

SPECIAL PROJECT PROGRESS REPORT

All the following mandatory information needs to be provided. The length should *reflect the complexity and duration* of the project.

Reporting year 2022

Project Title: COSMO and ICON Numerical Weather Prediction Test Suite

Computer Project Account: SPITRASP

Principal Investigator(s): Amalia Iriza-Burca (NMA,Romania) ¹
Ines Cerenzia (Arpae-SIMC, Italy) ²
Enrico Minguzzi (Arpae-SIMC, Italy)²

Affiliation: National Meteorological Administration (NMA) ¹
Regional Agency for Prevention, Environment and Energy of Emilia-Romagna – Hydro-Meteo-Climate Service (Arpae-SIMC) ²

Name of ECMWF scientist(s) collaborating to the project (if applicable) Umberto Modigliani and his staff, Andrea Montani

Start date of the project: 2021

Expected end date: 2023

Computer resources allocated/used for the current year and the previous one (if applicable)

Please answer for all project resources

		Previous year		Current year	
		Allocated	Used	Allocated	Used
High Performance Computing Facility	(units)	5 000 000	3932198.65 (78%)	5 000 000	0%
Data storage capacity	(Gbytes)	2000	11000	4000	11000

Summary of project objectives (10 lines max)

The COSMO and ICON Numerical Weather Prediction Test Suite Special Project intended to continue the activities started in the previous three special projects will ensure the usage of a homogeneous verification platform for both the COSMO and ICON models. This is meant as a benchmark in order to evaluate new versions of the model against existing operational ones, prior to their official release. The aim of using this type of controlled approach for standardized testing and verification for the COSMO and ICON models is to ease the comparison of corresponding model versions (operational against new), in an effort to assess the impact of new features introduced in the code. The set-up and configuration of the models will focus on minimising initial and lateral boundary conditions effect, also eliminating the data assimilation system. Through this approach, performance of each new model version can be thoroughly tested, with an emphasis on newly introduced code developments.

Summary of problems encountered (10 lines max)

No problems encountered.

Summary of plans for the continuation of the project (10 lines max)

The detailed guidelines for the proper use and execution of each NWP test using this platform prepared during previous special projects related to this activity will be revised considering the ICON model and corresponding model configurations. A detailed description of all steps will be included, from the compilation of a new ICON model test version to the final production of the graphics for the statistical scores extracted. Activities (including use of resources) will also include evaluating new official versions of the ICON model prior to their release as well as maintenance of the Test Suite, including migration to the new ATOS cluster. As versions v5.08 and v6.0 are the last for the COSMO model, the COSMO test suite will no longer be needed and as a result it will not be ported to the new ECMWF systems in Bologna.

List of publications/reports from the project with complete references

I. Cerenzia, E. Minguzzi, A. Iriza-Burca, R. Dragomir, F. Gofa, F. Fundel (contributors) - *"Numerical Weather Prediction Meteorological Test Suite": COSMO 5.08 vs. COSMO 5.06 and ICON-LAM v2.6.1*, COSMO-Model Report, June 2021
Massimo Milelli - *"WG6 (PP IMPACT + past PP CEL-ACCEL as needed +COSMO LEPS + COSMO Web)"*, The 23st COSMO General Meeting (teleconferences), Bucharest, Romania, 14 - 17 September 2021
Davide Cesari - *"ICON Test Suite"*, The 23st COSMO General Meeting (teleconferences), Bucharest, Romania, 14 - 17 September 2021

Summary of results

If submitted **during the first project year**, please summarise the results achieved during the period from the project start to June of the current year. A few paragraphs might be sufficient. If submitted **during the second project year**, this summary should be more detailed and cover the period from the project start. The length, at most 8 pages, should reflect the complexity of the project. Alternatively, it could be replaced by a short summary plus an existing scientific report on the project attached to this document. If submitted **during the third project year**, please summarise the results achieved during the period from July of the previous year to June of the current year. A few paragraphs might be sufficient.

Phases I&II: Set-up of the COSMO and ICON models, Configuration and Execution of Runs

The results reported below complete the ones already presented in last year's report, as they were obtained in the first part of 2021.

For a detailed description of the system set-up and model evaluation, we refer to the COSMO model report by Cerenzia et. al (2021) and the presentations by Milelli (2021) and Cesari (2021). The complete overview of all the statistical analysis (graphs and numbers) are available on the COSMO web-site (<https://cosmo-model.org/content/default.htm>).

COSMO and ICON Numerical Weather Prediction Test Suite is aimed at comparing the behaviour of new versions of the COSMO or ICON models with respect to the previous ones in real cases. Where possible, inter-comparison between the COSMO and ICON models are performed. Every test covers 2 months (summer and winter), with different horizontal grid steps (≈ 7 km, $\approx 2.8 / 2.5$ km). The fine runs are 1-way nested in the corresponding coarse run. As all runs are continuous; a restart file is produced every 5 days.

Both suites are implemented on ECMWF systems with cdp/ecflow workflow manager, based on COSMO-LEPS technology. ECMWF computer resources were used for the aim of this task both for simulation and for archiving purposes, through billing units provided by the members as part of the SPITRASP special project “COSMO and ICON numerical weather prediction test suite” approved for 2021-2023. Moreover, the COSMO test suite is based on the old sms/cdp software which will be dismissed after the migration of ECMWF data center to Bologna.

In order to ensure a fair comparison between the models, the horizontal grid steps are comparable for COSMO and ICON runs (0.0625 deg, 0.0025 deg for COSMO and R3B8, R2B10 for ICON ≈ 7 km, $\approx 2.8 / 2.5$ km), while the number of vertical levels is 40 (for 7 km resolution) and 50 (for 2.8 / 2.5 km) for both models. ICON was additionally run with a more state-of-the-art 65 levels at both grid steps. A summary of specifications for the COSMO and ICON test suites is presented in Figures 1 and 2.

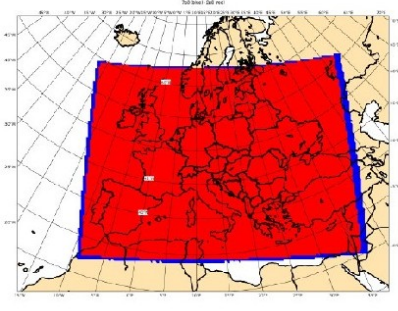

	<i>COSMO</i>	<i>ICON</i>
Lead time	<u>+24h starting daily at 00UTC using warm initialization</u>	<u>+31days using a restart file every 5 days</u>
Domain, Output and Post process		
File format	<u>Grib1</u>	<u>Grib2</u>
Coordinates	<u>Rotated lat-lon and structured grid</u>	<u>Unstructured grid (main option), As COSMO (alternative option)</u>
Domain		

Fig. 1 Summary of specifications for the COSMO and ICON test suites – domain, output and post-processing: COSMO (left) and ICON (right).

	<i>COSMO</i>	<i>ICON</i>
Horizontal and Vertical Resolution		
Coarse Res.	H: <u>7km</u> , V: 40 levels	H: R3B7 (<u>6.6km</u>), V: 40 levels
Fine Res.	H: <u>2.8km</u> , V: 50 levels	H: R2B9 (<u>2.5km</u>), V: 50 levels
Initial and Boundary conditions		
Model levels	from ECMWF HRES analyses at 00, 06, 12, 18UTC and forecasts at 03, 09, 15, 21UTC <u>with 3 hours forecast-range</u>	from ECMWF HRES analyses at 00, 06, 12, 18UTC and forecasts at 03, 09, 15, 21UTC <u>with 3 and 9 hours forecast-range</u>
Soil levels	Initial Condition fom ICON-EU, then free soil (both at 7km and 2.8km)	Initial Condition fom ICON-EU, then free soil (both at 6.6km and 2.5km)
Sea surface	From ECMWF HRES analysis (daily update at 00UTC)	

Fig. 2 Summary of specifications for the COSMO and ICON test suites – model set-up: COSMO(left) and ICON (right).

Since version 5.03 of the model (7 km resolution), all versions are implemented on the Cray HPC. The 2.8km horizontal resolution of the model is tested using the NWP Suite starting with version 5.04a (quasi 5.05), while the single precision configuration (7 km) is also included in the test suite starting with version 5.05. In order to reduce computational costs and reduce the spin-up effect caused by cold-start forecast runs, the forecast mode was replaced with the hindcast mode from version 5.06.

COSMO v5.08 was evaluated against COSMO v5.06 and ICON-LAM v2.6.1. For both COSMO versions, int2lm version 2.05 was used for the interpolation of initial and lateral boundary conditions, while DWD ICON Tools version 2.3.8 was employed for the same purpose for the ICON model. The model output in grib2 format obtained from the experiments is locally stored in the ECFS system. The necessary software used for NWP Test suite is implemented either on cca (MEC) or ecgate (Rfdbk). Model output in grib2 format was processed along with the corresponding observations using the MEC (Model Equivalent Calculator) software, in order to produce the required Feedback Files employed by the Rfdbk (DWD developed) software to produce verification scores.

Following the testing of COSMO version 5.08 against COSMO version 5.06 and ICON version 2.6.1, it was decided that there was no need to test the COSMO version 6.0, since the difference in the results are limited to sea-ice parameterization, which has no impact on the test suite.

As versions v5.08 and v6.0 are the last for the COSMO model, the COSMO test suite will no longer be needed and as a result it will not be ported to the new ECMWF systems in Bologna.

Phase III: MODEL OUTPUT VERIFICATION

The MEC (Model Equivalent Calculator)+Rfdbk verification procedure includes the conversion of observations from bufr to netcdf format (using *bufr2netcdf*), pre-processing of model output in grib format for ingestion in MEC, processing model output and corresponding observations to obtain feedback files (MEC), execution of verification procedures (Rfdbk) and transfer and visualisation of results on the COSMO shiny server. The verification results are periodically collected in a report.

The verification was performed with grid-to-point comparisons in order to compare gridded surface and upper-air model data to point observations. The selected NWP suite stations are situated in an area covering -25/24/65/65 (W/S/E/N) and are around 3200. Due to the specifications of the verification system for hindcast runs (single run), +24 hours lead time is shifted to 0.

The verification modules for testing the COSMO and ICON models include *surface continuous parameters*, *precipitation verification (6h and 12h)* and *upper air verification (TEMP based)*, as follows:

- 2m temperature (T2M), 2m dew point (TD2m), 10 meter wind speed (FF), total cloud cover (N), surface pressure (PS): mean error (ME), root mean square error (RMSE), mean absolute error (MAE), standard deviation (SD), R^2 , TCC (tendency correlation), LEN (number of observations used), OMEAN and FMEAN (observed and forecast mean), etc.;
- precipitation for selected thresholds (greater than 0.2, 0.4, 0.6, 0.8, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 25, 30): probability of detection (POD), false alarm rate (FAR), equitable threat score (ETS), frequency bias (FBI), Performance diagrams, etc.
- upper air temperature (T), relative humidity (RH) and wind speed (FF) for selected pressure levels (250., 500., 700., 850., 925., 1000.): BIAS, MAE, RMSE, SD, etc.

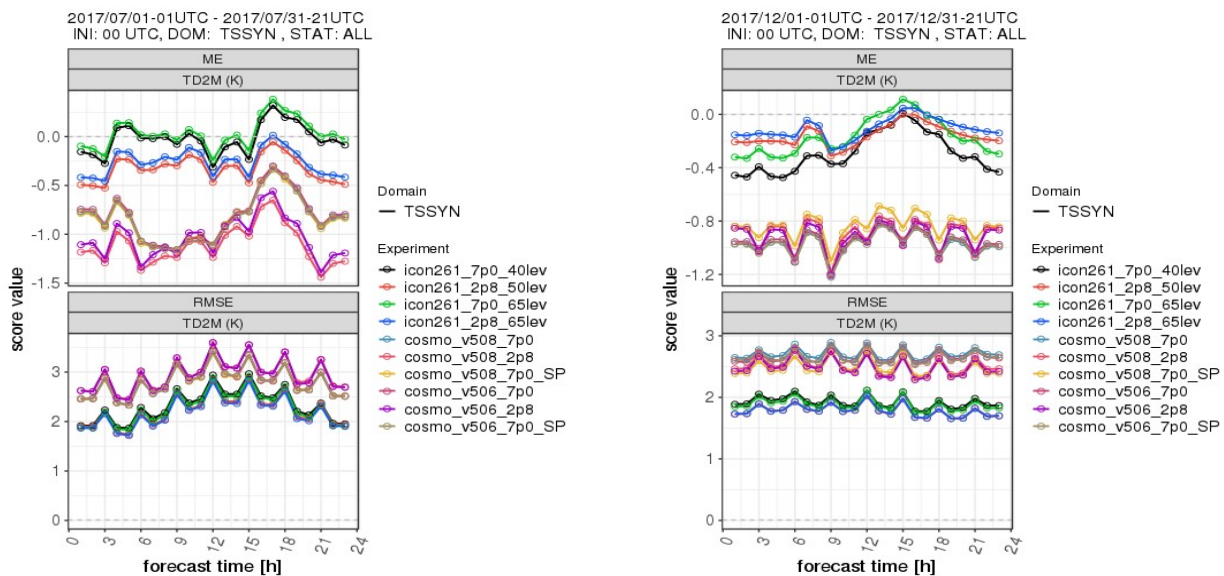


Fig. 3 2-meter dew point temperature verification results - July 2017 (left), Dec 2017 (right), for: ICON-7p0/40levels (black), ICON-2p8/50levels (red), ICON-7p0/65levels (blue), ICON-2p8/65levels (green), COSMO-v5.08-7p0 DP (purple), COSMO-v5.08-2p8 (orange), COSMO-v5.08-7p0 SP (yellow), COSMO-v5.06-7p0 DP (brown), COSMO-v5.06-2p8 (pink) and COSMO-v5.06-7p0 SP (grey); ME (top) and RMSE (bottom). Red/gray filled dots indicate a significant/insignificant (95% level) difference of scores between the model versions.

The evaluation of 2m Dew Point Temperature results (Figure 3) for COSMOv5.08/v5.06 shows that RMSE values are very similar for all model versions and setups. The 2.8km model performs slightly better, while errors with COSMO v5.08 seems lightly higher than for COSMO v5.06, especially during the winter period. All models/setups underestimate 2m dew point temperature, for both seasons, June 2022

while an increased underestimation is shown for the month of July; for this month, the underestimation is lower for COSMO v5.08 for both resolutions. With regards to the COSMO/ICON-LAM comparison, all ICON models outperform compared to COSMO ones for all setups and both seasons. RMSE is reduced in all cases by around 0.5deg with insignificant changes among ICON models with different resolution or model levels. This parameter is underestimated for all models and seasons, but with ICON-LAMs this tendency is much reduced.

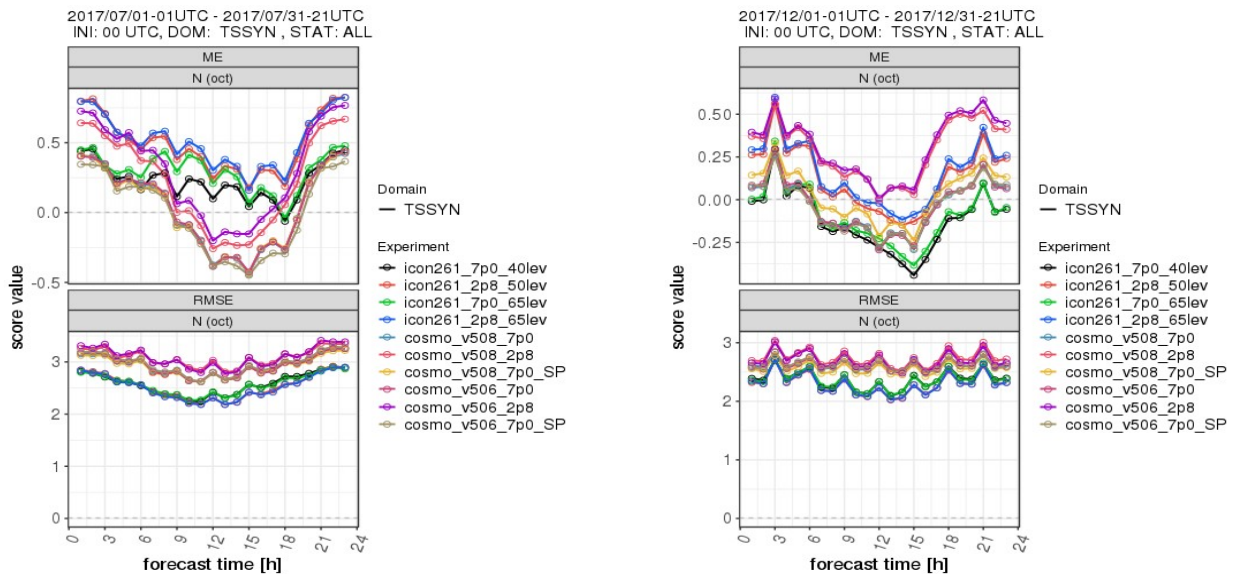


Fig. 4 Total cloud cover verification results - July 2017 (left), Dec 2017 (right), for: ICON-7p0/40levels (black), ICON-2p8/50levels (red), ICON-7p0/65levels (blue), ICON-2p8/65levels (green), COSMO-v5.08-7p0 DP (purple), COSMO-v5.08-2p8 (orange), COSMO-v5.08-7p0 SP (yellow), COSMO-v5.06-7p0 DP (brown), COSMO-v5.06-2p8 (pink) and COSMO-v5.06-7p0 SP (grey); ME (top) and RMSE (bottom). Red/gray filled dots indicate a significant/insignificant (95% level) difference of scores between the model versions.

Total Cloud Cover (Figure 4) simulated by the COSMO model exhibits no change in the results between the 5.06 and 5.08 versions for both resolutions/seasons/precisions. Summer forecasts underestimate the TCC during the warm hours of the day especially with the 7km resolution, while the opposite trend (overestimation) is shown during winter night hours. No significant difference between the two model versions. COSMO/ICON-LAM: Also for this parameter for all forecast hours and both seasons, ICON-LAM performs better than COSMO versions. RMSE is reduced in all cases while the behavior of TCC forecasts indicated through ME reveals an overestimation in the summer with ICON which differs from the behavior of COSMO model forecasts.

The scores for the performance of upper air parameters (relative humidity, temperature and wind speed) for COSMOv5.08 and v5.06 (Figures 5 and 6) show similar behaviour for both models in general. Relative Humidity exhibits small differences between the two model versions with the 7km resolution and the 2.8km one. For the winter period, there is a small reduction in the overestimation of RH during the night in the middle atmosphere. During the warm hours of the day there is also a reduced underestimation, but smaller. During summer, the new version performs again slightly better, with reduced overestimation of RH from the surface up to almost 500mb, while during the night there is no difference between the two model versions. RMSE values exhibit overall no significant difference. With the 2.8km version, values are almost identical for both months. Overall, there is a small positive impact of version 5.08.

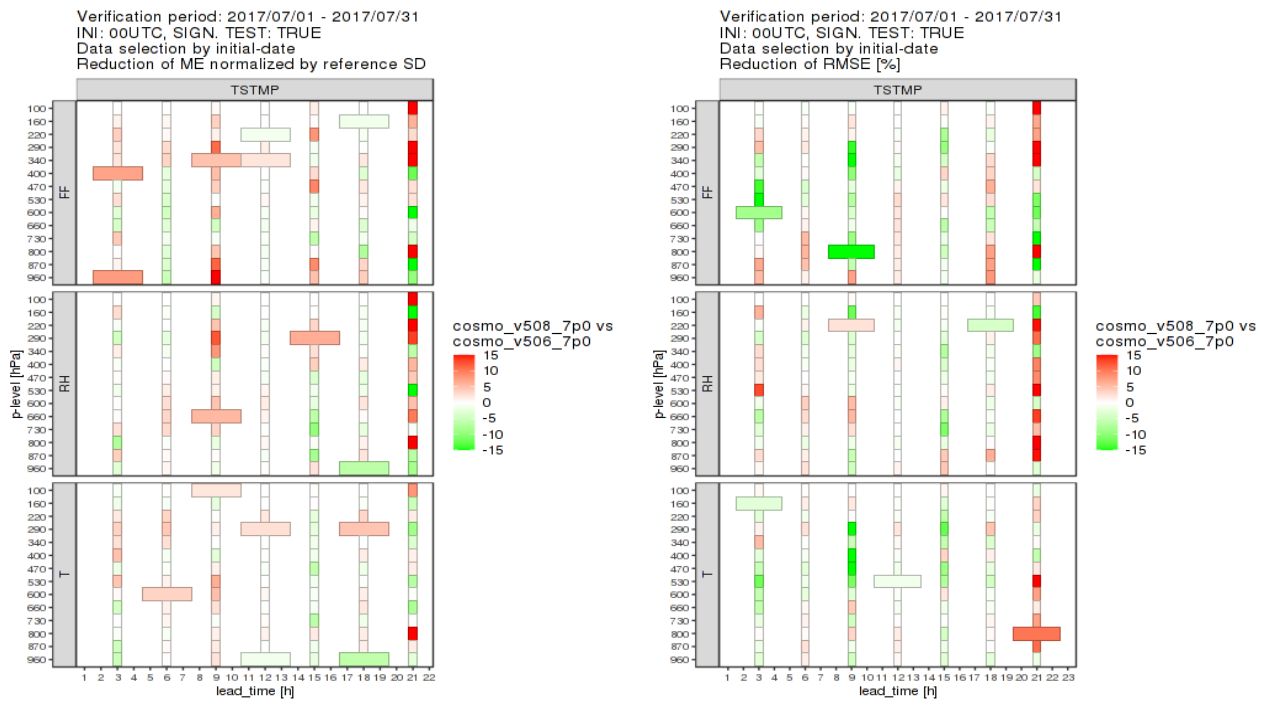


Fig. 5 Upper Air verification results - July 2017, reduction in ME (left) and RMSE (right) for FF, RH and T (top to bottom). Colors indicate COSMO-v5.08-7p0 DP better (green) and COSMO-v5.06-7p0 SP better (red).

Temperature comparison for the two model versions suggest insignificant differences in almost all cases for both resolutions. Wind speed performance is similar, with no strong trend in the impact on the performance in the 7km version model. RMSE values do not change significantly with the new model, while ME exhibits a steady behaviour for both resolutions in the winter, with overestimation that with the newer version is slightly reduced in the upper atmosphere. For the summer period, the behaviour exhibits no difference with the coarse resolution models. With the 2.8km resolution however, there is an increase in overestimation almost for all levels during summer, while there is no difference in the performance during winter.

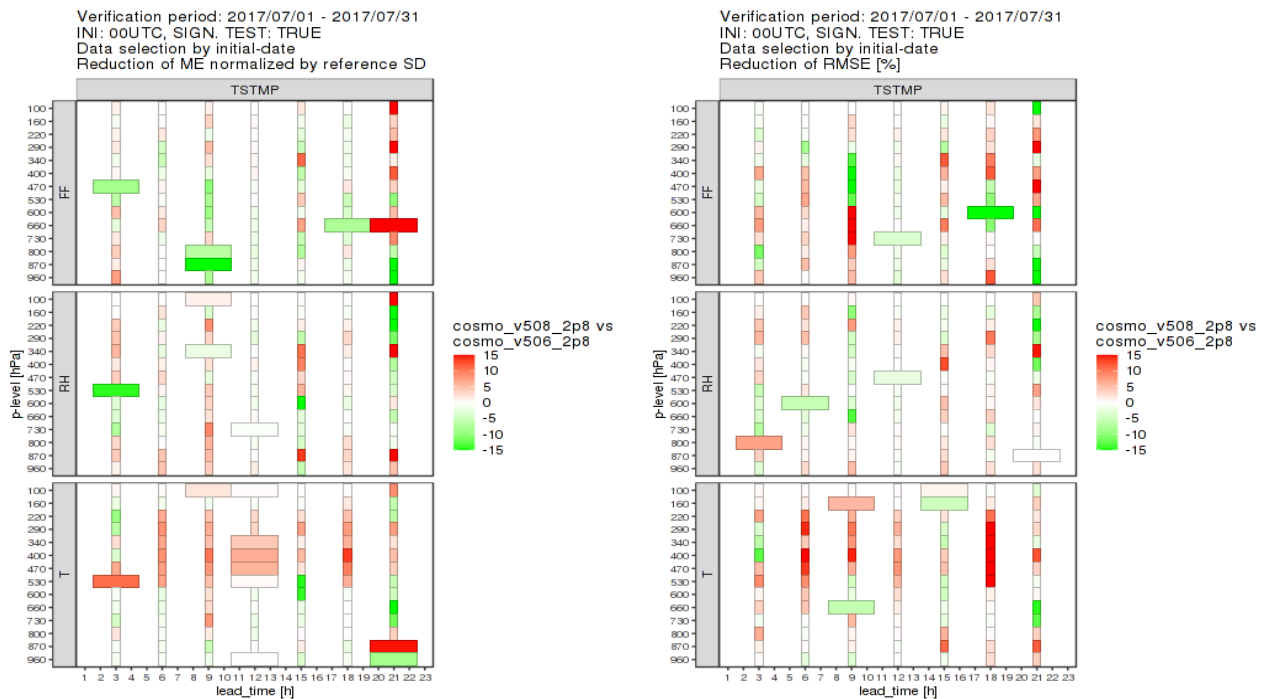


Fig. 6 Upper Air verification results - July 2017, reduction in ME (left) and RMSE (right) for FF, RH and T (top to bottom). Colors indicate COSMO-v5.08-2p8 DP better (green) and COSMO-v5.06-2p8 SP better (red).

Phase IV: Additional steps and further actions

Considering the upcoming switch off of Cray HPC in Reading, the ICON-test suite need to be migrated to the new Atos system (aa) in Bologna. As a consequence, effort has been dedicated to implementing and testing the activities currently running on the Cray HPCs (cca and ccb) to the new system, such as:

- implementation of the ICON Test Suite to the Atos system (model configuration and integration, processing of model output for production of feedback files)
- implementation of the MEC system for production of feedback files
- implementation of the Rfdbk package dedicted to the calculation of statistical scores.

It is planned to conclude these migrations within the end of 2022 or anyway before to evaluate any new official version of the ICON model.

As versions v5.08 and v6.0 are the last for the COSMO model, the COSMO test suite will no longer be needed. As a result it will not be ported to the new ECMWF systems in Bologna.

Activities (including use of resources) will also include evaluating new official versions of the ICON model, most likely during the second part of the year. A definitive time line for the release of a new ICON model version is not yet available, but it is anticipated that a new request will come within 2022. Otherwise, the resources for 2022 will be only in part used for the porting of the ICON-test suite on ATOS.