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	2019
Project Title:	**Testbed for the Evaluation of COSMO Model Versions**
Computer Project Account:	**SPITRASP**
Principal Investigator(s):	Amalia Iriza (NMA, Romania)[1] Ines Cerenzia (Arpae-SIMC, Italy)[2]
Affiliation:	National Meteorological Administration (NMA)[1] Environmental Agency 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:	2018
Expected end date:	2020

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 + 2 650 000	4 417 294.39 (57%)	5 000 000	3.146.320.06 (63%)
Data storage capacity	(Gbytes)	2000	2400	1000	1600

Summary of project objectives (10 lines max)

The main objective of the **"Testbed for the Evaluation of COSMO Model Versions"** Special Project is to perform testing of new COSMO model versions prior to their official release using the software environment built on the ECMWF platform during previous SPITRASP projects (2013-2015, 2016-2018). This evaluation of new model versions carried out according to source code management procedures and using the Test Suite platform is taken into account before any operational implementation and release of an official model version. The NWP test suite currently represents a benchmark for rigorous testing of all new model features and allows the model developers to produce guidelines for the selection of a new operational implementation of the model. Several model versions and configurations have been installed and tested up to now in the framework of the SPITRASP special projects, while more are expected to be evaluated using this platform.

Summary of problems encountered (10 lines max)

No problems encountered.

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

Activities (including use of resources) to test the new official version (5.06) of the COSMO model prior to its release in autumn 2019 are on-going.
Evaluation procedures will also be carried out in the second part of the year, when another release of the COSMO model is anticipated.
Implementation of model output also in grib2 format for the verification package.
Maintenance of the Test Suite.
Extending the Test Suite to evaluate ICON model official releases (migration to ICON-LAM: prepare an EcFlow suite on HPC, ICONTOOLS and ICON executable already available).

List of publications/reports from the project with complete references

F. GOFA - "WG5 overview of Activities", The 20th COSMO General Meeting, St. Petersburg, Russia, 3 - 6 September 2018
A. MONTANI - "NWP test Suite Some Updates", The 20th COSMO General Meeting, St. Petersburg, Russia, 3 - 6 September 2018
A. MONTANI, F. FUNDEL, M. BOGDAN, R.C. DUMITRACHE, F. GOFA, A. IRIZA-BURCA, F. BATIGNANI (contributors) - "Numerical Weather Prediction Meteorological Test Suite: COSMO 5.03 vs. 5.05", COSMO-Model Report, June 2018

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.

The NWP test suite procedure was adopted by COSMO in order to perform carefully-controlled and rigorous testing, including the calculation of verification statistics, for any COSMO model test-version. Following the source code management procedure, this testing phase should offer the necessary information on the model forecasting performance, in order to determine whether the upgrade of a model test-version to a new release version is possible. For previous testing procedures, the VERSUS system has been used to perform verification. All activities were

performed first during the frame of the COSMO Priority Task NWP Test Suite (2013-2015) and as part of special projects at ECMWF (2013-2015, 2016-2017).

Phases I & II: Model set-up & Model Configuration and Execution of Runs

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 "Testbed for the Evaluation of COSMO Model Versions" approved for 2018-2020. The platform has already been used for a large number of COSMO versions (starting from 5.00 up to 5.05). In recent years, the testing procedure has been enriched and modified.

Since version 5.03 of the model, all versions are implemented on the Cray HPC. Starting from version 5.04a (quasi 5.05) of the COSMO model, the 2.8km horizontal resolution of the model is also tested using the NWP Suite, in addition to the previously used setup at 7km. For version **5.05**, the double precision configurations (both 7 km and 2.8 km horizontal resolution) were tested against version 5.03 of the model, while the single precision (SP) setup for v5.05 was tested against the double precision (DP) one.

Details about the implementation and testing procedure for COSMO v5.05 (7km horizontal resolution only) can be found in the report by Montani et al. (2018). Finally, during this year (starting from version 5.06), the forecast mode applied to the test simulations until now is being replaced with the hindcast mode. This change is aimed at reducing the computational costs and time for testing.

Present on-going activities are dedicated to the evaluation of version **5.06** of the COSMO model. For these tests, version **5.05_1** (v5.05 with bug fixes) is used as benchmark/operational version. The int2lm 2.05 version was used for the interpolation of initial and lateral boundary conditions. The directory structure and the archiving procedures for the new COSMO-5.05 (single precision, forecast mode) and COSMO-5.05_1, COSMO-5.06 (hindcast mode) model versions follow the ones used for the previous implementations.

After completion of the testing procedure, model outputs were processed together with the corresponding observations using MEC (Model Equivalent Calculator), a software aimed at producing the necessary Feedback files, and Rfdbk (DWD developed), a software that utilizes R libraries to process Feedback.

For version 5.05 (DP and SP) the forecast period of each daily run is 72 hours for 7km and 48 hours for 2.8 km, on one daily cycle based on the 00UTC initializing data from the ECMWF HRES system. Simulations were performed for one month in winter (December 2017) and one month in summer (July 2017), 2 months in total for each model version. Table 1 summarizes the main features of the models involved in the COSMO NWP Test Suite for version 5.05 of the COSMO model.

Table 1: Main features of the models used in the NWP Test Suite

	ECMWF HRES	**COSMO 7p0**	**COSMO 2p8**
Grid points (nx x ny)	901 x 501	661 x 471	1587 x 1147
Model levels	137	40	50
Resolution (dx x dy)	0.1 x 0.1	0.0625 x 0.0625	0.025 x 0.025
Forecast range (h)	72	72	48

The integration domains at 7 km and 2.8 km are presented in figure 1.

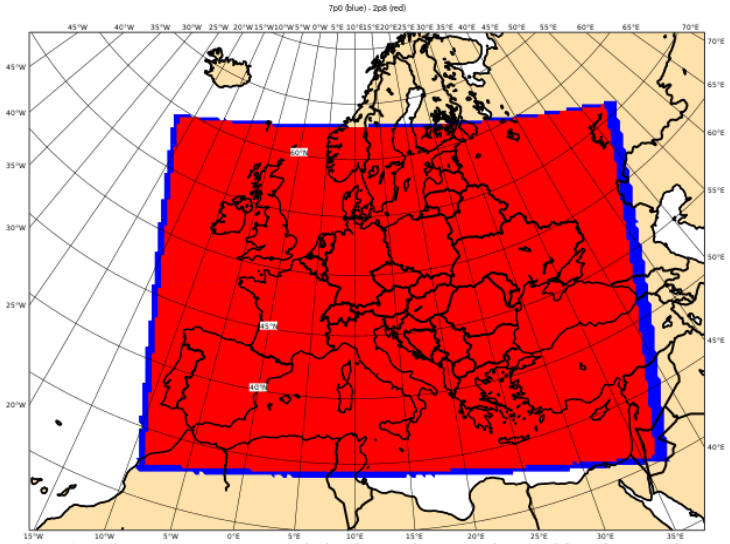

Fig. 1 Integration domains for the COSMO model: domain at 7 km of horizontal resolution (blue) is below the domain at 2.8 km of horizontal resolution (red).

For the evaluation of version 5.05 of the COSMO model (released in 2018), the following tests were performed using version 5.03 as benchmark:

- ✓ runs at 2.8 km, 50 model levels; 48h forecast range, forecast mode, DP
- ✓ runs at 7.0 km, 40 model levels; 72h forecast range, forecast mode, DP
- ✓ runs at 7.0 km, 40 model levels; 72h forecast range, forecast mode, SP (with version 5.05 DP as benchmark)

Starting from version 5.06 of the model, evaluations are performed only in hindcast mode in both double and single precision versions for the 7 km horizontal resolution setup and only in double precision for the 2.8 km horizontal resolution configuration. For the hindcast mode, initial conditions are provided by ECMWF HRES analysis, whereas lateral boundary conditions are introduced with a 3 hourly frequency and they include the ECMWF HRES analyses (at hours 00, 06, 12 and 18UTC) and short cut off analyses (at hours 03, 09, 15 and 21UTC) with soil initialized from ICON-EU, then free soil (both model resolutions).

The model output obtained from the experiments is locally stored in the ECFS system. All the necessary software (MEC, Rfdbk) used for NWP Test suite purposes are also implemented on ecgate.

To evaluate version 5.06 (test version), the following tests are being performed using version 5.05_1 as benchmark:

- ✓ runs at 7.0 km, 40 model levels; hindcast mode, DP
- ✓ runs at 7.0 km, 40 model levels; hindcast mode, SP
- ✓ runs at 2.8 km, 50 model levels; hindcast mode, DP

Cost of the ECMWF Suite

The new ECMWF Special Project SPITRASP has an allocation of 5.000.000/year (2018-2020). The costs of the suite in forecast mode for the COSMO configurations v5.03 and v5.05 with icon soil are presented in Table 2 for both the double precision and the single precision versions (7 km horizontal resolution), respectively. The costs of the test suite in hindcast mode for the configurations v5.05_1 and v5.06 are presented in Table 3 (double and single precision for 7 km, double precision only for 2.8 km).

Compared to the forecast mode, the hindcast mode costs in terms of BU and simulation time about one third and the half for running respectively the COSMO model at 7 km and 2.8 km horizontal resolutions. Differences are mainly due to reduced time range of hindcast simulations (24h long)

with respect to the 72h and 48h forecast range set for COSMO in forecast mode at 7km and 2.8km respectively.

Table 2 Cost of the suite for configurations v5.03 and v5.05 (on Cray).

INT2LM from IFS to COSMO-7km	
HRES → 7p0 ~ 47 BU, ~ 5min,	
EC_total_tasks=36, EC_nodes=1	
INT2LM from ICON to COSMO-7km	
ICON → 7p0 ~ 40 BU, ~ 6min	
EC_total_tasks=24, EC_nodes=1	
INT2LM from COSMO-7km to COSMO-2.8km	
7p0_DP→2p8_DP ~ 278 BU, ~ 14 min 24 sec	
EC_total_tasks=72, EC_nodes=2	
COSMO-5.03	***COSMO-5.05***
7p0_DP ~ 3000 BU ~ 28min	**7p0_DP** ~ 2500 BU, ~ 13min
EC_total_tasks=480, EC_nodes=20	EC_total_tasks=720, EC_nodes=20
2p8_DP ~ 36500 BU, ~ 1h 50min 16sec	**2p8_DP** ~ 27250, ~ 1h 45
EC_total_tasks=1296, EC_nodes=36	EC_total_tasks=972, EC_nodes=27
	7p0_SP ~ 1650 BU, ~ 9min
	EC_total_tasks=720, EC_nodes=20

Table 3 Cost of the suite for configurations v5.05_1 and v5.06 (on Cray).

INT2LM from IFS to COSMO-7km	
HRES → 7p0 ~ 17 BU, ~ 1.5min,	
EC_total_tasks=72, EC_nodes=1	
COSMO-5.05_1	***COSMO-5.06***
7p0_DP ~ 890 BU, ~ 5 min	**7p0_DP** ~ 800 BU, ~ 5min
EC_total_tasks=720, EC_nodes=20	EC_total_tasks=720, EC_nodes=20
7p0_DP→2p8_DP ~ 115 BU, ~ 6 min	**7p0_DP→2p8_DP** ~ 121 BU, ~ 6 min
EC_total_tasks=72, EC_nodes=2	EC_total_tasks=72, EC_nodes=2
2p8_DP ~ 13.535 BU, ~ 53 min	**2p8_DP** ~ 12.708 BU, ~ 50 min
EC_total_tasks=972, EC_nodes=27	EC_total_tasks=972, EC_nodes=27
7p0_SP ~ 712 BU, ~ 4 min	**7p0_SP** ~ 761 BU, ~ 4 min
EC_total_tasks=720, EC_nodes=20	EC_total_tasks=720, EC_nodes=20

Phase III: Model Output Verification

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 **3600** for this stratification (see figure 2). Suspect observation values had been previously created for each parameter (forecast-observation greater than a specific limit are excluded) and included in the verification test in order to eliminate errors that are connected with observations. For this test, the verification procedure based on the Rfdbk/MEC system has been completed and is presented in this report. Due to the requirements of MEC software, all observations were converted in netcdf format with the bufr2netcdf software.

The verification modules for testing v5.05 are the following:

- **surface continuous parameters** 2m temperature (T2M), 2m dew point (TD2m), wind speed (FF), total cloud cover (N), surface pressure (PS): BIAS (ME), RMSE, SD, R^2, TCC (tendency correlation), LEN (# of observations used), OMEAN and FMEAN (observed and forecast mean) – up to 72 hours forecast horizon;
- **precipitation verification** (6h, 12h) 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): ETS, FBI, Performance diagrams, etc. – up to 72 hours forecast horizon;
- **upper air verification (TEMP based)** – Temperature (T), dew point (TD), relative humidity (RH), wind speed (FF) and wind direction (DD) for selected pressure levels (250., 500., 700., 850., 925., 1000.): BIAS, MAE, RMSE. SD, etc. – up to 72 hours forecast horizon.

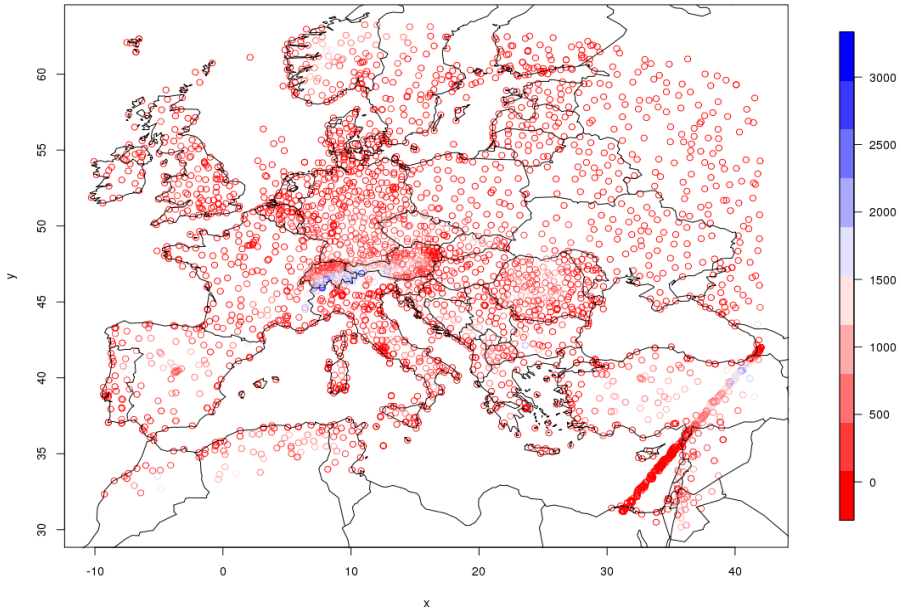

Fig. 2 Location of meteorological stations used for the verification.

Double precision runs for v5.05 (7km, forecast and 2.8km, forecast)

For the **double precision** run, verification was performed for the 5.05 version of the COSMO model at 7 km resolution against version 5.03, for the months of July and December 2017. A complete overview of all the statistical analysis regarding the comparison of the COSMO 5.03 versus COSMO 5.05 versions (graphs and numbers) is presented in the technical report by Montani et al. (2018), including the detailed description for the compilation of the new model version, model set-up configurations and the final production of the graphics for the statistical scores extracted.

Single precision runs for v5.05 (7km, forecast)

For the **single precision** run, the Rfdbk based verification was performed for the **5.05 version** of the COSMO model at **7 km resolution**, against the **double precision run of the same version**, for the months of July and December 2017. The same verification procedures as for the comparison of 5.05 against 5.03 (double precision, forecast mode) were employed. Some selected statistical results that were obtained through the R-based Rfdbk verification system are presented below, in figures 3 - 5, while a complete overview of all the statistical analysis (graphs and numbers) can be studied in the report by Montani et al. (2018).

A summary of the main findings for surface parameters is given below:

- 10m wind speed differences between the DP and the SP runs of the model both for the winter and summer season are insignificant, as is also the case for mean sea level pressure.
- Both COSMO 5.03 and COSMO 5.05 underestimate the values forecasted for 2m dew point temperature for the entire winter period and for most of the summer period. Apart from slightly smaller ME values for the winter period for the DP runs, differences between DP and SP runs are mostly insignificant.
- For 2m air temperature both runs display an underestimation of forecasted values for most of the winter period. With respect to ME and RMSE, the situation is similar to that of 2m dew point temperature, slightly smaller ME values for the winter period for the DP runs, while differences between DP and SP runs are mostly insignificant.
- Results for the Total Could Cover needed some further investigation due to differences in the DP verification compared to the previous results.

For the forecast of precipitation (6h and 12h accumulation), the statistics of the two model runs (DP and SP) are quite similar while some problematic behavior is again observed for some timesteps. As in the case of the previous DP verification (5.05 against 5.03), this cannot be attributed to the model performance as it is not systematic, so further analysis of the input data was required.

For the upper air parameters as well, differences between the DP and the SP runs of the 5.05 version are insignificant, both for the summer and the winter precision.

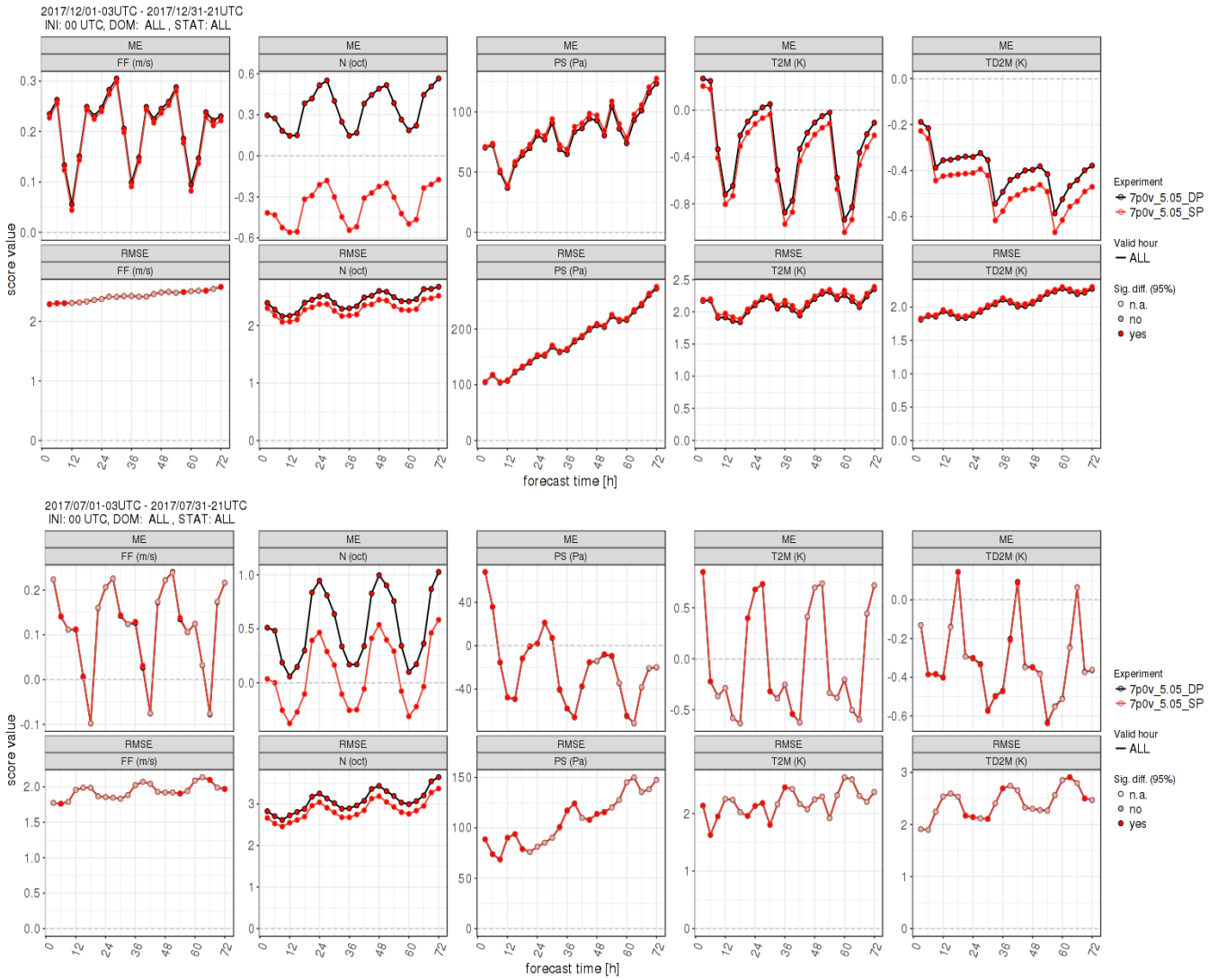

Fig. 3 COSMO-7km Continuous parameters verification results (00UTC run) - COSMO 5.05 DP (black) and SP (red) mean error (ME) and root mean square error (RMSE) for: Dec 2017 (top) July 2017 (bottom). Red/gray filled dots indicate a significant/insignificant (95% level) difference of scores between the 2 model runs.

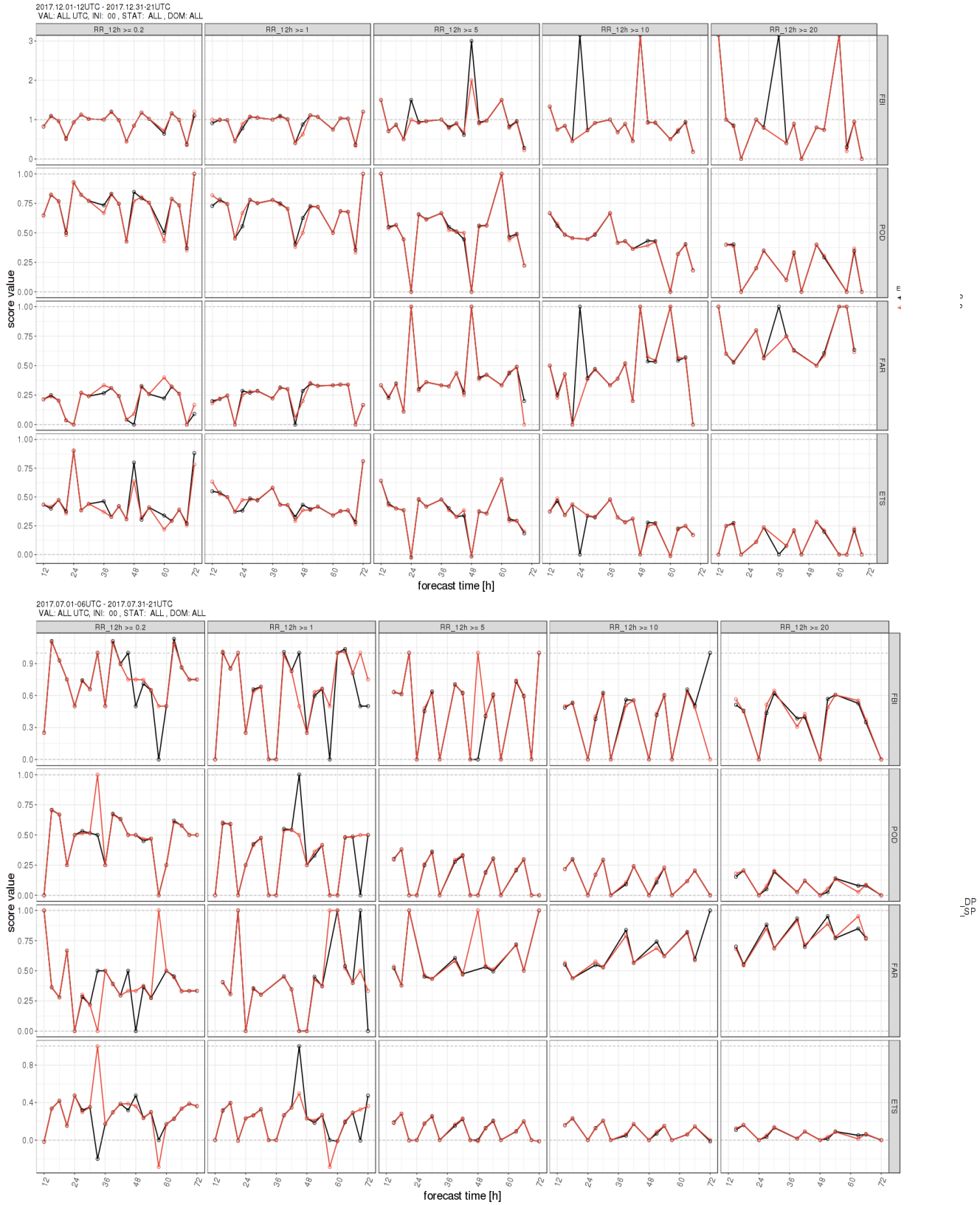

Fig. 4 COSMO-7km 12h precipitation verification results (00UTC run) - COSMO 5.05 DP (black) and SP (red): FBI, POD, FAR and ETS for thresholds 0.2, 1,5,10,20mm/12h, Dec 2017 (top), July 2017 (bottom).

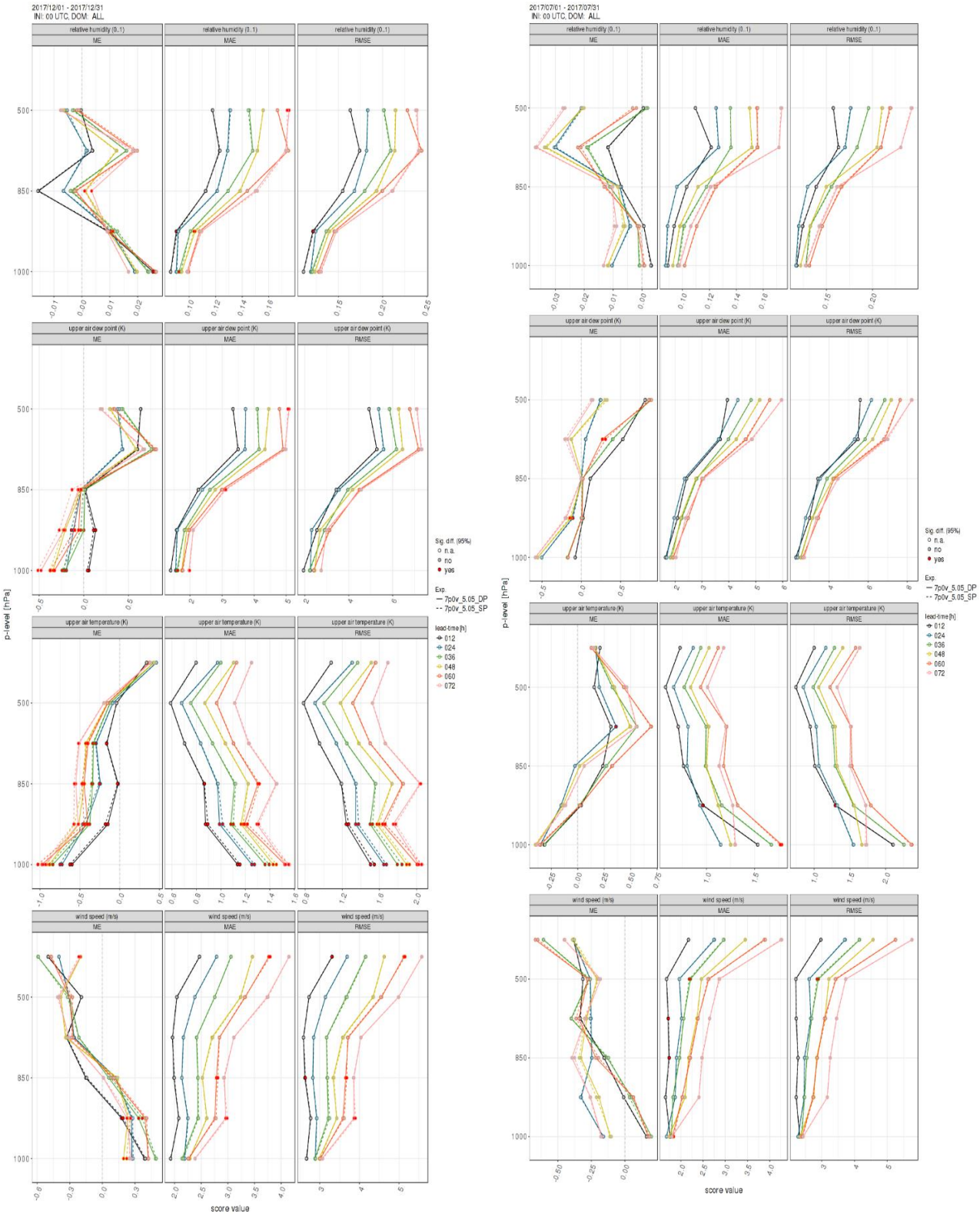

Fig. 5 COSMO-7km UpperAir ME-RMSE results (left: Winter, right: Summer) - COSMO 5.05 DP (continuous line) and SP (dashed line): RH, Td, T, Windsp.

Phase IV: Additional steps

Testing of version 5.06 is on-going in hindcast mode, for both double and single precision runs at 7 km horizontal resolution and double precision runs at 2.8 km horizontal resolution.