# SPECIAL PROJECT FINAL REPORT

All the following mandatory information needs to be provided.

| | |
|---|---|
| **Project Title:** | **NWP meteorological test suite** |
| **Computer Project Account:** | **spitrasp** |
| **Start Year - End Year :** | **2013 - 2015** |
| **Principal Investigator(s)** | Adriano Raspanti (cn9) [1]<br>Amalia Iriza (roz) [2] |
| **Affiliation/Address:** | Ufficio Generale Spazio Aereo e Meteorologia (USAM-CNMCA) / Aeroporto M. De Barnadi, Via di Pratica di Mare 45, 00040, Promezia (RM), Italy [1]<br><br>National Meteorological Administration (NMA) / Sos. Bucuresti-Ploiesti, nr. 97, Bucuresti, 013686, Romania [2] |
| **Other Researchers (Name/Affiliation):** | Flora Gofa (emo) / Hellenic National Meteorological Service (HNMS)<br>Angela Celozzi (cmo) / Centro Nazionale di Meteorologia e Climatologia Aeronautica (USAM-CNMCA)<br>Rodica Dumitrache (rod) / National Meteorological Administration (NMA)<br>Andrea Montani (itm) / Agenzia regionale per la prevenzione, l'ambiente e l'energia (ARPAE-SIMC) |

The following should cover the entire project duration.

## 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 you encountered any problems of a more technical nature, please describe them here. )

With regards to the running and maintenance the SMS suite, we experienced, especially in May and June 2014, some problems with the network, with frozen windows every now and then. This made more difficult to work remotely on ECMWF machines. Also at the end of 2014 we experienced some problems with the network and remote connection to the VERSUS virtual machine using NoMachine.
With regards to the installation of VERSUS and model output verification we encountered a few problems:

- restricted rights to some directories and configuration files where certain components of VERSUS are installed
- setup of the virtual machine on ecgate (actually 8Gb and 2 CPU's)
- remote web access issues and problems with connection to the virtual machine
- very slow connection with the GUI (fixed using NoMachine)
- installation of firefox for multiple users (fixed)
- problems with the registration of the stations stratification (3600 stations-fixed)
- memory allocation of some important variables for VERSUS performance
- size of grib files crucial for uploading speed and dependant on the number of stations (need for splitting in smaller units, fixed)
- allocation of data (and log files) on another file system than the VERSUS one (use of symbolic links)
- problems with permissions read/write resulted since the installation of VERSUS patch 4.2 (August 2015)
- problems with access permission for stop/start pending jobs (always to be performed through communication with ECMWF personnel)

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

## Experience with the Special Project framework

(Please let us know about your experience with administrative aspects like the application procedure, progress reporting etc.)

We consider that collaboration with your administrative and support team was very good. From our point of view, progress report procedures are clear and helpful.

## Summary of results

(This section should comprise up to 10 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 will be 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 versions 5.0, 5.01 (5.1), 5.03 (5.3) - in comparison with the previous one - versions 4.26, 5.0 and 5.01 (5.1) respectively. 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 was considered validated or accepted if the set of verification results showed a positive impact on the common domain or if the results are neutral.

### 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.
Several versions of the COSMO model at 7km resolution (versions 4.26, 5.00, 5.01 (5.1) and 5.03 (5.3)) were implemented and used for these tests, following the operational resolution in most meteorological services. The INT2LM 2.0 version was used for the interpolation of initial and lateral boundary conditions.

Main activities of this phase consisted of the preparation of the model installation. For this purpose the following steps were performed:

- compilation all the necessary external libraries and tools for file managing
- compilation of the interpolation program (INT2LM)
- compilation of each COSMO version tested

The directory structure and the archiving procedures were set similarly for all model versions . After the completion of each 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 application running the COSMO experiments comprises several tasks with interdependencies between them. The workflow was organised and maintained through the **xcdp/sms** suite (Xbox Community Developer Program/Supervisor Monitoring Scheduler - ECMWF's monitoring and scheduling software). The COSMO numerical tests were performed on the IBM platform for model versions 4.26, 5.00 and 5.01 (5.1) and on the Cray platform for model version 5.03 (5.3).

### Model Set-up and Configuration of Test Runs

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

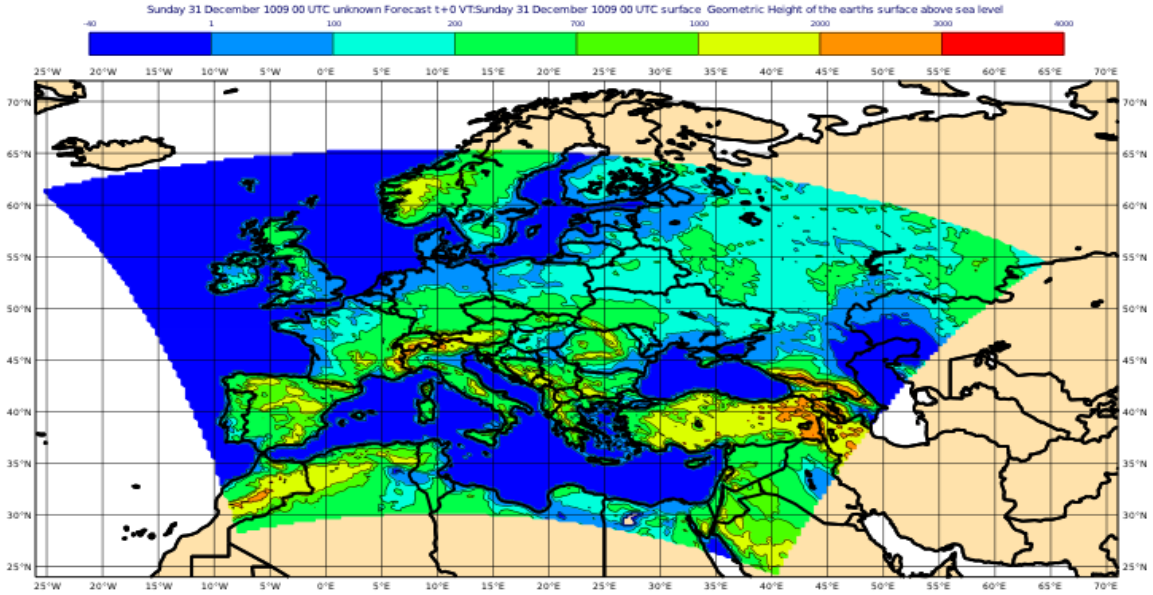



Fig. 1 Integration domain for the COSMO model (all versions) used for the tests.

The ECMWF Special Project SPITRASP had an allocation of 400.000 units for 2013, 1.000.000 units for 2014 and 1.000.000 units for 2015. Data storage capacity allocated was of 80GB for 2013, 180GB for 2014 and 180 GB for 2015. Depending on the model configuration, the costs of the suite were the following:

*COSMO versions 4.26, 5.0 and 5.01 (5.1) - on IBM*

- Interpolation for COSMO-4.26: about 80.0 BU per run (takes approx. 8 min)
- Interpolation for COSMO-5.0, COSMO-5.01 (5.1): about 81.5 BU per run (takes approx. 8 min)
- COSMO-4.26: about 2434 BU per run (takes approx. 30 min)
- COSMO-5.0: about 2350 BU per run (takes approx. 29 min)
- COSMO-5.01 (5.1): about 2284 BU per run (takes approx. 28 min)
- total_tasks = 64 and node = 1 for INT2LM
- total_tasks = 512 and node = 8 for COSMO-4.26, COSMO-5.0 and COSMO-5.01 (5.1)

*COSMO version 5.03 (5.3) - on Cray*

- Interpolation for COSMO-5.03 (5.3): about 40 BU per run (takes approx. 6 min)
- COSMO-5.03 (5.3): about 3600 BU per run (takes approx. 28 min)
- EC_total_tasks = 24 and EC_nodes = 1 for INT2LM
- EC_total_tasks = 480 and EC_nodes = 20 for COSMO-5.3 (5.3)



- total_tasks = 64 and node = 1 for int2lm
- total_tasks = 512 and node = 8 for COSMO

**Model Output Verification**

For the model verification VERification System Unified Survey (VERSUS) software was implemented at ECMWF. VERSUS is LAMP open source software. LAMP is an acronym referring to the first letters of Linux (operating system), Apache HTTP Server, MySQL (database software) and PHP, principal components to build a viable general purpose web server. The solution that was recommended that could fulfil the needs of the software was to install it on a virtual machine that was based on ECGATE linux system. As ECMWF personnel had no prior experience in hosting external software, there were many technical and security issues that had to be overcome in order to complete the installation of all necessary software that is linked with VERSUS system.The final product allows the remote access of each user not only to the virtual machine but also the execution of the verification suite through the web graphic interface.

After the successful installation of the main packages several activities devoted to the proper setup of the system were performed until a final configuration was achieved. The preparatory work for the installation and communication with ECMWF started already in 2013, while the actual installation of the virtual machine and the system started in February 2014. The installation of the main package was considered completed at the end of May 2014. Subsequently, the updates VERSUS 3.1, 3.2, 4.2 and 5.0 were successfully installed.

The VERSUS virtual machine was setup with 2 CPU's and a total RAM of 8Gb. The tests executed have showed that the performance (in terms of stability of the system and speed) are satisfactory. The VERSUS virtual machine was setup in order to have a different disk space for DB (quite huge) and another one even bigger to store the data (grib and bufr). A dedicated space was created, on which the usual directories tree that VERSUS requests are created. As VERSUS requests to have grib and bufr files in the same Filesystem but the space was too small, symbolic links were created to the directory where the data actually are stored.

To connect to the VERSUS virtual machine (ms-versus) the **NoMachine** software was used. This feature allows multiple GUI users. Other connections to the VERSUS virtual machine can be established simply using "ssh versus@ms-versus". Only the user "versus" has the access to the VERSUS machine (apart from root, available only for ECMWF). This is accessible from the usual Ecgate cluster only by the users that at the moment are allowed without any password (other users needs password, available at the moment only for ECMWF): Adriano Raspanti (cn9), Flora Gofa (emo), Amalia Iriza (roz), Andrea Montani (itm) and Angela Celozzi (cmo).

In order to perform the verifications, the VERSUS sistem was set for the ingestion of all necessary data (model and observations). 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. The observations for the predefined periods of testing were uploaded in the DB using the world-wide WMO flatfile. Suspect observations values were 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 observation 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 |

For each NWP model version, the operational one and the new test version, the registration was done with the version number it has, in order to follow the evolution of model versions/tests. Both models have the same grid characteristics but as it is a prerequisite of the software, they were each assigned a different model id. For each model tested, the new model is registered using the same topography file but with a new model-id code that is in the namelist of the model.

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. Due to the large size of the forecast data, the original grib model outputs were split in smaller files, using the wgrib facility. This action is performed before the uploading phase, in order to speed up the latter.

The verification modules for each of the tested model versions were as follows:

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

For the model output verification, the following steps were performed:

- Registration of the COSMO model versions
- Configuration of separate FEs for each COSMO version
- Acquisition of forecast data
- 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

**Verification Results**

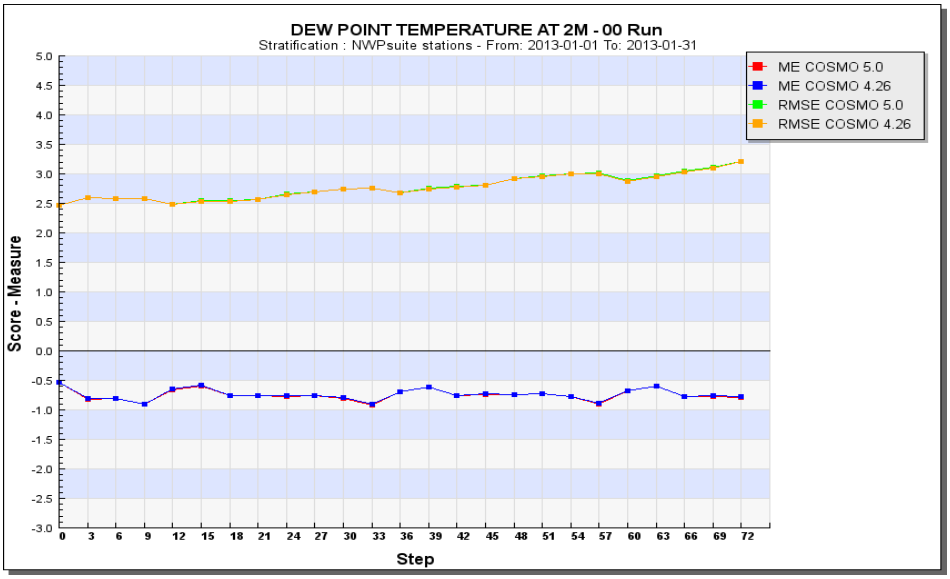
As previously mentioned, the verifications for the two model versions were performed for January 2013 and July 2013. The verifications were performed for the following COSMO versions:

- COSMO-4.26 against COSMO-5.0
- COSMO-5.0 against COSMO-5.01 (5.1)
- COSMO-5.01 (5.1) against COSMO-5.03 (5.3)
- tests for COSMO-5.03 (5.3) against COSMO-5.04a are currently undergoing

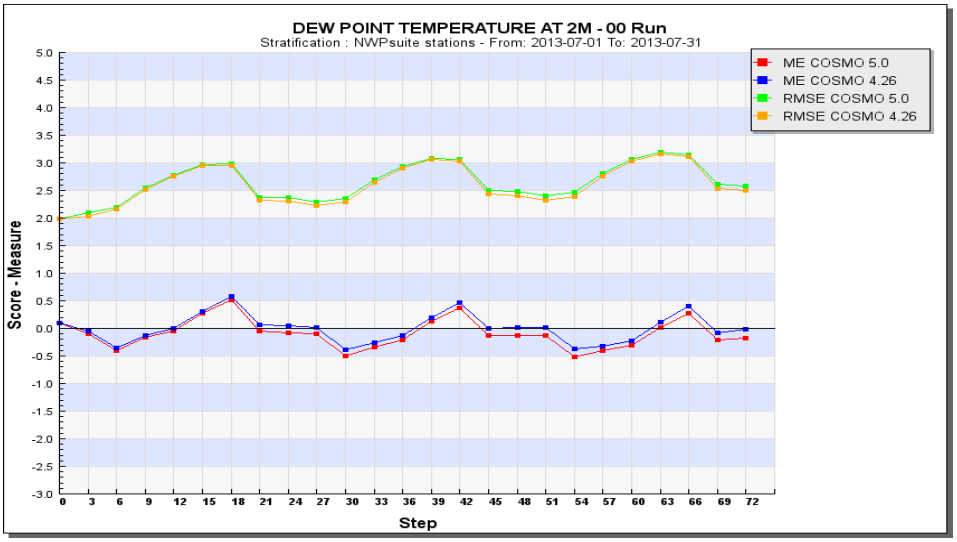
The graphics included in this report represent only a selection from the full range of statistical scores obtained for the comparison of the COSMO versions mentioned above. The entire set of results can be retrieved from the VERSUS system at the ECMWF. Some of the results for model output verification obtained after running the VERSUS system are presented below.

<u>COSMO-4.26 against COSMO-5.0</u>

The statistics of the two versions of the model were almost identical with some slight differences mainly for the summer period (e.g. figure 2).




(a)




(b)

Fig. 2 Dew point temperature at 2m verification results (00UTC run) – COSMO-4.26 and COSMO-5.0 ME and RMSE for: (a) January 2013 (b) July 2013

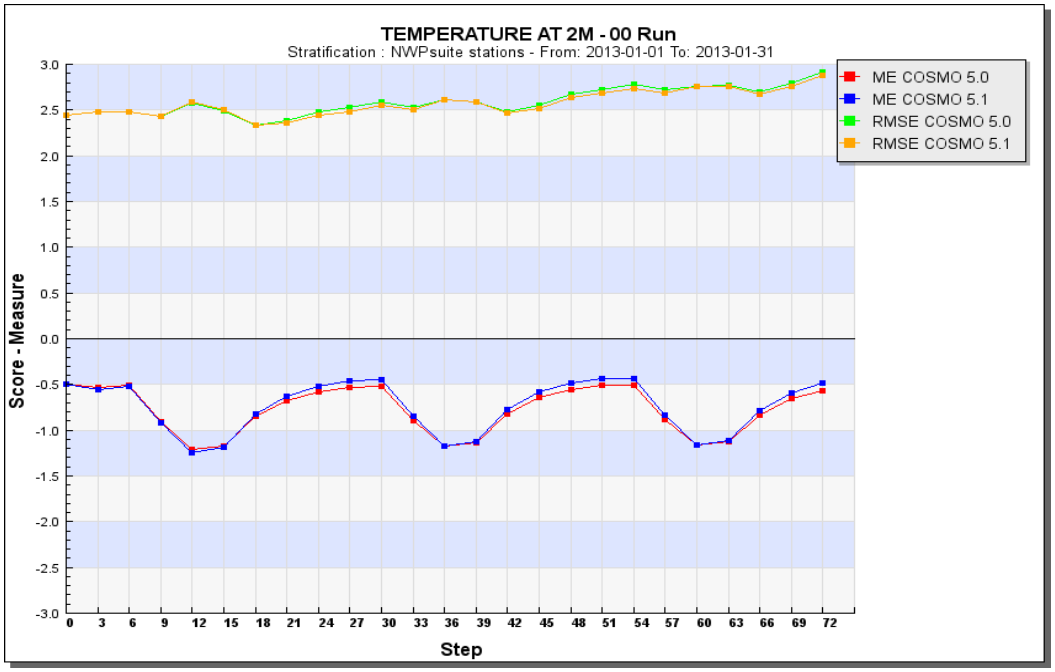
COSMO-5.0 against COSMO-5.01 (5.1)

Overall, the statistics of the two versions of the model were very similar, with only some slight differences. With respect to 2m temperature, COSMO-5.01 (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. 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-5.01 (5.1) is slightly improved compared to that of COSMO-5.0 for the night – anticipations 21 - 27 and 45 – 51 (figure 3).

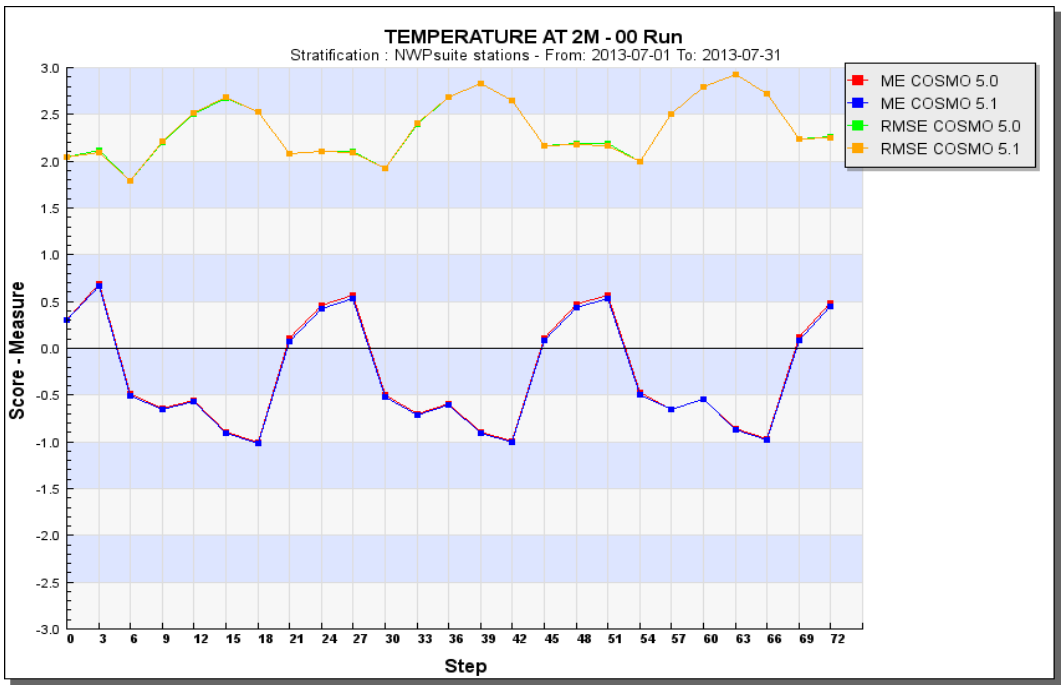
For the forecast of precipitation (6h and 24h accumulation periods), the statistics of the two versions of the model are very similar (overestimation in small thresholds [>1mm] but underestimation of precipitation amounts for higher thresholds [<5mm], higher FAR and lower POD with increasing threshold) with some slight differences, mainly in the summer period (figure 4).

The analysis of the scores for 10 meter wind speed show no significant differences between the two versions of the COSMO model.

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. The RMSE difference between the 5.0 and 5.01 (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 (e.g. figure 5).




(a)




(b)

Fig. 3 Temperature at 2m verification results (00UTC run) – COSMO-5.01 (5.1) and COSMO-5.0 mean error (ME) and root mean square error (RMSE) for: (a) January 2013 (b) July 2013

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 | 38325 | 0.27 | 0.26 | 0.01 |
| 48 | 1 | 38325 | 0.3 | 0.29 | 0.01 |
| 48 | 5 | 38325 | 0.39 | 0.4 | -0.01 |
| 48 | 10 | 38325 | 0.44 | 0.45 | -0.01 |
| 48 | 20 | 38325 | 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 |

Fig. 4 24h precipitation for all thresholds FAR differences 5.0 vs 5.01 (5.1): Jan(left)-July(right)
(red:worsening, green:improvement, yellow:neutral)

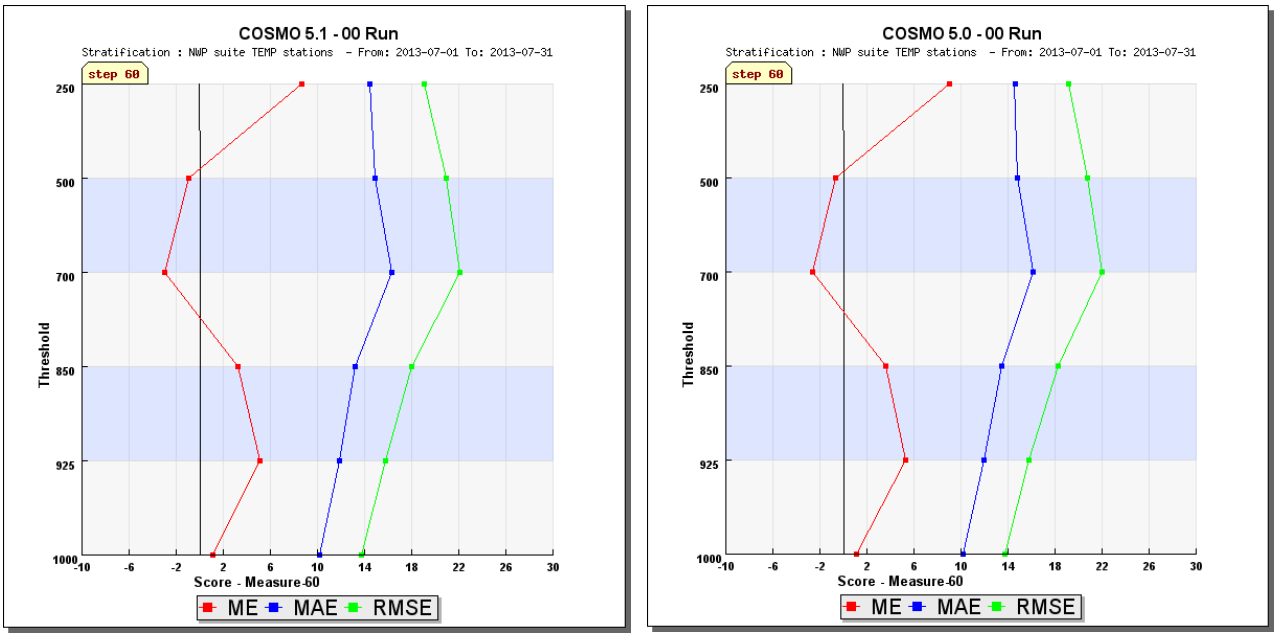



Fig. 5 Upper air verification for July 2013 - 60 UTC: Relative humidity
COSMO-5.01 (5.1) (left) / COSMO-5.0 (right)

COSMO-5.01 (5.1) against COSMO-5.03 (5.3)

As for the two verifications performed previously, the statistics of the two model versions 5.01 (5.1) and 5.03 (5.3) were quite similar, with some differences, presented below.

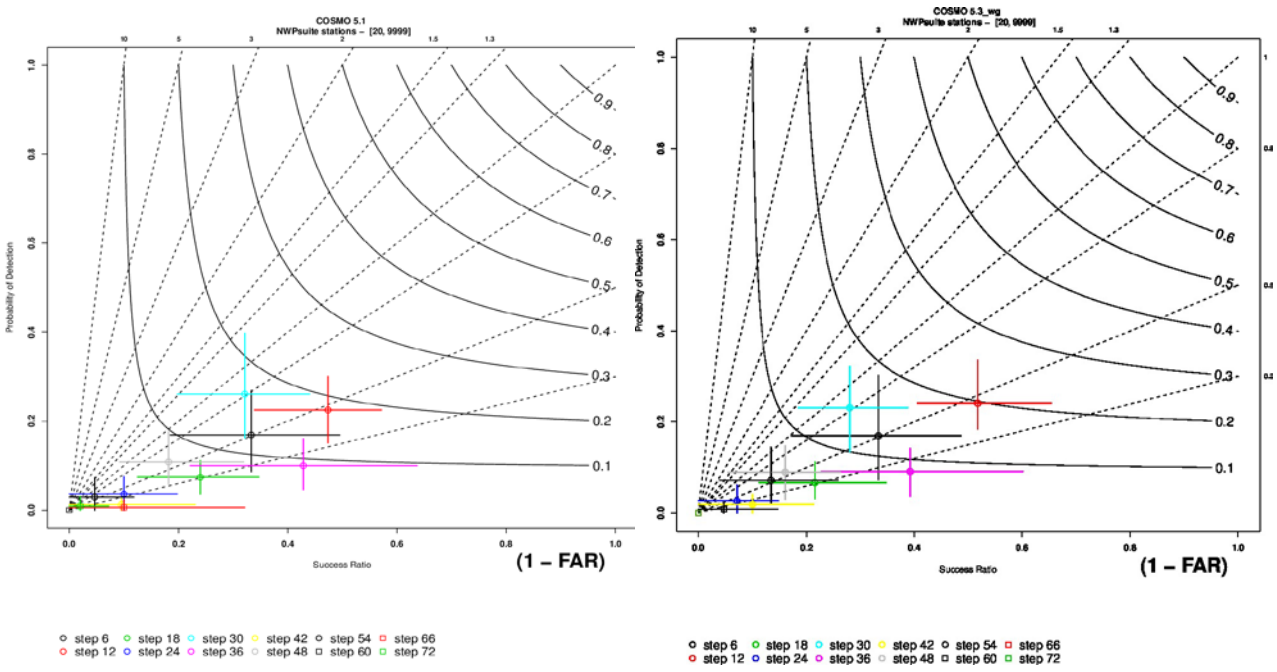
For the forecast of precipitation (6h and 24h accumulation periods), the statistics of the two versions of the model are very similar (overestimation in small thresholds [>1mm] but underestimation of precipitation amounts for higher thresholds [<5mm], higher FAR and lower POD with increasing threshold) with some insignificant differences mainly associated with False Alarm Rate score (e.g. figure 6).

With respect to 10 meter wind speed, mean error values for the winter period are worsened in the 5.03 (5.3) version but this is not noticeable in the summer period. Overall the comparison of scores shows neutral impact resulting from the introduction of the new version.

2m temperature differences for the winter season are insignificant, with both models underestimating values forecasted for the entire period, with a profound daily cycle of the errors. For the summer period again both models exhibit a tendency to underestimate T2m values during the day and overestimate during the night. However, looking at the numerical differences, the forecast of COSMO-5.03 (5.3) is slightly worsened compared to that of COSMO-5.01 (5.1) for most of the day in the winter period while slightly improved or neutral for almost all timesteps for the summer one. It has to be noted however that the difference of the statistical scores is marginal (apparent in the 3rd decimal place).

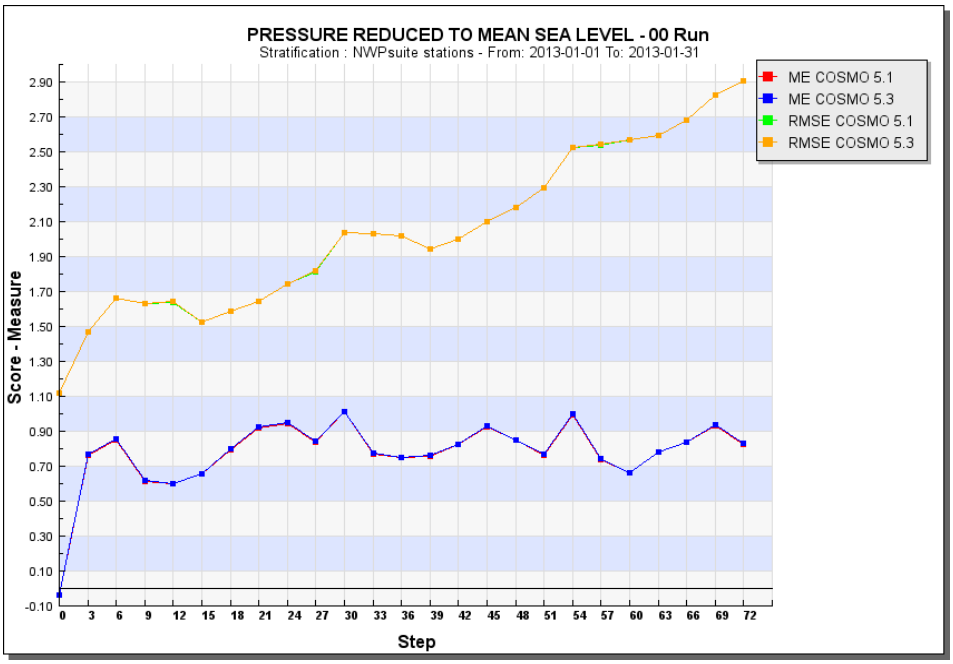
For mean sea level pressure, both model versions exhibit the same behaviour during both periods analysed, mainly overestimation and increasing RMSE with forecast lead time, especially in winter. As in the case of the previous parameters analysed, the model shows no improvement for the winter period. However, again for the summer period, COSMO-5.03 (5.3) displays a small improvement, while the amplitude of errors is slightly reduced during most of the forecast intervals (figure 7).

The scores for the forecast of upper air parameters (relative humidity, temperature and wind speed) also show similar behaviour for both. The numerical difference of scores was also calculated. Temperature comparison of ME and RMSE for the two model versions gave insignificant differences (lower than 0.02 degrees). Similar is the outcome from Wind Speed performance comparison, while it is indicated for the winter period there a slight negative impact from the 5.03 (5.3) implementation in the higher atmospheric levels (>500mb). Similarly for RH, the differences were minimal with no steady inclination towards either of the two versions (figure 8).



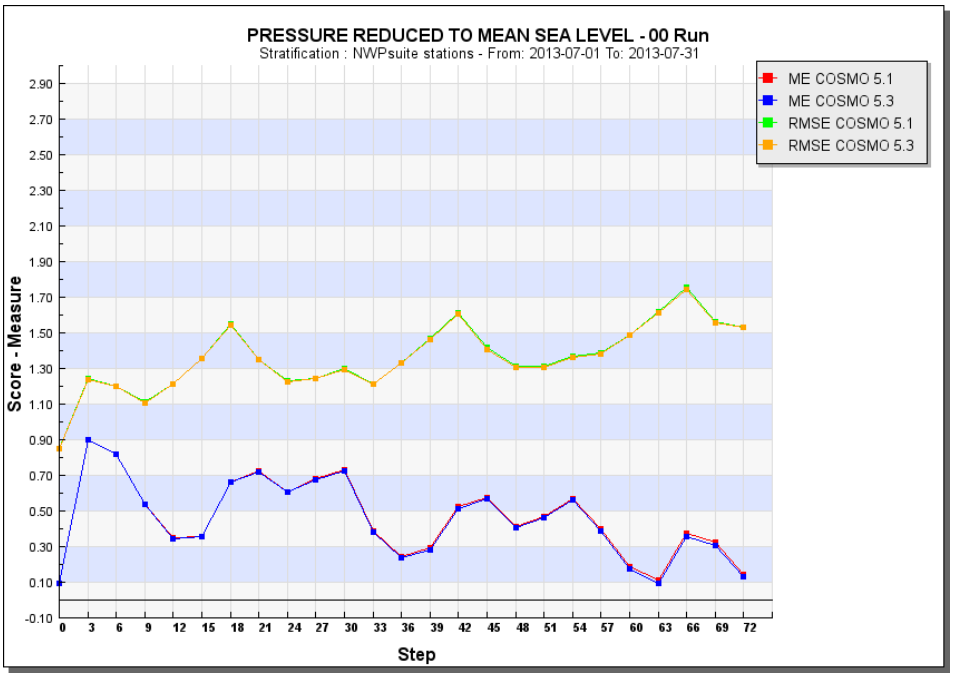

(a) (b)

Fig. 6 6h precipitation > 20mm verification results (00UTC run) – July 2013, PD for:
(a) COSMO-5.01 (5.1) (b) COSMO-5.03 (5.3)




| MSLP | winter | | |
|---|---|---|---|
| ME | 5.3 | 5.1 | |
| 0 | -0.036 | -0.036 | 0.000 |
| 3 | 0.767 | 0.765 | -0.002 |
| 6 | 0.855 | 0.853 | -0.003 |
| 9 | 0.618 | 0.615 | -0.003 |
| 12 | 0.601 | 0.599 | -0.002 |
| 15 | 0.658 | 0.656 | -0.002 |
| 18 | 0.798 | 0.796 | -0.002 |
| 21 | 0.923 | 0.920 | -0.003 |
| 24 | 0.949 | 0.946 | -0.004 |
| 27 | 0.843 | 0.840 | -0.004 |
| 30 | 1.014 | 1.011 | -0.003 |
| 33 | 0.773 | 0.770 | -0.003 |
| 36 | 0.749 | 0.747 | -0.002 |
| 39 | 0.761 | 0.759 | -0.002 |
| 42 | 0.827 | 0.826 | -0.002 |
| 45 | 0.930 | 0.928 | -0.002 |
| 48 | 0.851 | 0.848 | -0.003 |
| 51 | 0.768 | 0.765 | -0.003 |
| 54 | 0.997 | 0.994 | -0.003 |
| 57 | 0.741 | 0.737 | -0.004 |
| 60 | 0.665 | 0.661 | -0.004 |
| 63 | 0.784 | 0.780 | -0.004 |
| 66 | 0.839 | 0.836 | -0.003 |
| 69 | 0.935 | 0.932 | -0.004 |
| 72 | 0.831 | 0.827 | -0.004 |
| RMSE | 5.3 | 5.1 | |
| 0 | 1.117 | 1.117 | 0.000 |
| 3 | 1.468 | 1.467 | -0.001 |
| 6 | 1.665 | 1.664 | -0.001 |
| 9 | 1.633 | 1.632 | -0.001 |
| 12 | 1.641 | 1.641 | -0.001 |
| 15 | 1.526 | 1.526 | -0.001 |
| 18 | 1.588 | 1.587 | -0.001 |
| 21 | 1.643 | 1.642 | -0.001 |
| 24 | 1.744 | 1.742 | -0.001 |
| 27 | 1.816 | 1.816 | -0.001 |
| 30 | 2.038 | 2.038 | -0.001 |
| 33 | 2.029 | 2.029 | 0.000 |
| 36 | 2.018 | 2.018 | -0.001 |
| 39 | 1.945 | 1.944 | -0.001 |
| 42 | 2.001 | 2.000 | -0.001 |
| 45 | 2.098 | 2.097 | -0.001 |
| 48 | 2.181 | 2.180 | 0.000 |
| 51 | 2.295 | 2.295 | 0.000 |
| 54 | 2.527 | 2.527 | -0.001 |
| 57 | 2.541 | 2.540 | 0.000 |
| 60 | 2.568 | 2.568 | 0.000 |
| 63 | 2.593 | 2.594 | 0.001 |
| 66 | 2.680 | 2.681 | 0.002 |
| 69 | 2.827 | 2.827 | 0.000 |
| 72 | 2.904 | 2.904 | 0.000 |

(a)




| MSLP | summer | | |
|---|---|---|---|
| ME | 5.3 | 5.1 | |
| 0 | 0.091 | 0.091 | 0.000 |
| 3 | 0.900 | 0.900 | 0.000 |
| 6 | 0.817 | 0.818 | 0.001 |
| 9 | 0.537 | 0.539 | 0.002 |
| 12 | 0.346 | 0.349 | 0.003 |
| 15 | 0.354 | 0.358 | 0.004 |
| 18 | 0.661 | 0.666 | 0.005 |
| 21 | 0.720 | 0.724 | 0.004 |
| 24 | 0.604 | 0.608 | 0.004 |
| 27 | 0.677 | 0.681 | 0.004 |
| 30 | 0.725 | 0.731 | 0.005 |
| 33 | 0.380 | 0.386 | 0.006 |
| 36 | 0.238 | 0.247 | 0.008 |
| 39 | 0.283 | 0.293 | 0.011 |
| 42 | 0.515 | 0.527 | 0.012 |
| 45 | 0.567 | 0.577 | 0.010 |
| 48 | 0.405 | 0.412 | 0.007 |
| 51 | 0.463 | 0.466 | 0.003 |
| 54 | 0.563 | 0.568 | 0.005 |
| 57 | 0.388 | 0.399 | 0.011 |
| 60 | 0.177 | 0.190 | 0.013 |
| 63 | 0.095 | 0.112 | 0.017 |
| 66 | 0.357 | 0.376 | 0.019 |
| 69 | 0.306 | 0.325 | 0.019 |
| 72 | 0.129 | 0.147 | 0.018 |
| RMSE | 5.3 | 5.1 | |
| 0 | 0.853 | 0.853 | 0.000 |
| 3 | 1.240 | 1.241 | 0.000 |
| 6 | 1.199 | 1.200 | 0.001 |
| 9 | 1.109 | 1.110 | 0.001 |
| 12 | 1.214 | 1.215 | 0.001 |
| 15 | 1.357 | 1.358 | 0.001 |
| 18 | 1.544 | 1.547 | 0.003 |
| 21 | 1.349 | 1.351 | 0.002 |
| 24 | 1.226 | 1.229 | 0.002 |
| 27 | 1.241 | 1.243 | 0.002 |
| 30 | 1.296 | 1.299 | 0.003 |
| 33 | 1.212 | 1.215 | 0.002 |
| 36 | 1.328 | 1.332 | 0.004 |
| 39 | 1.463 | 1.467 | 0.004 |
| 42 | 1.607 | 1.614 | 0.007 |
| 45 | 1.409 | 1.416 | 0.007 |
| 48 | 1.308 | 1.313 | 0.005 |
| 51 | 1.308 | 1.311 | 0.003 |
| 54 | 1.365 | 1.369 | 0.004 |
| 57 | 1.381 | 1.385 | 0.004 |
| 60 | 1.486 | 1.490 | 0.004 |
| 63 | 1.614 | 1.618 | 0.004 |
| 66 | 1.746 | 1.755 | 0.010 |
| 69 | 1.554 | 1.562 | 0.008 |
| 72 | 1.530 | 1.534 | 0.004 |

(b)

Fig. 7 Pressure reduced to mean sea level verification results (00UTC run) – COSMO5.01 (5.1) and COSMO-5.03 (5.3) ME and RMSE for: (a) January 2013 (b) July 2013. Numerical scores and differences on the right pane. Colors indicate: **red** - worsening, **green** - improvement, **yellow** - neutral.

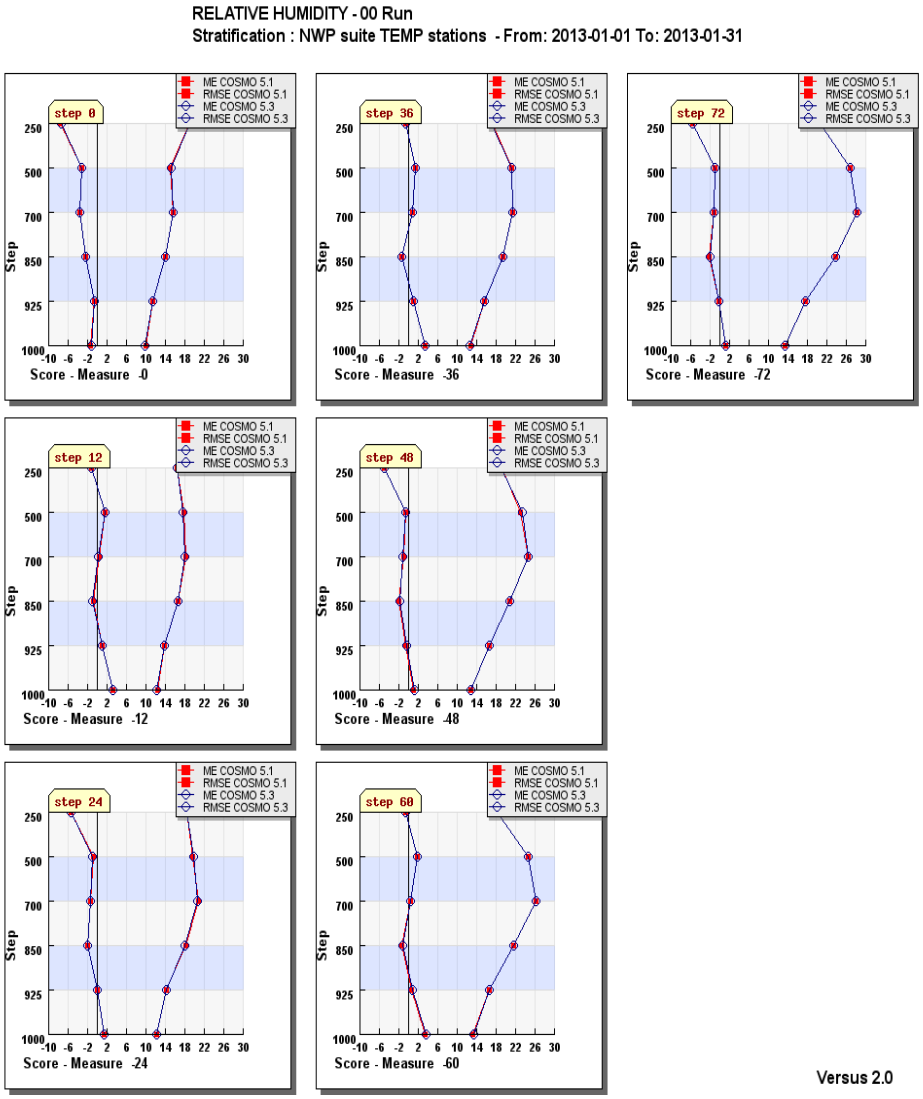



Fig. 8 Upper air verification for January 2013: Relative humidity COSMO-5.01 (5.10) - red / COSMO-5.03 (5.3) - blue

## List of publications/reports from the project with complete references

A. MONTANI, A. IRIZA, M. BOGDAN, A. CELOZZI, R. DUMITRACHE, F. GOFA - "Numerical Weather Prediction Meteorological Test Suite": COSMO 5.3 vs. 5.1, COSMO-Model Report, December 2015,
http://cosmo-model.org/content/model/documentation/NWPSuiteReports/cosmo v5.3-vs-5.1.pdf

A. MONTANI, A. IRIZA, R. DUMITRACHE, F. GOFA - "Numerical Weather Prediction Meteorological Test Suite": COSMO 5.1 vs. 5.0, COSMO-Model Report, June 2015,
http://cosmo-model.org/content/model/documentation/NWPSuiteReports/cosmo v5.1-vs-5.0.pdf

A. MONTANI, A. RASPANTI, F. GOFA, R. C. DUMITRACHE, A. IRIZA - Final Report of the COSMO Priority Task "Numerical Weather Prediction Meteorological Test Suite", Technical report, Consortium for Small-scale Modeling, 2014,
http://cosmo-model.org/content/consortium/generalMeetings/general2014/plenary/nwp_iriza.pdf
*Detailed guidelines for the proper use and execution of each NWP test using this platform were included in the report presented at the 16th COSMO General Meeting, 8-11 September 2014, Eretria, Greece.*

## Future plans

(Please let us know of any imminent plans regarding a continuation of this research activity, in particular if they are linked to another/new Special Project.)

- 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, both at 7km and 2.8km horizontal resolution.
- Performing model evaluation for the next versions of the model.
- Wider simulation area for the COSMO convection permitting horizontal resolution.
- Perform additional verification activities.
- A continuation project is currently registered at ECMWF as „Special Project spitrasp: COSMO NWP meteorological test suite”.
- At present, the tests for a new version (5.04a) of the COSMO model are undergoing. For this version, the model will be tested both at 7km and 2.8km horizontal resolution. For this purpose, tests for COSMO-5.03 (5.3) at 2.8km horizontal rezolution will also have to be performed, in order to provide an operational version against which the new model (5.04a) can be tested.