SPECIAL PROJECT PROGRESS REPORT

Progress Reports should be 2 to 10 pages in length, depending on importance of the project. All the following mandatory information needs to be provided.

Reporting year	2018
Project Title:	**Testbed for the Evaluation of COSMO Model Versions**
Computer Project Account:	**SPITRASP**
Principal Investigator(s):	Amalia Iriza (NMA,Romania)[1] Andrea Montani (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
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	2395074.93 (50%)	5.000.000	4155131.28 (83%)
Data storage capacity	(Gbytes)	1000	1200	1000	2400

Summary of project objectives

(10 lines max)

In the frame of the **"Testbed for the Evaluation of COSMO Model Versions"** Special Project the main objective 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). The 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 moreare expected to be evaluated using this platform.

Summary of problems encountered (if any)

(20 lines max)

No problems encountered.

Summary of results of the current year (from July of previous year to June of current year)

This section should comprise 1 to 8 pages and can be replaced by a short summary plus an existing scientific report on the project

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 previously registered for 2016-2017 (COSMO NWP Meteorological Test Suite, project account SPITRASP, ended in 2017) and the follow-up special project "Testbed for the Evaluation of COSMO Model Versions" approved for 2018-2020. 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. Given the fact that the test suite was developed starting with 2013 and a large number of model versions have already been evaluated using this platform, activities in this period were also dedicated to improving the use of the test suite by migrating to a more recent evaluation period and additional model configurations that can offer even more insight regarding the impact of newly implemented features on the development of the COSMO model.

From version 5.03 of the model onwards, all COSMO versions are implemented on the Cray HPC. The current operational model version (5.03) was previously implemented for evaluation against COSMO version 5.01 and was also used as reference version (operational) to evaluate versions 5.04a and 5.04e, using initial and lateral boundary conditions from the IFS model. Versions 5.04x of the model (quasi 5.05) were not meant specifically for operational use, but rather to test various new model developments which would be available in the official 5.05 version. As a consequence, versions 5.03 will again be used as reference version (operational) against which COSMO-5.05 will also be evaluated.

For the current tests (modifications to the test suite and testing of versions 5.05), initial and lateral boundary conditions were obtained from the ICON global model for the 7 km resolution. For the 2.8 km resolution, initial and lateral boundary conditions were interpolated from the 7 km model output.

The forecast period of each daily run is 72 hours for 7km and 48 hours for 2.8 km. While previous tests and evaluations of model runs were performed for January and July 2013, current and future test suite simulations will now be performed for July 2017 and December 2017, 2 months in total (for each model resolution, each model version and configuration). As a consequence, a series of new tests were performed for version 5.03 of the COSMO model (both 7 km and 2.8 km horizontal resolution) to prepare for new adaptations of the test suite. The verification procedures were performed both with VERSUS and the new R-based DWD software in order to obtain an objective comparison of the results from the two verification solutions.

At present, there are on-going activities for the testing and evaluation of version 5.05 of the COSMO model (to be released this summer). For these activities, new tests will be performed, as follows:

- ✓ runs at 2.8 km, 50 model levels; 48h forecast range, forecast mode, double precision (model integration complete for versions 5.03 – operational and versions 5.05 – new test version)
- ✓ runs at 7.0 km, 40 model levels; 72h forecast range, forecast mode, double precision (model integration complete for versions 5.03 – operational and versions 5.05 – new test version)
- ✓ runs at 7.0 km, 40 model levels; 72h forecast range, forecast mode, single precision (versions 5.03 – operational and 5.05 – test)
- ✓ runs at 7.0 km, 40 model levels, hindcast mode (30 days forecast range), single precision (versions 5.03 – operational and 5.05 – test)
- ✓ runs at 2.8 km, 50 model levels; hindcast mode (30 days forecast range), double precision (versions 5.03 – operational and 5.05 – test)
- ✓ runs at 7.0 km, 40 model levels; hindcast mode (30 days forecast range), double precision (versions 5.03 – operational and 5.05 – test)

The int2lm 2.05 version was used for the interpolation of initial and lateral boundary conditions. Due to restrictions from the initial and lateral boundary conditions, the integration domains were slightly restricted compared to previous tests (figure 1).

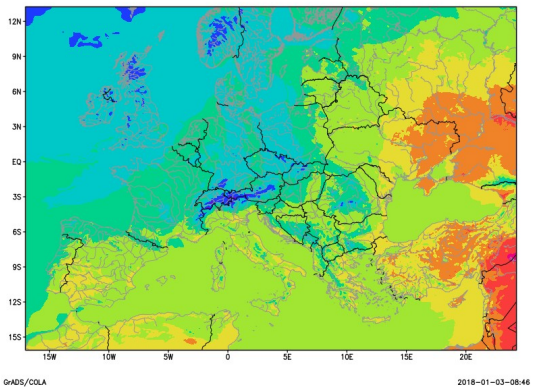

Fig. 1 Integration domain for the COSMO model used in current tests.

- COSMO@7p0: ie_tot=661, je_tot=471, 40ML; dlon=dlat=0.0625; fc+72h;
- COSMO@2p8: ie_tot=1587, je_tot=1147, 50ML; dlon=dlat=0.125; fc+48h.

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 for the COSMO configuration v5.03 with icon soil are presented in table 3 for both 7km and 2.8km versions, respectively.

Table 3 Cost of the suite in the present configuration (on Cray).

INT2LM from ICON to COSMO-7km	***INT2LM from COSMO-7km to COSMO-2.8km***
about 40 BU per run (~ 6min)	about 278 BU per run (~ 14 min 24 sec)
EC_total_tasks=24, EC_nodes=1	EC_total_tasks=72, EC_nodes=2
COSMO-5.03-IS	**COSMO-5.03-2p8-IS**
about 3000 BU per run (~ 28min)	about 36500 BU per run (~ 1h 50min 16sec)
EC_total_tasks=480, EC_nodes=20	EC_total_tasks=1296, EC_nodes=36

Phase III: Model Output Verification - The VERSUS System

The VERSUS system installed on the dedicated virtual machine at ECMWF was updated with three more additional patches that have updated some of the features already available in the system and have helped in solving some of the existing problems.

Following the procedure submitted in the Final Report of the PT, the verification performed with VERSUS used grid-to-point comparisons in order to compare gridded surface and upper-air model data to point observations. The selected NWP suite stations are distributed in the entire model domain and are presented in figure 2. Surface (SYNOP) and upper air (TEMP) observations were extracted from the MARS database and uploaded to the VERSUS system.

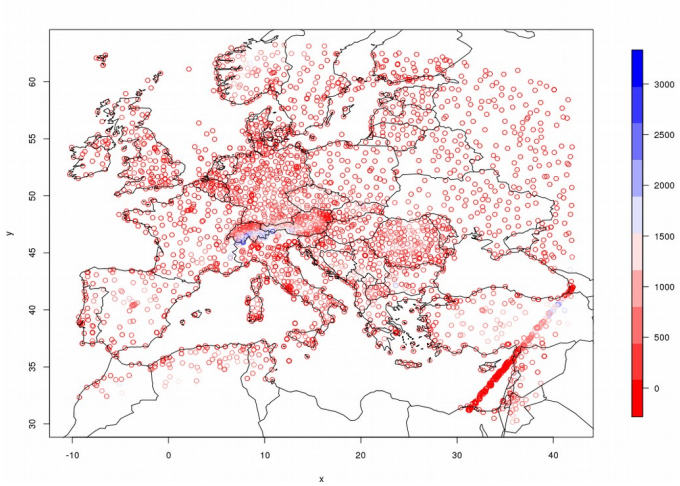

Fig. 2 Location of meteorological stations used for the verification.

Suspect observation values had been previously created for each parameter (observations are excluded when forecast-observation values are greater than specific limits) and included in the verification test in order to eliminate errors that are connected with observations.

For the improvement of the test suite and preparation on new model configurations to be tested, the operational model version 5.03 (lateral and boundary conditions from the ICON model) was registered in order to follow the evolution of model versions/tests and avoid confusions with previously tested versions (with IFS lateral and boundary conditions). For these tests, one model was taken into account: already operational version 5.03, which was assigned a new model id:

- ✓ COSMO 5.03 – 7km: process ID 113
- ✓ COSMO 5.03 – 2.8km: process ID 114

For the evaluation of the new 5.05 version of the COSMO model (on-going) against operational version 5.03, to avoid confusion with other tests previously performed, two models (2 resolutions each) were taken into account and assigned new model ids:

- ✓ COSMO 5.03 – 7km, double precision runs, forecast mode: process ID 117
- ✓ COSMO 5.03 – 2.8km, double precision runs, forecast mode: process ID 118
- ✓ COSMO 5.05 – 7km, double precision runs, forecast mode: process ID 115
- ✓ COSMO 5.05 – 2.8km, double precision runs, forecast mode: process ID 116

New front-ends (FEs) are registered for the tests, four for each model resolution for version 5.03 (process IDs 113 and 114), four for each model resolution for version 5.03 double precision runs - forecast mode and four for each model resolution for version 5.05 double precision runs - forecast mode (24 new FEs currently).

FEs were created separately because of the different interpolation methods used in each case: separate FEs for precipitation (interpolation method - 15 km radius), cloud cover (interpolation method - 30 km radius) and other surface parameters ((interpolation method – nearest grid point, 3D optimized) and separate FEs for the upper air data (interpolation method – nearest grid point). Original grib model outputs were split in smaller grib files in order to speed up the uploading process for the VERSUS system.

The verification modules for the tests are:

- ✓ **surface continuous parameters** (2mT, Dew Point T, WindSp, TCC, MSLP): BIAS, RMSE – up to 72 hours anticipation for COSMO-7km, up to 48 hours anticipation for COSMO-2.8km;
- ✓ **precipitation verification** (6h, 12h, 24h) 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 – up to 72 hours anticipation for COSMO-7km, up to 48 hours anticipation for COSMO-2.8km;
- ✓ **upper air verification** - T, RH, WindSp for selected pressure levels (250., 500., 700., 850., 925., 1000.): BIAS, MAE, RMSE – up to 72 hours anticipation for COSMO-7km, up to 30 hours anticipation for COSMO-2.8km.

<u>Verification Results</u>

As previously mentioned, the verification results for the two model versions were performed for the months of July and December 2017. A sample of the derived statistics is presented in figures 3 – 5. 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) which is currently being drafted. This technical report will include a detailed description of all steps, from the compilation of the new model version, model set-up configurations and the final production of the graphics for the statistical scores extracted.

While verification activities are still on-going, a summary of the current results is given below.

2m temperature differences for the winter season are significant (figure 3), with a small reduction of RMSE in the warm hours of the day after the implementation of model version 5.05 and a smaller underestimation of the values for this parameter, while the daily cycle of error is reduced. In the summer, this behavior is still present but to less extend. For the finer resolution (figure 4) however, there is no improvement during the winter, the daily cycle of errors is slightly reduced, but the underestimation in the night hours for the new version is a slightly higher one than for the 5.03 version. The 7 km resolution produces larger errors for both model versions (5.05, 5.03) in the winter season in comparison to the finer one.

Both COSMO 5.03 and COSMO 5.05 underestimate the values forecasted for 2m dew point temperature for the entire winter period. During winter, the 7 km model exhibits a smaller RMSE for the new version. For July (figure 5), there is an improvement in the degree of underestimation during the night with the new version and a small increase in the overestimation during warm hours with the 7 km version.

For mean sea level pressure, both model versions exhibit the same behavior during both periods analyzed, mainly overestimation of MSLP, especially in winter and also for this season an increase of error with lead time. COSMO 5.05 exhibits a significant improvement in ME and RMSE for the winter season for both resolutions (more with the 7km one).

With respect to 10 meter wind speed, behavior of NWP test results exhibit almost identical behaviour for both resolutions and both seasons. RMSE values show no change with the new version, while there is a clear improvement in the steadily overestimating behavior of wind speed values that is more obvious with the 7 km model resolution.

Total Could Cover is a parameter that exhibits small changes in the error magnitude or behaviour with the implementation of version 5.05. There is a significantly smaller bias of cloud cover for the new 2.8 km COSMO version in December. However, especially in December, the 2.8 km version shows an overall larger bias than the 7 km version.

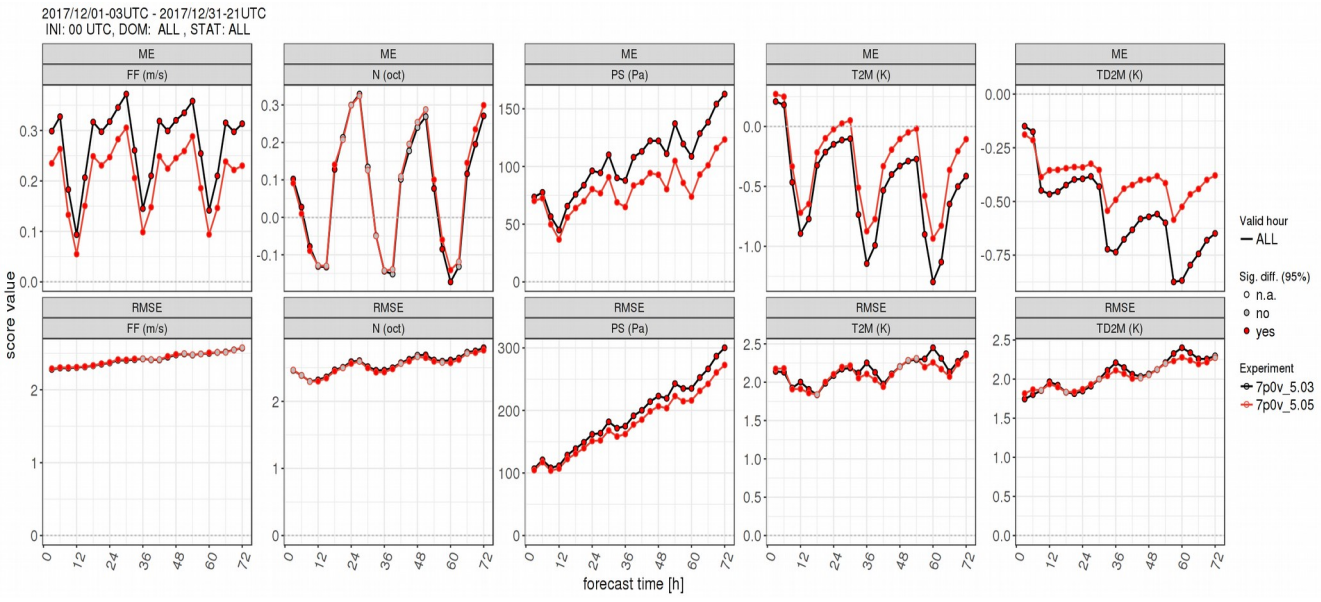

Fig. 3 COSMO-7km Continuous parameters verification results (00UTC run) - COSMO 5.05 and COSMO 5.03 mean error (ME) and root mean square error (RMSE) for December 2017. Red/gray filled dots indicate a significant/insignificant (95% level) difference of scores between the 2 model versions.

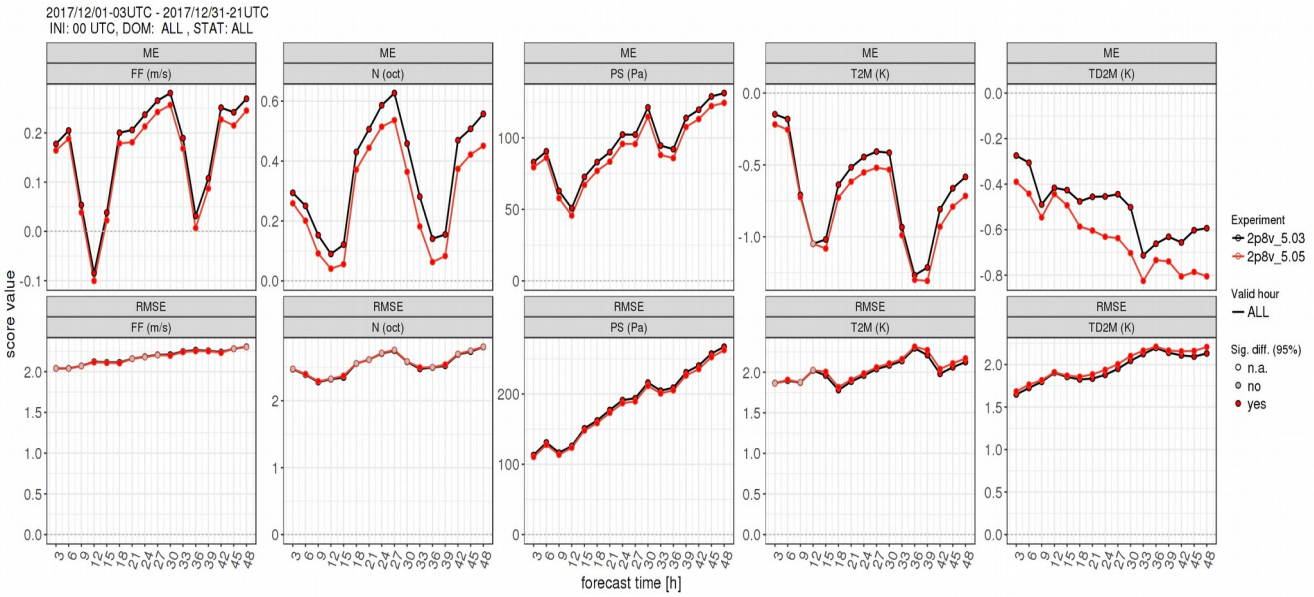

Fig. 4 COSMO-2.8km Continuous parameters verification results (00UTC run) - COSMO 5.05 and COSMO 5.03 mean error (ME) and root mean square error (RMSE) for December 2017.

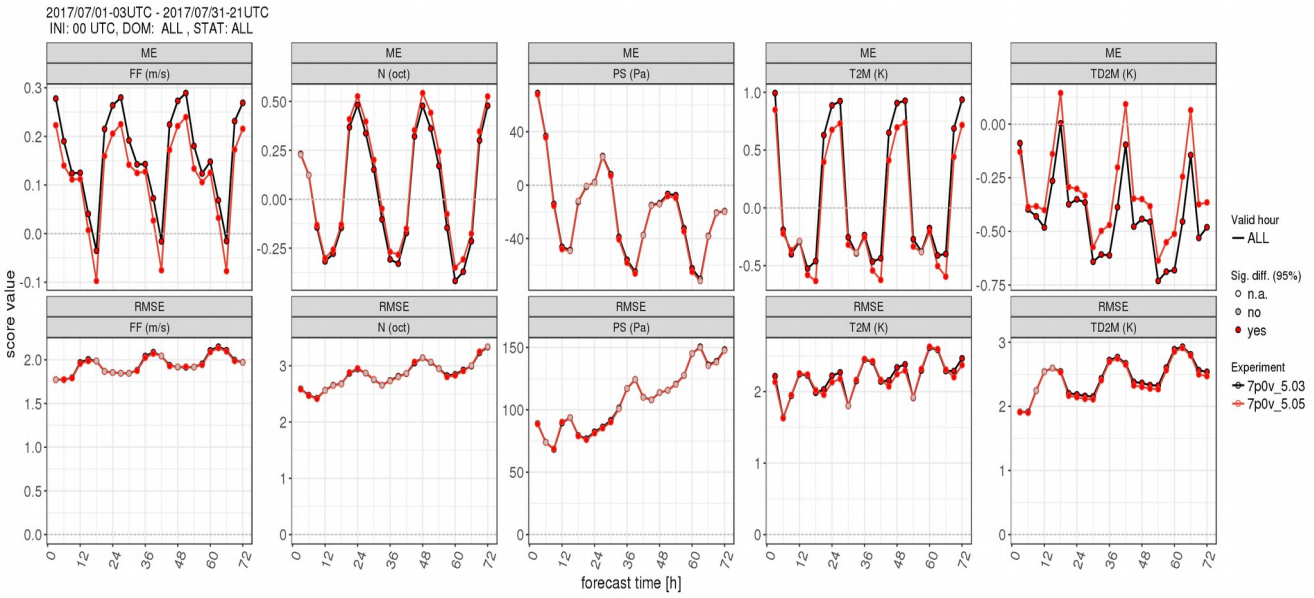

Fig. 5 COSMO-7km Continuous parameters verification results (00UTC run) - COSMO 5.05 and COSMO 5.03 mean error (ME) and root mean square error (RMSE) for July 2017.

Phase IV: Additional steps

As previous mentioned, testing of version 5.05 will also include single precision runs both in forecast and hindcast mode for the 7 km resolution and double precision runs in hindcast mode (both model resolutions). These activities will be performed after the completion of the verification activities for the double precision runs previously mentioned.

At the same time, a new verification software is currently being tested and will eventually replace VERSUS. RFDBK is an assortment of R libraries and scripts that utilizes Feedback Files (combined forecast and observation information on specific station points). Both Rfdbk and MEC (Model Equivalent Calculator for the preparation of the Feedback files), are installed on ecgate and tests are being performed to compare the results with those derived from VERSUS software.

List of publications/reports from the project with complete references

F. GOFA - "WG5 overview of Activities", The 19th COSMO General Meeting, Jerusalem, Israel, 11 - 14 September 2017

M. MILELLI - "WG6 overview", The 19th COSMO General Meeting, Jerusalem, Israel, 11 - 14 September 2017

A. MONTANI – "NWP Test Suite Status", ICCARUS 2018, Offenbach, Germany, 26 - 28 February 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 plans for the continuation of the project

(10 lines max)

Activities (including use of resources) to test the new official version (5.05) of the COSMO model prior to its release in summer 2018 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 (version 5.06 expected by the end of 2018).
Maintenance of the Test Suite.
Future versions of the COSMO model and future VERSUS releases to be installed as soon as they are available.
Replacing of VERSUS verification Tasks with MEC/Rfdbk as mentioned above.
Performing model evaluation for the next versions of the model.