSE scores were calculated for Continuous parameters: T2m (2m Temperature), MSLP (Mean Sea Level Pressure), Td (dew point Temperature), WS (Wind Speed), TCC (Total Cloud Cover).

FBI, ETS scores were calculated for Dichotomic parameters: 6h cumulated Precipitation, main thresholds: 0.2, 1, 5, 10, 15, 20 mm/6h.

Dichotomic scores FBI, ETS, CSI were also applied to continuous parameters TCC and Wind Gust by setting intervals. Specifically, for TCC intervals are set to [0,25], (25, 75), [75,100], and for Wind Gust > 12.5, >15, >20 m/sec.

For more information about scores analysis, see the interactive plots and final report for 2018-2019 and previous years on [1].

3 Results

Detailed analysis on the models performance and comparison with previous years can be found in [1]. Only the main results for Common Area 1 are presented in this short report.

2m Temperature: Despite the differences among models, the common feature is that T diurnal cycle amplitude is underestimated, similarly to the previous years. In JJA (Figure 2a) and SON, the bias is positive in the daytime (overestimation), and slightly negative (underestimation) or close to zero during night hours. In DJF (Figure 3b) and MAM, the temperature is underestimated in the daytime, while the values are close to zero at night in DJF and slightly overestimated at night in MAM. The RMSE also exhibits a diurnal cycle with maxima depending on the season and following the absolute maxima of the ME cycle. ICON and ICON-EU perform relatively better compared to COSMO models: The bias cycle is weaker with values closer to zero. RMSE values are also lower (Figure 3).

It is worth noting, that there are some differences in RMSE of ICON model diurnal cycle between JJA and DJF. During the summer, ICON and ICON-EU performance follows COSMO models one exhibiting similar RMSE diurnal cycle, with higher values in the daytime. However, in winter period, ICON and ICON-EU RMSE values exhibit a slight peak during night hours. Regarding the tendency of the scores compared to the previous year, RMSE slightly decreased for JJA, SON and MAM. A small increase was noticed for DJF period though[1].

Dew point temperature at 2 m. Similarly, to the previous years Td exhibits a diurnal cycle in errors, with ME and RMSE peaking in early afternoon.

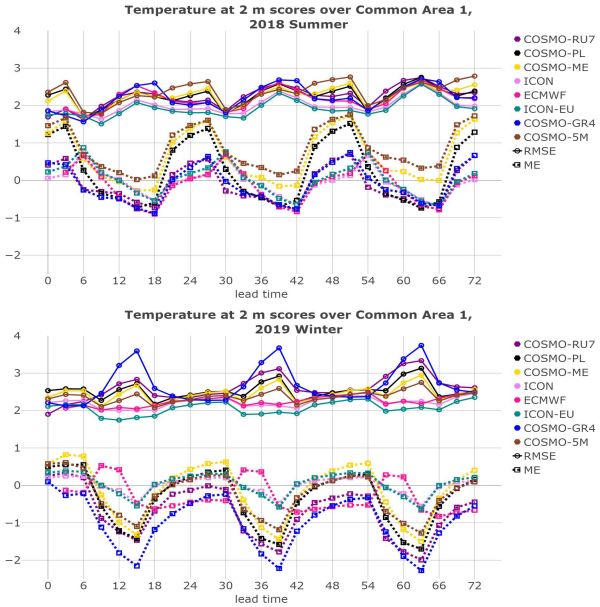

Figure 2: Temperature at 2 m Continuous scores during summer(a) and winter(b)

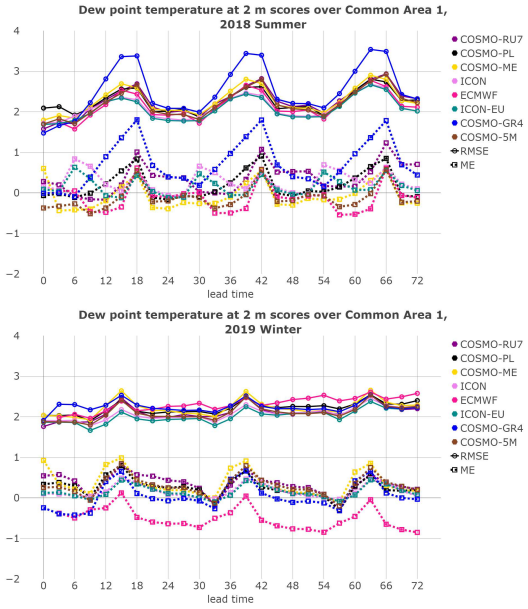

Figure 3: Dew point temperature at 2 m Continuous scores during summer(a) and winter(b)

All models have the tendency to overestimate dew point temperature at 2 m. However, COSMO-PL and COSMO-RU7 underestimate observed values in fall season [1] while IFS underestimates it in winter. ME has negative values increasing with lead time (Figure 3). ICON and ICON-EU scores are better than COSMO for all seasons. Regarding the score tendency compared to the previous year, RMSE slightly decreased for JJA season while no significant changes are shown for MAM. However, a small increase was noticed for DJF and SON seasons [1].
It is worth noting that ICON and ICON-EU elevated ME values found in the year 2016-2017, are no longer present in the two more recent periods.

Mean Sea Level Pressure. The Bias evolution with lead time differs among models and seasons and does not have a very clear diurnal cycle. The only season where a noticeable trend is apparent is SON, that ICON, ICON-EU and IFS underestimate MSLP while all COSMO models tend to overestimate it.

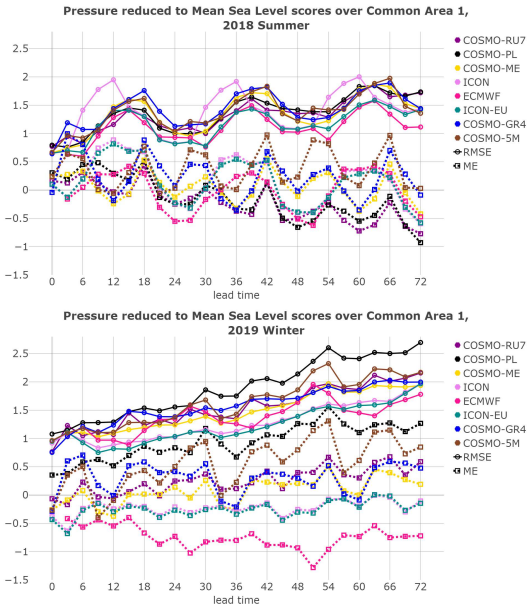

Figure 4: MSLP at 2 m Continuous scores during summer(a) and winter(b)

The RMSE variability is more obvious for this parameter. Similarly to the previous years, RMSE increases with forecast time especially in SON, DJF, MAM [1] and exhibits an afternoon maximum in JJA and in MAM seasons (ICON-EU, ICON and IFS peak only during JJA) (Figure 4). ICON, ICON-EU and IFS RMSE values are lower than COSMO models, especially at higher lead times.
Regarding the RMSE trend compared to the previous years, there was a decrease of RMSE in JJA, SON and DJF for all models. However, a slight RMSE increase is noticed for all models in MAM.

Total Cloud Cover (Note that nighttime observations are limited). ICON and IFS performance is generally better than the one of COSMO models for all seasons especially in JJA, with ME values closer to zero. (Figure 5).

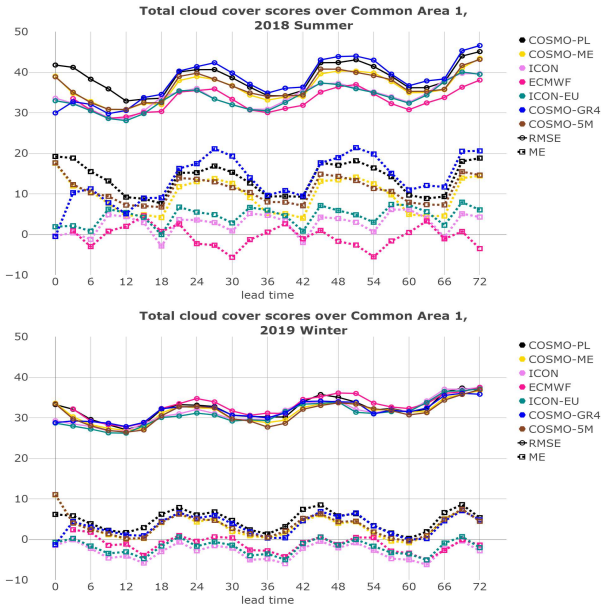

Figure 5: Total Cloud Cover continuous scores during summer and winter

RMSE values for ICON model are slightly lower than COSMO in warm seasons and quite comparable in cold seasons. RMSE for all models is slightly higher in JJA especially during nighttime. Regarding the tendency compared to the previous year, RMSE decreased during MAM for all models, COSMO RMSE increased during JJA and SON (ICON did not change significantly), and there were no significant changes for DJF season.

The dichotomic scores TCC analysis showed that:
TCC<25% (including clear sky) events are generally underforecasted especially at night in JJA when the FBI values are lower.
TCC >75% (including overcast) events are generally overforecasted especially in the summer around noon (FBI >1).
For 25-75% TCC events, the FBI values (mainly<1) indicate underforecasting of events especially for COSMO models. This category is the hardest to forecast especially in the summer, exhibiting the lowest TS (Threat score) [1]. Overall, ICON models perform better than COSMO. In Figure 6, the FBI values for ICON are shown.

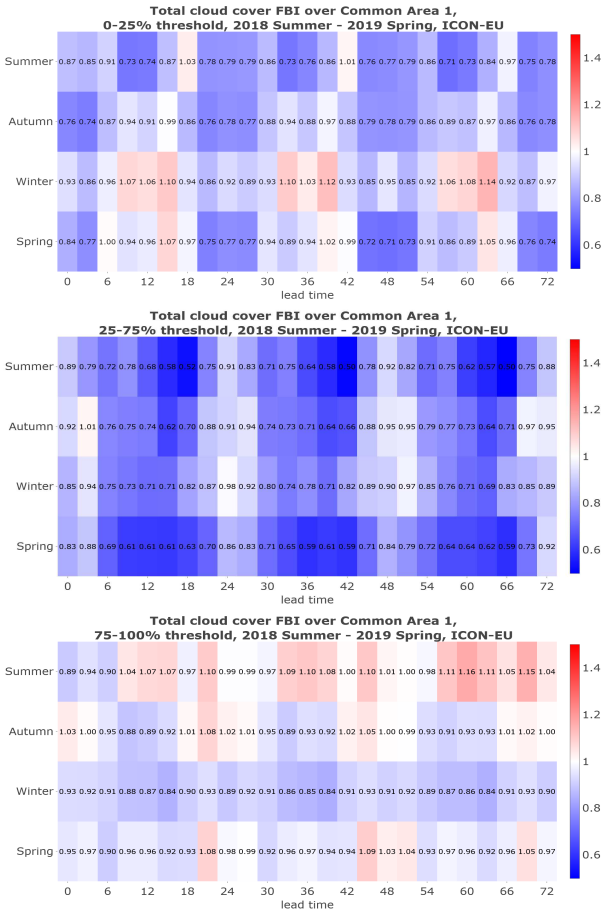

Figure 6: ICON-EU Total Cloud Cover FBI (Frequency Bias Index) for different thresholds

Wind speed at 10 m: COSMO and IFS models exhibit a diurnal bias cycle, with a slight wind speed overestimation at night and early morning, and a slight underestimation around noon hours. ICON models generally underestimate wind speed especially at night, with a minimal diurnal cycle (Figure 7). All models have similar RMSE values and variation, with a slight diurnal cycle in summer period. There is a notable RMSE difference between ICON and ICON-EU during winter, with ICON exhibiting higher values than ICON-EU limited area model. The RMSE trend compared to the previous year was a slight decrease in JJA and SON seasons, and a slight increase in spring. No significant changes for winter months.

10m Wind Gust: The dichotomic scores (FBI, TS) of the 10m wind gust simulations [1] exhibit a general tendency to underestimate 10m wind gust FBI especially for higher thresholds. TS values decrease with increasing threshold.

6h precipitation: Overestimation of low thresholds and underestimation of higher thresholds of precipitation events, is a common feature that is shown similarly to the previous years, derived from the fact that FBI decreases for higher precipitation thresholds (Figures 8, 9). FBI values exhibit a pronounced diurnal cycle in JJA season with higher values around noon which means that mainly noon time precipitation events are overestimated.

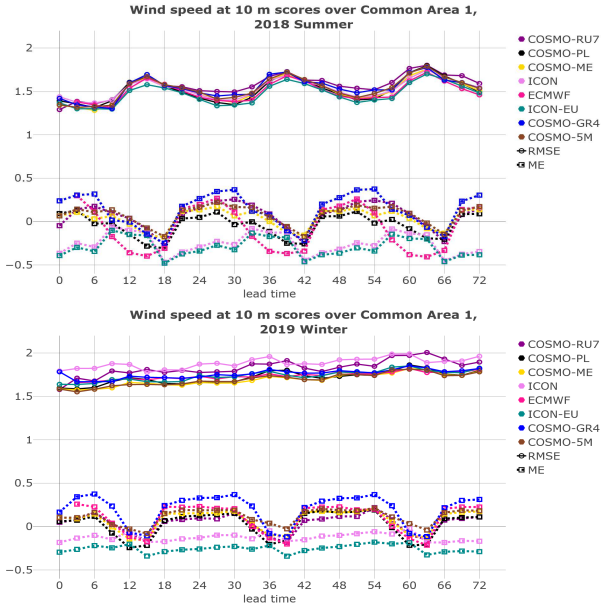

Figure 7: Wind speed at 10 m Continuous scores during summer and winter

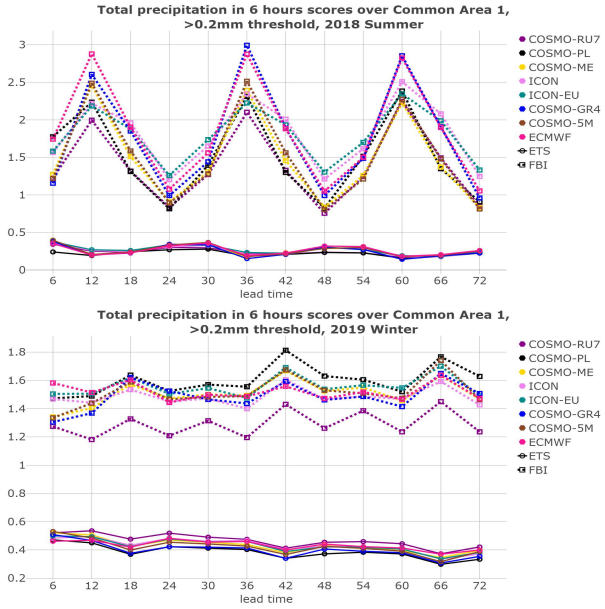

Figure 8: Total precipitation in 6 hours >0.2 mm threshold scores (FBI-ETS) during summer and winter

COSMO models underestimate the frequency of higher threshold precipitation events during summer and autumn, less than ICON model does (better FBI values). ETS (Equitable Threat score) for all models is lower for higher thresholds which means that the forecast quality is worse with increasing threshold. COSMO ETS values are worse than for ICON models.

ETS precipitation scores are improved during winter and are worsen during summer period. The worst ETS values generally occur at forecast times 12-18 UTC, while the best values during the period 00-06 UTC.

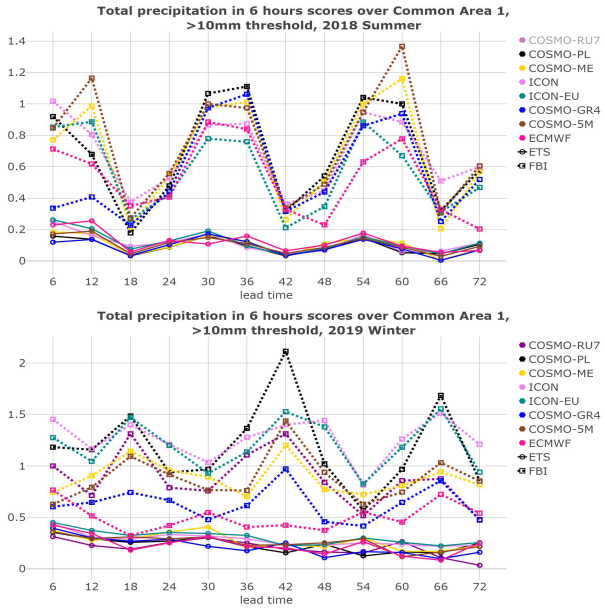

Figure 9: Total precipitation in 6 hours >10 mm threshold scores during summer and winter FBI

4 Conclusions

Common Plots Verification activity which initiated after the 14th COSMO General Meeting in 2012, aims to track systematic errors and long term trends of COSMO model, During Summer 2018 – Spring 2019 while transition from COSMO to ICON model is ongoing, one of the main points of Common Verification activity was the comparison of COSMO and ICON models over the same geographic region and time periods.

COSMO statistical scores behavior (diurnal cycle, evolution with forecast time, seasonal variability) for almost all parameters are persistent over the years. However, the scores absolute values have the tendency to decrease over the years, which denotes the model improvement in subsequent model versions.

ICON model(s) performance is overall improved compared to that of COSMO and IFS models especially for continuous parameters with reduced diurnal cycle of errors and lower RMSE values. Precipitation scores are however comparable. There was some improvement of ICON performance over the last 2 years due to tuning that has been applied as shown in [1] (e.g. dew point ME reduction).

References

[1] Consortium for Small-scale Modeling Verification tasks page.http://cosmo-model.org/content/tasks/verification.priv/default.htm