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	2014
Project Title:	Development of a perturbation strategy for convection-permitting ensemble forecasting over Italy
Computer Project Account:	SPITCONV
Principal Investigator(s):	Chiara Marsigli
Affiliation:	ARPA-SIMC, Bologna, Italy
Name of ECMWF scientist(s) collaborating to the project (if applicable)	…………………………………………………………… ……………………………………………………………
Start date of the project:	01.01.2013
Expected end date:	31.12.2015

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)	2,300,000	2,147,290	1,500,000	370,000
Data storage capacity	(Gbytes)	200	150	200	20

Summary of project objectives
(10 lines max)

The aim of this project is to study a good perturbation strategy for the development of convention-permitting ensemble forecasting based on COSMO over an Italian domain (COSMO-IT-EPS).

The main objectives are:

- test the performance of perturbed Initial Conditions (ICs) derived from the LETKF scheme for the km–scale of the COSMO Consortium (KENDA system), which is implemented on ECMWF computers for this purpose
- test new COSMO model perturbation approaches for the convection-permitting scale (SPPT, soil perturbations)
- test high-resolution Boundary Conditions (BCs) from ECMWF ENS experimental runs and further test the impact of an intermediate nesting with COSMO-LEPS

Summary of problems encountered (if any)
(20 lines max)

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

The general aim of this project is the development of COSMO-IT-EPS, addressing the issues of: IC perturbation, model perturbation, BC perturbation.

As a first step of this work, a "reference" ensemble, with no IC perturbations at the small scale was run, to allow the assessment of the impact of any further development. This was done (and it has been already reported) in the SP "Testing the impact of model perturbations applied to the COSMO model at a convection-permitting scale over Italy", where an ensemble with a basic set-up was run for the whole period SOP1.1. of the Hymex project (2 months).

During the reporting period (July 2013 – June 2014) several experiments aimed at the development of the COSMO-IT-EPS ensemble have been carried out thanks to the SPITCONV SBUs. In particular, the use of the LETKF (Localised Ensemble Transform Kalman Filter) scheme of COSMO (KENDA, Km-Scale Ensemble-Based Data Assimilation) as a mean to provide Initial Condition (IC) perturbations to the 2.8km ensemble under development has been investigated (section initial condition perturbations), as well as the impact of different Boundary Condition (BC) perturbations for the forecasting ensemble (section boundary condition perturbations).

Initial condition perturbation

During the first reporting period (Jan-Jun 2013) it was presented the implementation of the data-assimilation scheme based on KENDA on the ECMWF machines. The structure of the DA cycle is sketched in Figure 1.

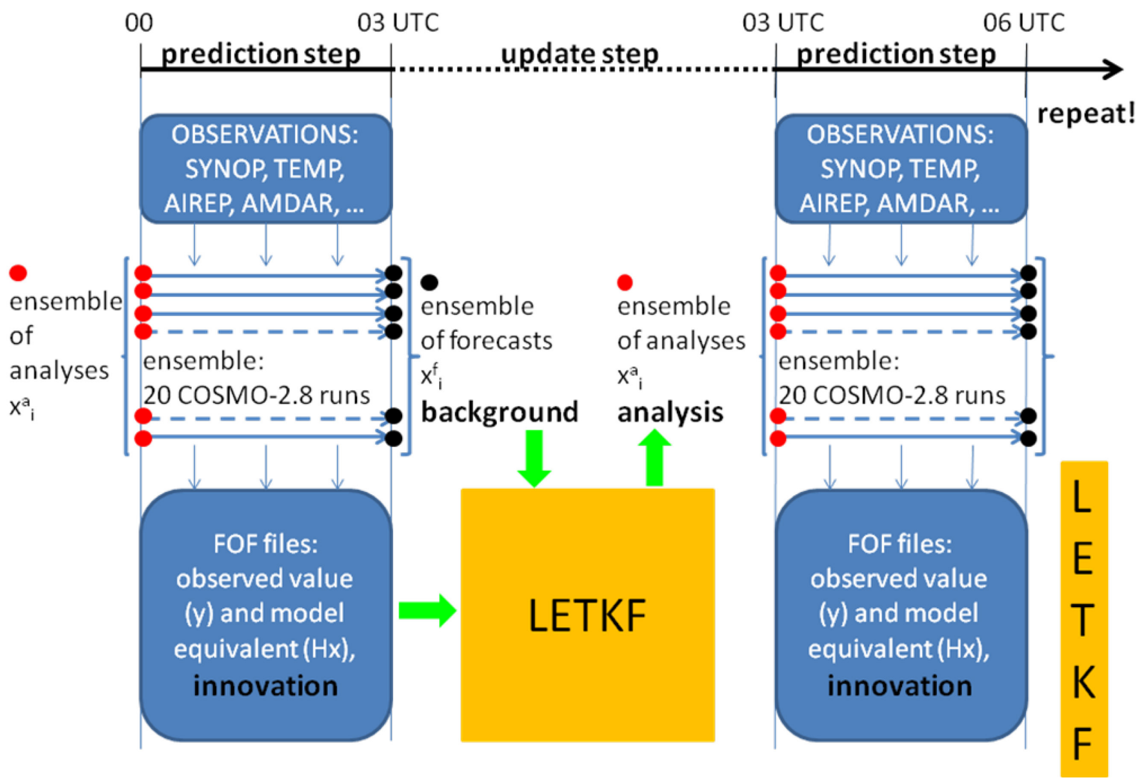

Figure 1. Set-up of the KENDA-based data assimilation cycle.

A prediction step, consisting of 20 runs of the COSMO model at 2.8km of horizontal resolution at +3h, is followed by an update step, when the KENDA code is run to produce the 20 perturbed analyses. This cycle is repeated as much times as needed to produce a good set of analyses to be used as ICs for the forecasting ensemble. In the experiments, usually 36h of DA are run to produce the analyses. In the operational set-up, the assimilation will be continuous.

An sms suite (iteps_2p8_suite, user mce) has been set-up, to perform the whole KENDA data assimilation chain on ECMWF machines (figure 2).

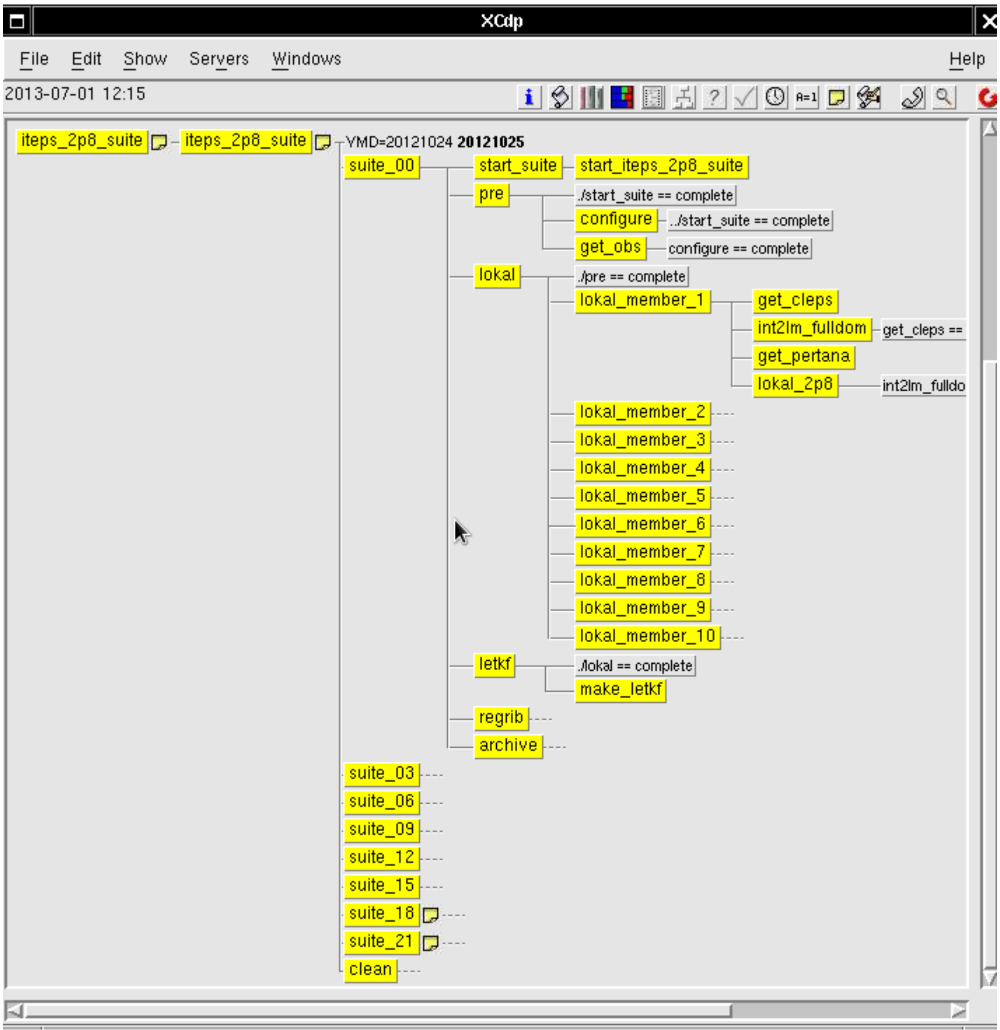

Figure 2. The sms suite for running the KENDA LETKF cycle (old suite with 10 members only).

BCs are provided either by COSMO-LEPS members or by ENS members, depending on the set-up, observations are retrieved from MARS archive and converted into NetCDF format with bufr2netcdf, IC for the first run (cold start) are provided either by COSMO-LEPS members or by ENS members, in accordance with the BCs.

In Table 1 are reported the dates of the cases considered up to now for the tests, together with few information on the case and on the duration of the DA cycle.

I.T. ensemble run	Area of the main meteorological phenomenon	KENDA cycle
2012092512	Alps	2012092400 2012092512
2012092612	Liguria Tuscany	2012092500 2012092612
2012101100	Corse Central Italy	2012101000 2012101100
2012101112	Corse Central Italy	2012101000 2012101112
2012102512	Northern Italy (IOP16)	2012102400 2012102512

Table 1. Dates of the cases considered for the KENDA testing.

For each case, the spectra of the KENDA-derived IC perturbations have been computed and compared with the spectra of the IC perturbations obtained by a downscaling of the global perturbed analyses (the analyses which constitute the ICs of ENS).

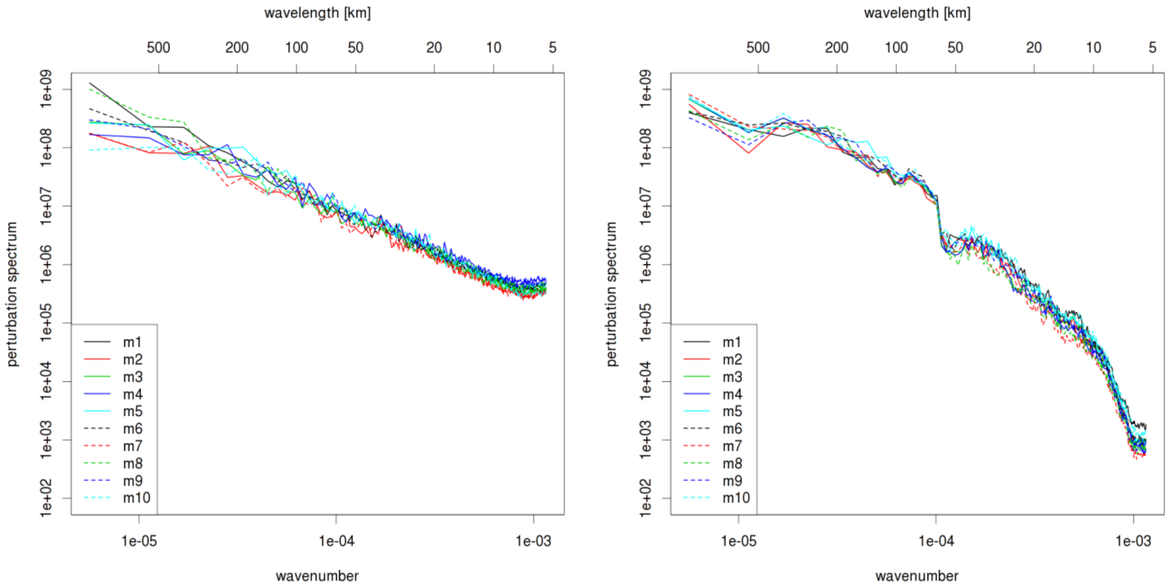

Figure 3. Spectra of the perturbations of ICs computed by KENDA assimilation cycle (left) and of the ICs of ENS interpolated on the COSMO grid at 2.8 km (right), for the Z500 fields.

The spectra of the perturbations are quite different starting from a scale of about 60 km and going towards smaller scales. It is evident that small scale perturbations are included in the KENDA-derived analyses, while ENS analyses are perturbed up to a scale of about 60 km, then a sharp decrease of the perturbation amplitude is visible. This provides an indication of the fact that the KENDA assimilation is able to introduce small scale perturbations in the COSMO analyses.

It has also been tested, on a couple of cases, the impact of provide ICs to the 2.8km COSMO ensemble from the KENDA analyses. For this test, the ensemble has been run with the same set-up adopted for the Hymex SOP1.1, only initial conditions are different.

In figure 4 the forecasts issued by 10 members of the reference ensemble (no ICs derived from KENDA, just downscaling) are shown, in terms of precipitation accumulated over 1 h for a forecast range 4-5 h. Observed precipitation over the same period is also shown as a separate plot.

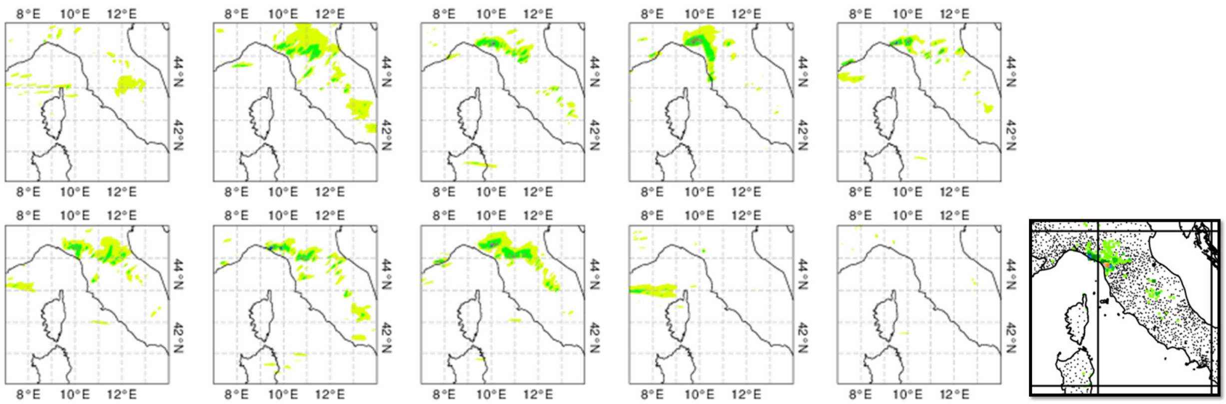

Figure 4. Forecast precipitation accumulated over 1 h for a forecast range 4-5 h for the 2.8 km ensemble starting at 12 UTC of the 20121011. Observed precipitation over the same period is also shown as a separate plot.

Precipitation is forecast by several ensemble members over the area hit by intense precipitation, even though the localization of the maximum is slightly misplaced.

In figure 5 it is shown the precipitation forecast issued by the same 2.8 km ensemble, but with ICs provided by KENDA analyses. In the top 10 panels, the analyses have been computed with a 36h KENDA cycle using only 10 members in the prediction step, while in the bottom 10 panels, the analyses have been computed with a 36h KENDA cycle using 20 members in the prediction step.

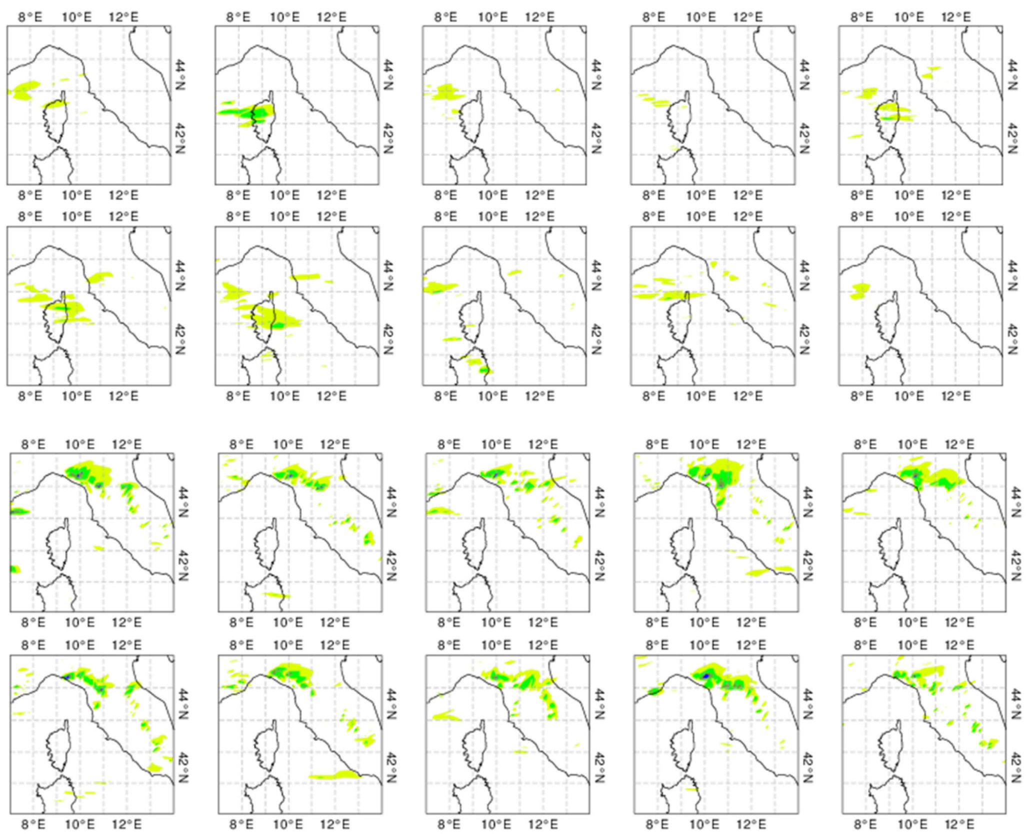

Figure 5. Same as in figure 4, but for an ensemble with ICs provided by KENDA analysis computed with 10 (top 10 plots) and 20 (bottom 10 plots) members in the data assimilation ensemble.

It is clear that 10 members in the data assimilation ensemble are not enough to allow a good assimilation of the observations, deteriorating the quality of the initial conditions. Instead, with 20 member the ensemble data assimilation is working properly, and the analyses are good enough to permit the initialization of the forecasting ensemble. The forecast in this case is at least as good as the one obtained with the downscaling, with some improvement in the fact that now all ensemble members correctly forecast intense precipitation of the area, with also a better localization of the maximum.

Then, it is presented a case in which the ensemble with ICs provided by KENDA is overperforming the ensemble with downscaled ICs.

In figure 6 it is shown the hourly precipitation forecast by the 10 members of the 2.8 km forecasting ensemble with downscaled ICs (left) and KENDA-derived ICs (right). Observed precipitation is also plotted as black dashed thick line. Precipitation is averaged over an area where intense precipitation was observed, including eastern Liguria and northern Tuscany.

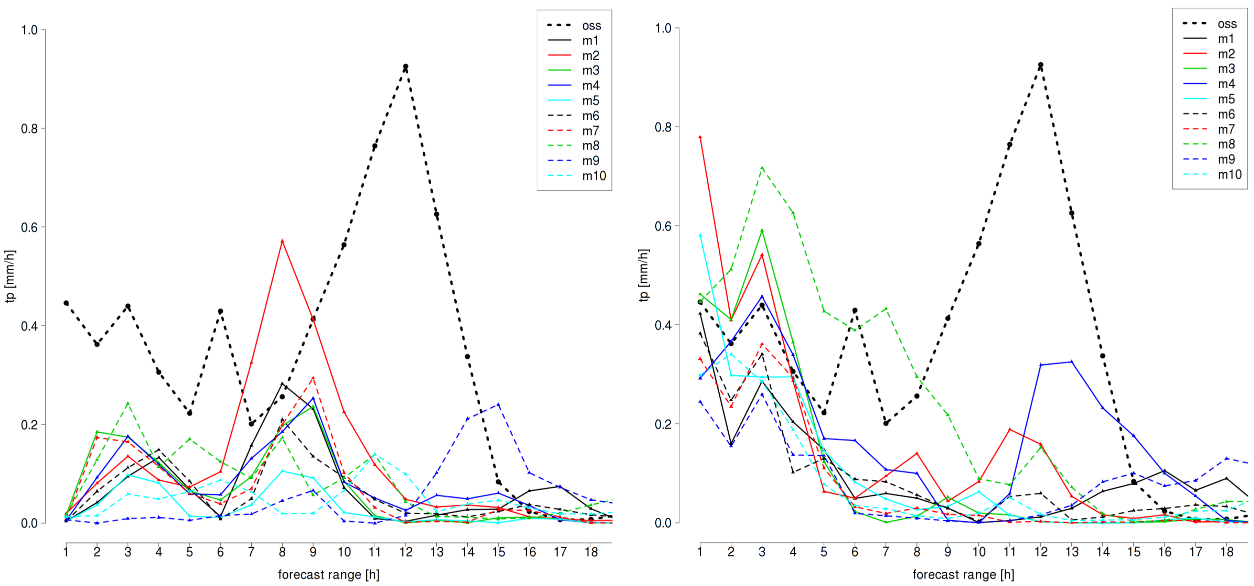

Figure 6. Same as in figure 4, but for an ensemble with ICs provided by KENDA analysis computed with 10 (top 10 plots) and 20 (bottom 10 plots) members in the data assimilation ensemble.

In this case the added value provided by the KENDA-based ICs is evident in the good performance of the ensemble in the first hours of forecast, when the downscaling ensemble was not able to forecast any intense rain over the area of interest.

These tests have been performed with a "standard" set-up of the KENDA algorithm, without any optimization. Instead, we want to study how to set-up the system in order to improve its performance over our area and for the purpose of providing ICs for the CP ensemble. In order to achieve this, an OSSE has been planned, aiming at studying the reactions of the KENDA assimilation to the values of the parameters regulating the assimilation (localization radius, inflation factor, number of members, duration of the cycles, …). This work is done in the framework of the

COSMO Priority Project KENDA. This experiment is now on-going and the runs of the system is this idealized framework are done thanks to the SPITCONV SBUs.

Boundary condition perturbation

in the framework of the C-SRNWP coordination and thanks to the help of ECMWF, it has been studied the effect of higher resolution BCs from ENS, for driving convection-permitting ensembles. ECMWF has performed the runs of 2 version of the EPS (T639, EPS-R and T1279, EPS-H) over the three periods in a configuration slightly different from the operational one, and the data were made available to the Countries.

At ARPA-SIMC, hourly BCs from this run have been used to provide perturbed BCs to 20 runs of the COSMO model at 2.8km horizontal resolution and 50 vertical level, over a domain including a large part of Italy and of the Tyrrhenian Sea and Switzerland.

Results have been presented during a meeting organized in the framework of the C-SRNWP Programme held at ECMWF in December 2013.

The set-up of the ensemble for this test is here summarised:

- IC and BCs from ECMWF ENS:
    - with operational resolution (T639): epsR experiment
    - with high resolution (T1279): epsH experiment
- hourly BCs
- no parameter perturbations
- no data assimilation
- 2.8 km, 50 levels
- 20+1 members
- initial time 12 UTC
- 36h forecast range

The Ranked Probability Score (RPS) and Skill Score (RPSS) for the precipitation forecasts accumulated over 12 hours are shown in Table 2, together with the percentage of outliers. The scores are computed against synop observations covering the whole domain.

+ 6-18h	RPS	RPSS	Outliers
epsR	0.157	0.449	14%
epsH	0.153	0.463	13%

+ 18-30h	RPS	RPSS	Outliers
epsR	0.183	0.339	10%
epsH	0.181	0.348	10%

Table 2. Dates of the cases considered for the KENDA testing.

RPS and RPSS are slightly in favour of the higher-resolution BCs configuration, for both forecast ranges. Considering the performance in dependence of the threshold, in terms of Brier Skill Score (figure 7), the result is confirmed, showing an higher score value for the ensemble nested on higher-resolution BCs for all the thresholds and both forecast ranges.

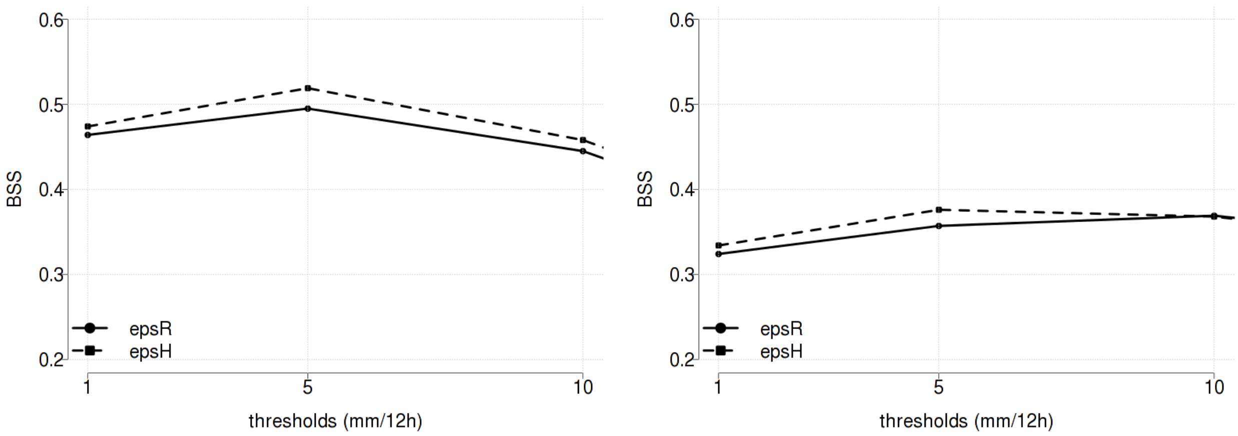

Figure 7. Brier Skill Score of the precipitation forecasts issued by the 2.8 km COSMO ensembles nested on epsR (solid) and epsH (dashed) as a function of the precipitation threshold.

The spread-skill relation of the ensemble has also been considered. In figure 8 the values relative to 2m temperature forecasts are plotted, as a function of the forecast range.

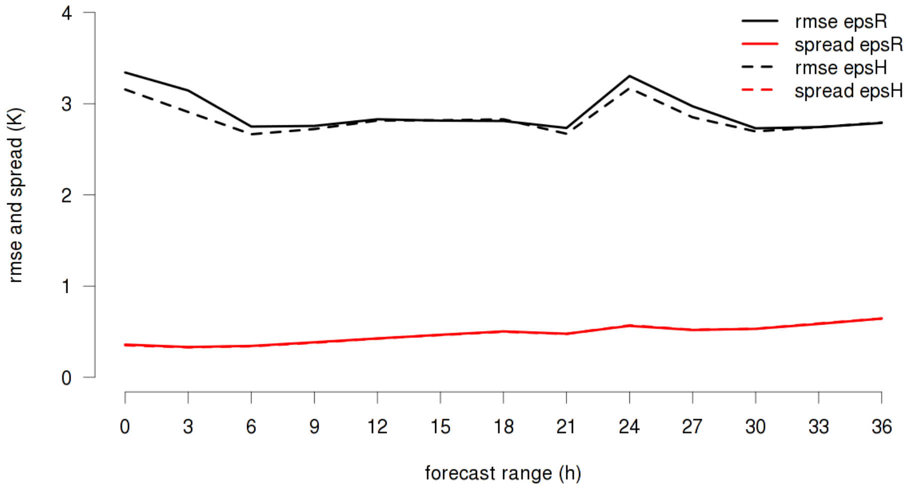

Figure 8. Root-mean-square error of the ensemble mean (black limes) and root-mean-square spread (red lines) for the 2.8 km COSMO ensembles nested on epsR (solid) and epsH (dashed) as a function of the forecast range.

The two spread lines are laying on each other, while the ensemble nested on epsH shows slightly smaller rms error.

Another experiment related to BCs for the convection-permitting ensemble has been carried out, aiming at defining the best set-up to be used in the ensemble configuration.

During the Hymex SOP 1.1 the COSMO-H2-EPS ensemble implemented for that project received ICs and BCs from 10 members of COSMO-LEPS, which, in turn, receive ICs and BCs from ENS members. It was decided to study the impact of driving the CP ensemble, instead, with the corresponding ENS members directly, skipping the intermediate step with the COSMO model at 7 km of horizontal resolution. Thanks to the SPITCONV resources (an increment of which has been required last year for this purpose), it has been possible to compare the two set-ups over a period of about 1 month, precisely 21 starting dates, with 36h forecast range ensemble runs.

The results of this comparison are presented in a paper recently published (Marsigli et al, 2014). Therefore, only a short summary is here provided.

Results show that the variability introduced in the geopotential field by the direct nesting is usually contained within the uncertainty described by the standard ensemble, and differences between pairs of members following different nesting approaches are generally smaller than the ensemble error, computed with respect to analysis. The relation between spread and error is even improved by the direct nesting approach (figure 9).

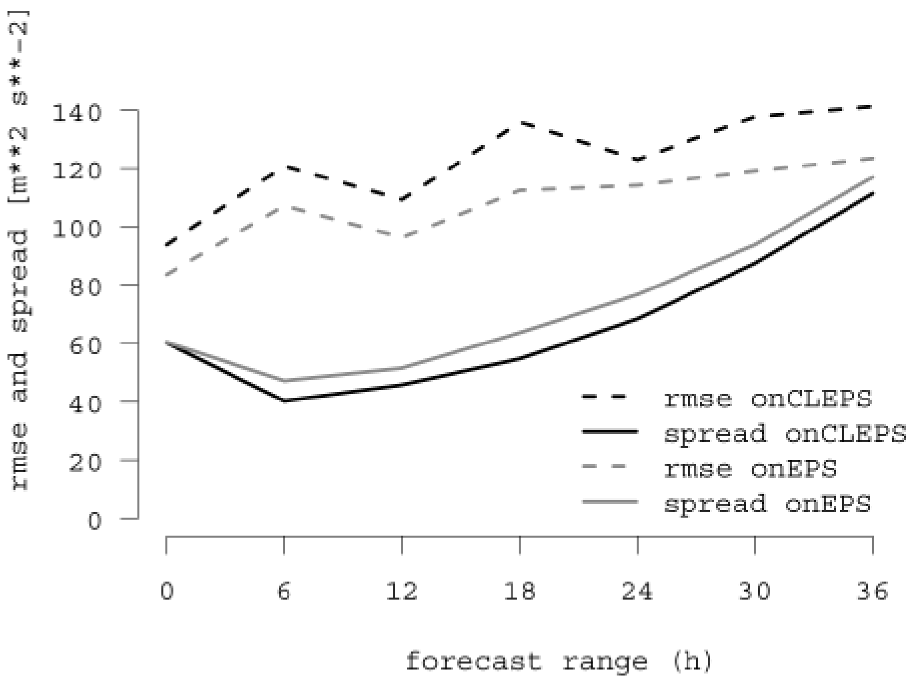

Figure 9. Root-mean-square error of the ensemble mean (dashed lines) and root-mean-square spread (solid lines) for the 2.8 km COSMO ensembles nested on COSMO-LEPS (black) and ENS (gray) as a function of the forecast range.

The skill of the CP ensemble precipitation forecasts, evaluated by means of an objective verification, is comparable. Therefore, the overall quality of the 2.8 km ensemble for the specific application is not deteriorated by the provision of lower resolution lateral boundary conditions directly from the global ensemble.

List of publications/reports from the project with complete references

Marsigli C., Montani A., Paccagnella T., 2014: Provision of boundary conditions for a convection-permitting ensemble: comparison of two different approaches. *Nonlin. Processes Geophys.*, **21**, 393–403.

Summary of plans for the continuation of the project

(10 lines max)

The idealized experiment (OSSE) is being carried out and will be the main task of this year with regards to KENDA, to study how the scheme react to changes in the assimilation procedure and algorithm, both in the update step (localization, inflation, observation used, quality control, observational error) and in the prediction step (number of members, duration of the cycles, boundary conditions). On top of this, other tests will be performed by initializing the convection-permitting forecast ensemble with the KENDA analyses, and results will be compared with the downscaling ensemble.

Beside this study, the impact of model perturbations will be addressed, mainly by testing the "stochastic tendencies" (SPPT) algorithm which has been implemented in the COSMO model. This task has been delayed up to now since it was decided to invest in the KENDA experimentation, first, also waiting for an official version of the COSMO code with the SPPT option.

Finally, the BCs issues will be further address, depending on the availability of other experimental dataset of ENS from ECMWF.

This extensive testing is aimed at defining an "optimised" set-up of KENDA for providing perturbed IC to the 2.8km ensemble, a strategy for model perturbations for COSMO at 2.8km, a choice for the BC perturbation, which could be combined in a test on a "complete" ensemble set-up.

Therefore, in 2015 an ensemble suite at 2.8 km adopting the defined perturbation strategy will be run and evaluated over a long enough period, in order to permit a robust statistical evaluation of its performances.