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	2015
Project Title:	NWP meteorological test suite
Computer Project Account:	SPITRASP
Principal Investigator(s):	Adriano Raspanti (cn9) Amalia Iriza (roz)
Affiliation:	Ufficio Generale Spazio Aereo e Meteorologia (USAM-CNMCA) National Meteorological Administration (NMA)
Name of ECMWF scientist(s) collaborating to the project (if applicable)	Umberto Modigliani and his staff
Start date of the project:	2013
Expected end date:	2015

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)	1000000	356420.40	1000000	6354.24
Data storage capacity	(Gbytes)	180	20	180	0

Summary of project objectives

(10 lines max)

The aim of the project NWP Meteorological Test Suite is to build up a software environment to perform carefully-controlled and rigorous testing, including the calculation of verification statistics, for any COSMO (Consortium for Small-scale Modeling) model test-version. The individual objectives in this project include the COSMO Model set-up (Phase I), Configuration of test runs (Phase II), Model Output Verification (Phase III) and Additional steps (Phase IV). The NWP meteorological test suite platform will provide the COSMO community with standards against which the impacts of new developments in the model should be evaluated. Moreover, this test procedure could serve as a benchmark to monitor the progress of mesoscale forecast improvement through periodic re-testing as the COSMO system evolves.

Summary of problems encountered (if any)

(20 lines max)

With regards to the running and maintaining of the test suite, we experienced, especially at the end of 2014 some problems with the network and remote connection to the VERSUS virtual machine using NoMachine.

Problems were also encountered during the installation of a newer version of the VERSUS software on the ecgate virtual machine, with the web server.

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

In the framework of NWP Meteorological Test Suite, a platform was developed, on which present and future versions of the COSMO model are tested for their forecasting performance, within a well-defined framework. This platform will be the tool used for performing the tests that will upgrade a model test-version to a new release.

The test suite only addresses the statistical quality of a COSMO version (in this case version 5.1) in comparison with the previous one (version 5.0). The statistical measures are defined within the task itself; this concerns not only the type of scores to be used but also the array of parameters (e.g. 500 hPa geopotential, precipitation, 2m temperature). The comparison of the model versions for validation was carried out on a common domain. The new version of the model is considered validated or accepted if the set of verification results show a positive impact on the common domain or if the results are neutral.

1. TECHNICAL SPECIFICATIONS

ECMWF computer resources were used for the aim of this task both for simulation and for archiving purposes, by using billing units provided by the members as part of the SPITRASP special project previously registered. Version 5.1 of the COSMO model (7km horizontal resolution) was implemented, following the procedure presented in the Final Report of the priority task. Version 5.0 had been previously implemented for the first tests of the suite. The int2lm 2.0 version was used for the interpolation of initial and lateral boundary conditions (like for the previous COSMO 5.0 tests).

The directory structure and the archiving procedures for the new version of the COSMO model (5.1) followed the ones used for the previous versions (4.26 and 5.0). After the completion of the testing procedure, model outputs were transferred to the machine with the installed VERSUS software in order to perform the statistical analysis. The data (model outputs) obtained from the experiments are locally stored in the ECFS system.

The domain involved in calculation covers the COSMO countries and a good part of European Russia (see figure 1). Grid definition (total number of grid points is 383761) is as follows:

- 751x511 grid points
- 40 vertical levels
- rotated coordinates:
  - pol latitude = 40
  - pol longitude = -170
- coordinates of the lowest left corner
  - start latitude = -16.125
  - start longitude = -15.75

The forecast period of each daily run is 72 hours, on one daily cycle based on the 00UTC initializing data. The simulation was performed for one month in summer (July 2013) and one month in the winter season (January 2013), in total 2 months (the simulations for the COSMO 5.0 version were previously performed). The initial and boundary data are provided by the ECMWF IFS system:

- area=75./-25.0/25.0/70.0 (761x401 points)
- 137 vertical levels
- horizontal resolution 0.125/0.125
- from fc+0h to fc+72h every 3h ---> size approx. 11.6 GB

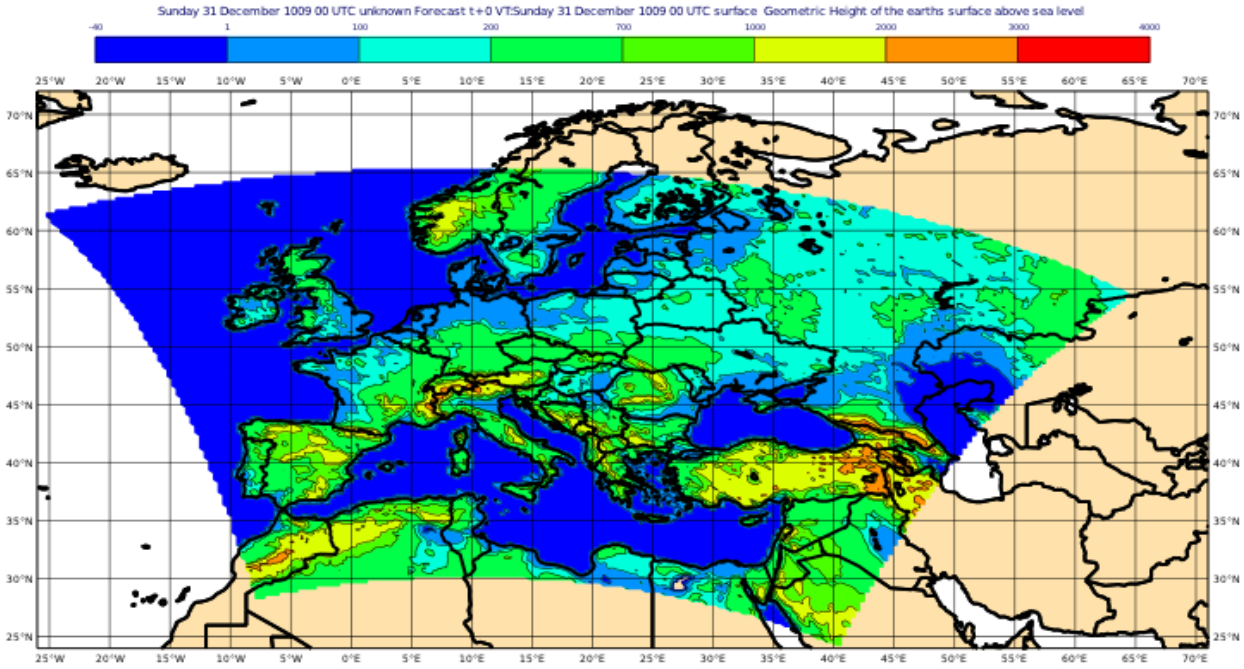


Fig. 1 Integration domain for the COSMO model (both versions) used for the current tests.

In the present configurations the costs of the suite are the following:

- Interpolation: about 81.5 BU per run (and takes approx. 8 min)
- COSMO-5.0: about 2350 BU per run (and takes approx. 29 min)
- COSMO-5.1: about 2284 BU per run (and takes approx. 28 min)
- total_tasks = 64 and node = 1 for int2lm
- total_tasks = 512 and node = 8 for COSMO

2. 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 stations are situated in an area covering -25/24/65/65 (W/S/E/N) and are around 3600 for the stratification that was used. Suspect observations values had been already created for each parameter (forecast-observation greater than a specific limit were excluded) and included in the verification test in order to eliminate errors that are connected with observations (table 1).

Table 1 Suspect observations values.

Parameter	**|Fcst-Obs| < value**
FF_10M	50. m/s
MSLP	2500. Pa
PS	2500. Pa
TD_2M	30. deg C
T_2M	30. deg C

The new model version was registered with the version number (COSMO-5.1), in order to follow the evolution of model versions/tests. For the present tests, two models were taken into account: 5.0 (operational) and 5.1 (new test version). Both models (5.0 and 5.1) have the same grid characteristics but they were each assigned **a different model id**.

There are three frontends (FE) registered for the new test version of the model and a separate FE for the upper air data for this version. Previous to the uploading phase, the original grib model outputs were split in smaller files, using the wgrib facility.

The verification modules for the current test were as follows:

- surface continuous parameters (2mT, Dew Point T, WindSp, TCC, MSLP): BIAS, RMSE;
- precipitation verification (6h, 12h, 24h) for selected thresholds (greater than0.2, 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;
- upper air verification - T, RH, WindSp for selected pressure levels (250.,500.,700.,850.,925.,1000.): BIAS, MAE, RMSE.

The steps that were followed are:

- Configuration of all standard surface verification tests
- Configuration of standard upper air verification tests
- Execution of above mentioned verifications in a batch mode
- Configuration of Cross model verification: interactively and batch mode
- Configuration of related graphics
-

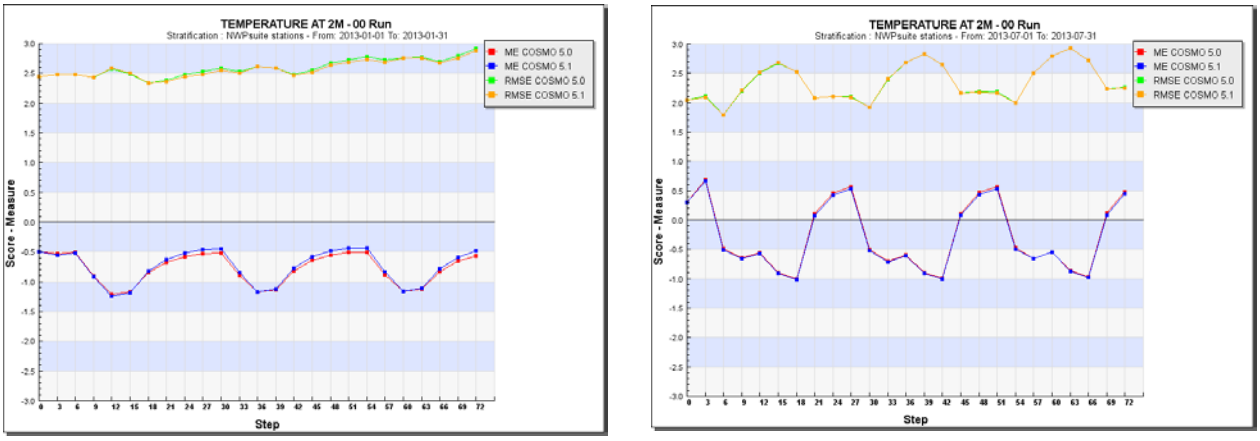
3. VERIFICATION RESULTS

As previously mentioned, the verifications for the two model versions were performed for January 2013 and July 2013. The results for model output verification obtained after running the VERSUS system (surface and upper air) are presented below. Overall, the statistics of the two versions of the model are very similar, with only some slight differences. Analytically:

- With respect to 2m temperature, COSMO 5.1 maintains the underestimation in the winter season for the entire forecast period with a profound daily cycle of the errors, as COSMO 5.0, but the statistics show slightly improved scores, especially for evening hours – forecast time 18 - 30 and 42 - 57 (figure 2). For the summer period again both models show a tendency to underestimate T2m values during the day and overestimate during the night. As before, the forecast of COSMO 5.1 is slightly improved compared to that of COSMO 5.0 for the night – anticipations 21 - 27 and 45 - 51 (figure 2).
- For 2m dew point temperature both COSMO 5.0 and COSMO 5.1 underestimate the values for the entire forecast period in January while for July both models underestimate values during the day but overestimate during the night. COSMO 5.1 exhibits an improvement in the forecast of this parameter for almost all forecast hours for winter, and between 21-24 hours anticipation and 45-48 hours anticipation for summer (figure 3).
- The forecast of COSMO 5.1 for mean sea level pressure exhibits the same behaviour as COSMO 5.0 during both periods analysed (overestimation, increasing RMSE with forecast lead time, especially in winter). However, the quality of the COSMO 5.1 forecast for this parameter is slightly worse exhibiting a higher overestimation (0.1-0.2Hpa) for almost all forecast hours (figure 4).
- The analysis of the scores for 10 meter wind speed show no significant differences between the two versions of the COSMO model (figure 5).
- For the forecast of precipitation (6h and 24h accumulation periods), the statistics of the two versions of the model are very similar (overestimation in small thresholds [>1mm] but underestimation of precipitation amounts for higher thresholds [<5mm], higher FAR and lower POD with increasing threshold) with some slight differences, mainly in the summer period (figures 6-7).

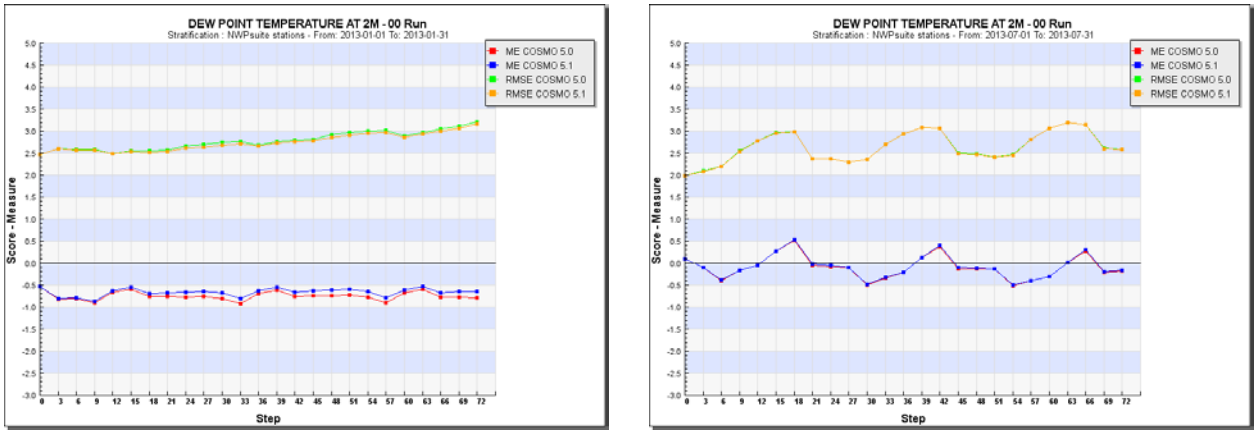
- The scores for the forecast of upper air parameters (relative humidity, temperature and wind speed) also show similar behaviour for both models, with no important differences (figures 8 – 9). The numerical difference of scores was calculated and those for RH are given in figure 23. The RMSE difference between the 5.0 and 5.1 versions of the model for Temp and Windspeed was inconsequential (the second decimal place), but was larger for RH (the first decimal place), with no steady inclination towards either of the two versions.

The graphics included in this report represent only a selection from the full range of statistical scores obtained for the comparison of COSMO 5.0 versus COSMO 5.1. The entire set of results can be retrieved from the VERSUS system at the ECMWF.



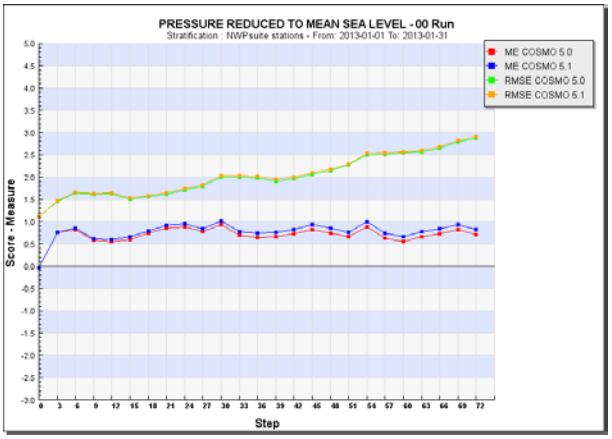
(a) (b)

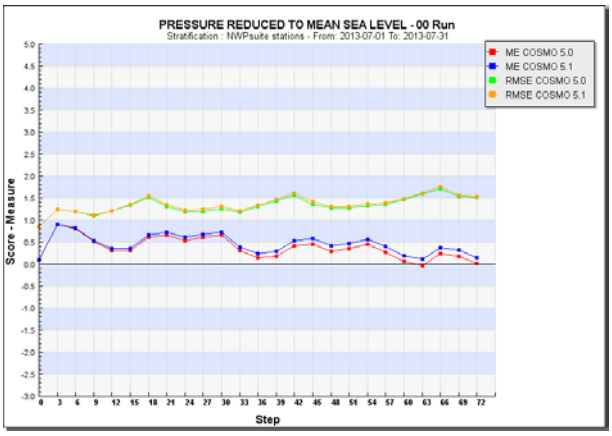
Fig. 2 Temperature at 2m verification results (00UTC run) - COSMO 5.1 and COSMO 5.00 mean error (ME) and root mean square error (RMSE) for: (a) January 2013 (b) July 2013



(a) (b)

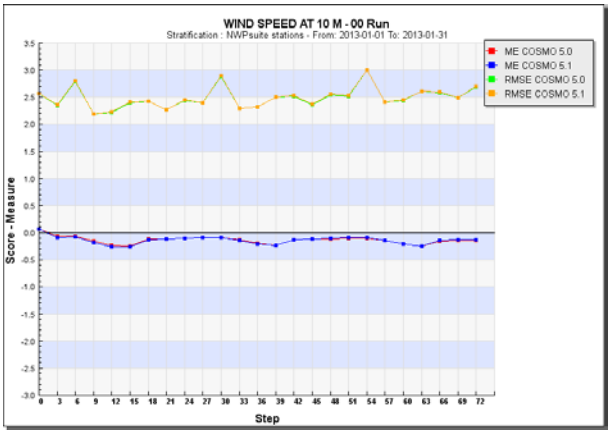
Fig. 3 Dew point temperature at 2m verification results (00UTC run) - COSMO 5.1 and COSMO 5.00 ME and RMSE for: (a) January 2013 (b) July 2013

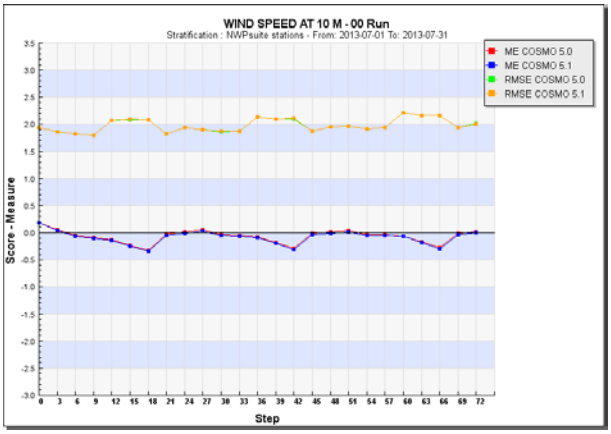




(a) (b)

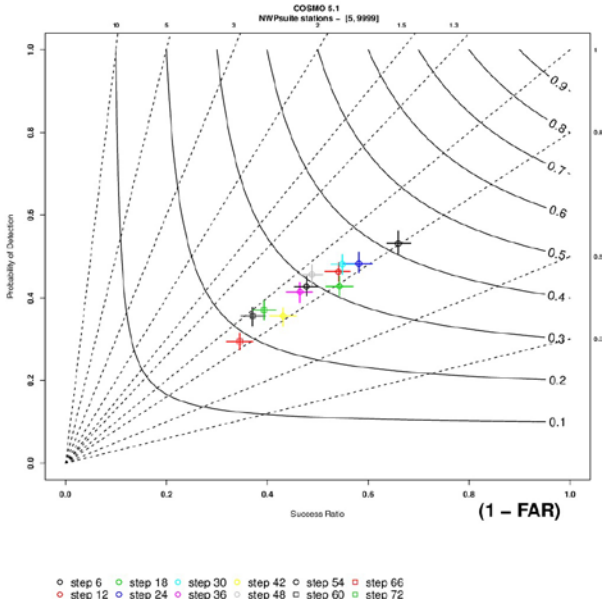
Fig. 4 Pressure reduced to mean sea level verification results (00UTC run) - COSMO 5.1 and COSMO 5.00 ME and RMSE for: (a) January 2013 (b) July 2013

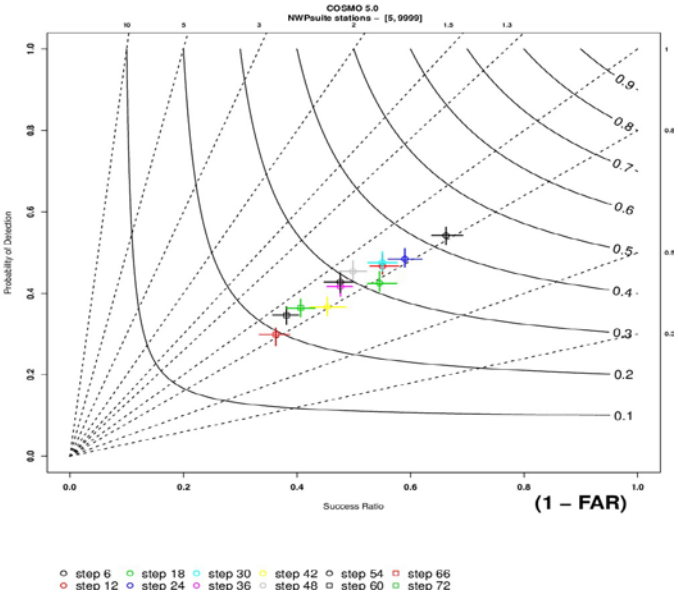




(a) (b)

Fig. 5 Wind Speed at 10 m verification results (00UTC run) - COSMO 5.1 and COSMO 5.00 ME and RMSE for: (a) January 2013 (b) July 2013





(a) (b)

Fig. 6 6h precipitation > 5mm verification results (00UTC run) – January 2013, (1-FAR) for: (a) COSMO 5.1 (b) COSMO 5.00

Index: ETS (ID: 12)
From: 2013-01-01 To: 2013-01-31

Fct Time	Threshold	Sample	vs. 5.0	vs. 5.1	Diff
24	0.2	39712	0.43	0.44	-0.01
24	1	39712	0.51	0.51	0
24	5	39712	0.46	0.46	0
24	10	39712	0.42	0.42	0
24	20	39712	0.32	0.32	0
48	0.2	38825	0.36	0.38	-0.02
48	1	38825	0.44	0.45	-0.01
48	5	38825	0.41	0.4	0.01
48	10	38825	0.36	0.36	0
48	20	38825	0.25	0.25	0
72	0.2	36948	0.33	0.34	-0.01
72	1	36948	0.4	0.41	-0.01
72	5	36948	0.33	0.32	0.01
72	10	36948	0.28	0.28	0
72	20	36948	0.23	0.22	0.01

Index: ETS (ID: 12)
From: 2013-07-01 To: 2013-07-31

Fct Time	Threshold	Sample	vs. 5.0	vs. 5.1	Diff
24	0.2	39462	0.35	0.35	0
24	1	39462	0.37	0.36	0.01
24	5	39462	0.3	0.29	0.01
24	10	39462	0.26	0.26	0
24	20	39462	0.23	0.22	0.01
48	0.2	38141	0.36	0.36	0
48	1	38141	0.36	0.35	0.01
48	5	38141	0.27	0.27	0
48	10	38141	0.22	0.22	0
48	20	38141	0.16	0.16	0
72	0.2	36826	0.32	0.32	0
72	1	36826	0.31	0.31	0
72	5	36826	0.23	0.22	0.01
72	10	36826	0.17	0.15	0.02
72	20	36826	0.1	0.07	0.03

Index: FAR (ID: 7)
From: 2013-01-01 To: 2013-01-31

Fct Time	Threshold	Sample	vs. 5.0	vs. 5.1	Diff
24	0.2	39712	0.22	0.21	0.01
24	1	39712	0.24	0.23	0.01
24	5	39712	0.33	0.33	0
24	10	39712	0.36	0.36	0
24	20	39712	0.42	0.42	0
48	0.2	38825	0.27	0.26	0.01
48	1	38825	0.3	0.29	0.01
48	5	38825	0.39	0.4	-0.01
48	10	38825	0.44	0.45	-0.01
48	20	38825	0.53	0.54	-0.01
72	0.2	36948	0.28	0.27	0.01
72	1	36948	0.31	0.31	0
72	5	36948	0.46	0.47	-0.01
72	10	36948	0.52	0.52	0
72	20	36948	0.59	0.6	-0.01

Index: FAR (ID: 7)
From: 2013-07-01 To: 2013-07-31

Fct Time	Threshold	Sample	vs. 5.0	vs. 5.1	Diff
24	0.2	39462	0.38	0.38	0
24	1	39462	0.4	0.4	0
24	5	39462	0.44	0.43	0.01
24	10	39462	0.46	0.45	0.01
24	20	39462	0.46	0.46	0
48	0.2	38141	0.38	0.38	0
48	1	38141	0.4	0.4	0
48	5	38141	0.48	0.48	0
48	10	38141	0.52	0.52	0
48	20	38141	0.56	0.55	0.01
72	0.2	36826	0.4	0.4	0
72	1	36826	0.44	0.44	0
72	5	36826	0.52	0.53	-0.01
72	10	36826	0.6	0.61	-0.01
72	20	36826	0.7	0.75	-0.05

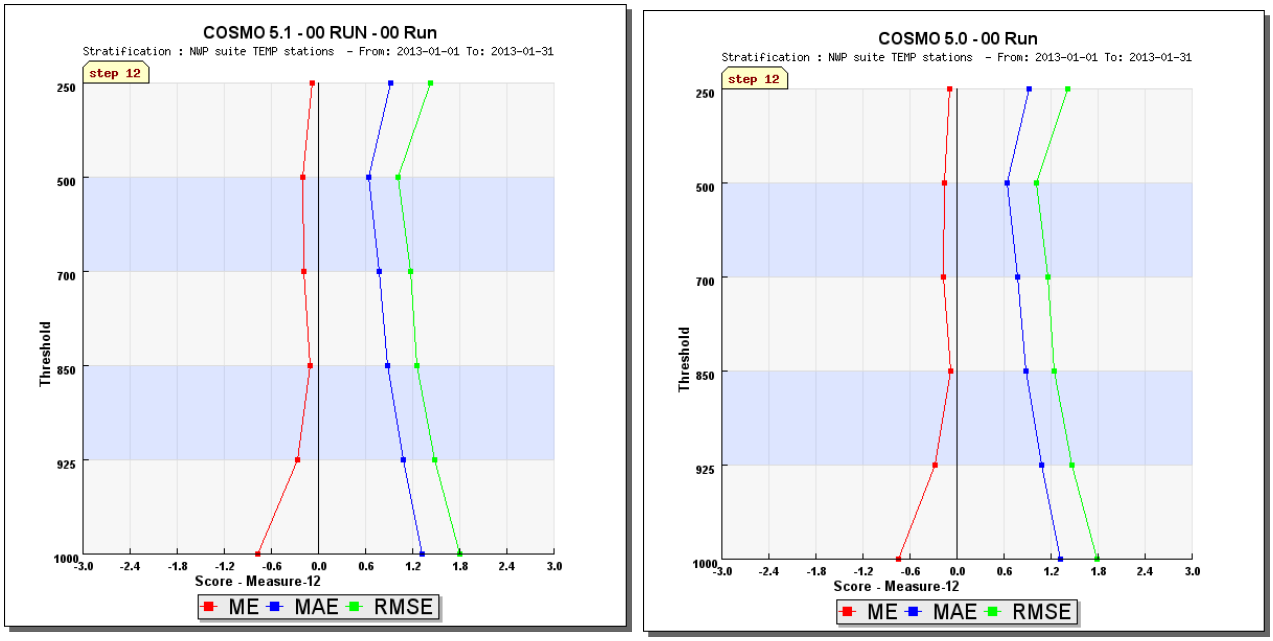
Index: POD (ID: 4)
From: 2013-01-01 To: 2013-01-31

Fct Time	Threshold	Sample	vs. 5.0	vs. 5.1	Diff
24	0.2	39712	0.87	0.86	0.01
24	1	39712	0.85	0.84	0.01
24	5	39712	0.7	0.7	0
24	10	39712	0.59	0.59	0
24	20	39712	0.43	0.43	0
48	0.2	38825	0.88	0.87	0.01
48	1	38825	0.84	0.84	0
48	5	38825	0.7	0.69	0.01
48	10	38825	0.57	0.57	0
48	20	38825	0.37	0.37	0
72	0.2	36948	0.85	0.85	0
72	1	36948	0.81	0.81	0
72	5	36948	0.62	0.62	0
72	10	36948	0.48	0.49	-0.01
72	20	36948	0.36	0.35	0.01

Index: POD (ID: 6)
From: 2013-07-01 To: 2013-07-31

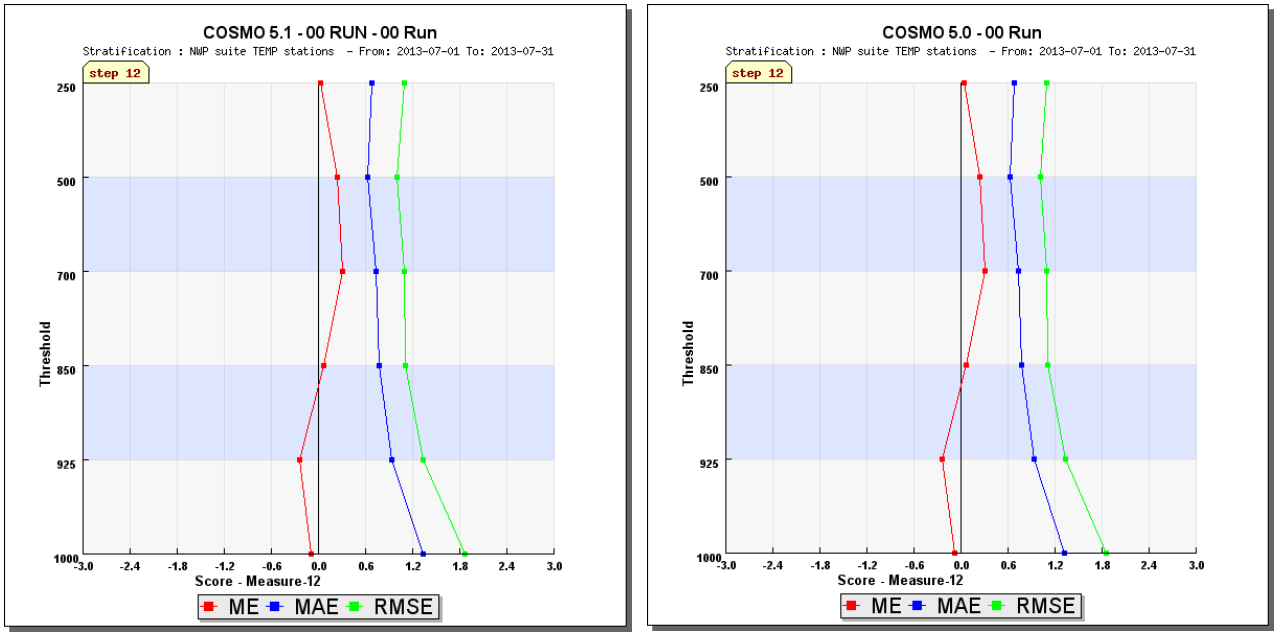
Fct Time	Threshold	Sample	vs. 5.0	vs. 5.1	Diff
24	0.2	39462	0.68	0.68	0
24	1	39462	0.63	0.63	0
24	5	39462	0.45	0.44	0.01
24	10	39462	0.36	0.35	0.01
24	20	39462	0.29	0.28	0.01
48	0.2	38141	0.68	0.68	0
48	1	38141	0.62	0.61	0.01
48	5	38141	0.42	0.41	0.01
48	10	38141	0.32	0.32	0
48	20	38141	0.21	0.21	0
72	0.2	36826	0.63	0.63	0
72	1	36826	0.56	0.55	0.01
72	5	36826	0.37	0.36	0.01
72	10	36826	0.26	0.23	0.03
72	20	36826	0.13	0.1	0.03

Fig. 7 24h precipitation for all thresholds ETS, FAR, POD differences 5.0 vs 5.1: Jan (left) – July (right) (red: worsening, green: improvement, yellow: neutral)



(a) (b)

Fig. 8 Upper air verification for January 2013 - 12 UTC: Temperature
(a) COSMO 5.1 (b) COSMO 5.0



(a) (b)

Fig. 9 Upper air verification for July 2013 - 12 UTC: Temperature
(a) COSMO 5.1 (b) COSMO 5.0

List of publications/reports from the project with complete references

A detailed report regarding the comparison of the COSMO 5.0 versus COSMO 5.1 versions using this platform were submitted to the COSMO Steering Committee. The technical report includes a detailed description of all steps, from the compilation of the new COSMO 5.1 model testing version to the final production of the graphics for the statistical scores extracted.

Summary of plans for the continuation of the project

- The main activity (including use of resources) will be concentrated in the second part of the year, when two more releases of the COSMO model are anticipated.
- Maintenance of the Test Suite
- Future versions of the COSMO model need to be installed as soon as they are available
- Maintenance of Versus - Future VERSUS releases will have to be installed as soon as they are available
- Running of the NWP test suite for subsequent versions of the model.
- Performing model evaluation for the next versions of the model.