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:	Investigation of case studies during Sochi Olympic Games using COSMO-based ensemble prediction systems.
Computer Project Account:	SPCOLEPS
Principal Investigator(s):	Montani Andrea(Italy)
Affiliation:	ARPA-SIMC
Name of ECMWF scientist(s) collaborating to the project (if applicable)	……………………………………………………….…… ……………………………………………………….……
Start date of the project:	2015
Expected end date:	2017

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)			1.000.0000	140.000
Data storage capacity	(Gbytes)			50	2

Summary of project objectives
(10 lines max)

It is planned to investigate the performance of the COSMO-based ensemble systems implemented during Sochi Olympic games (namely COSMO-S14-EPS and COSMO-RU2-EPS) for a number of cases occurred during February and March 2014, varying the configurations of the ensembles. For cases of poor performance of COSMO-S14-EPS, the ensemble size will be increased to 20 and to 51 and the probabilistic predictability of the system will be assessed. It will be also investigated the added value of the higher resolution of COSMO-RU2-EPS with respect to COSMO-S14-EPS. We will also run COSMO-S14-EPS for the missing dates during the Olympic season (from 15 January 2014 to 15 March 2014).

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)

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 billing units of the project were used to close the "holes" in the archive of COSMO-S14-EPS so as to perform a multi-model exercise considering the performance of different participating systems. This is described in the attached report (SCI-REPORT_spcoleps_2015.pdf).

List of publications/reports from the project with complete references

……

Summary of plans for the continuation of the project

(10 lines max)

It is planned to test the sensitivity of the forecast skill to the ensemble size of COSMO-S14-EPS and possibly of COSMO-RU2-EPS.

Performance of multi-model ensemble systems during Sochi-2014 Olympics: preliminary studies at ARPA-SIMC

Andrea Montani[1], D. Alferov[2], E. Astakhova[2], C. Marsigli[1], T. Paccagnella[1]

[1] *ARPA–SIMC, Bologna, Italy*

[2] *Roshydromet, Moscow, Russia*

1 Introduction

The past Winter Olympics and Paralympic Games took place in Sochi, Russia, from 7 to 23 February 2014 and from 7 to 16 March 2014. In the framework of these events, the **FROST-2014** (**F**orecast and **R**esearch in the **O**lympic **S**ochi **T**estbed; more info under `http://frost2014.meteoinfo.ru/`) was set up under the auspices of WMO (Kiktev, 2011). Several NWP systems were deployed in the Olympic area so as to advance the understanding of short-range prediction processes over complex terrain. As for limited-area ensemble prediction, six different systems were implemented and ran in real time during the Olympic season. All systems delivered both fields and plots which were used by the Sochi forecasters via the FROST-2014 Web-site (`http://frost2014.meteoinfo.ru/forecast/goomap`) for the prediction of high-impact weather over the Olympic mountainous areas up to day 3. While four systems were based on parameterised convection, two of them (namely, COSMO-RU2-EPS and HARMON-EPS) suuported an explicit representation of convection processes. Table 1 reports the main features of the particupating systems, while more details can be found under `http://frost2014.meteoinfo.ru/`.

Among the several post-Olympic activities, an intercomparison of the performance of some participating ensembles was initiated. As a pilot study, COSMO-S14-EPS (Montani et al., 2011 and 2013) and ALADIN-LAEF were selected and intercompared in terms of probabilistic prediction of precipitation. The attention was initially focussed on these two sysyems, because both of them were implemented and ran at ECMWF, the access to the fields to be compared being simpler and faster.

Table 1: Main characteristics of limited-area ensemble prediction systems.

System name	ensemble size	resolution (km)	forecast length (h)	boundary conditions	runs (UTC)
COSMO-S14-EPS	10	7	72	ECMWF ENS	00,12
ALADIN-LAEF	17	11	72	ECMWF ENS	00,12
GLAMEPS	54	11	54	ECMWF ENS	06,18
NMMB	7	7	72	GEFS	00,12
COSMO-RU2-EPS	10	2.8	48	COSMO-S14-EPS	00,12
HARMON-EPS	14	2.5	36	ECMWF ENS	06,18
combined	**27**	**7**	**72**		**12**

2 Main features of the intercomparison

In this section, the preliminary results of the intercomparison exercise are presented. The skills of the different ensemble systems were assessed over the period 15 January – 15 March

2014 (the official verification period of 2014 Olympics). In addition to COSMO-S14-EPS and ALADIN-LAEF, we also verified the performance of the multi-model ensemble (referred to as **"combined"** in Table 1) obtained by pooling together the members of COSMO-S14-EPS and ALADIN-LAEF. This enabled the generation of a 27–member ensemble, whose elements were reinterpolated onto a common 7-km grid. For these three systems, we considered the probabilistic prediction of 12-hour precipitation exceeding a number of thresholds for several forecast ranges. For reason of brevity, only the results relative to the 12UTC runs are reported. As for observations, it was decided to use the data obtained from the SYNOP reports available on the Global Telecommunication System (GTS). This enabled the possibility to assess the performance of the systems over a relatively dense observation dataset (73 stations), since the verification domain was restriced to an area centred over the Olympic venue (40-50N, 35-45E), shown in Fig. 1. As for the comparison of model forecasts against observations, we selected the grid-point closest to the observation.

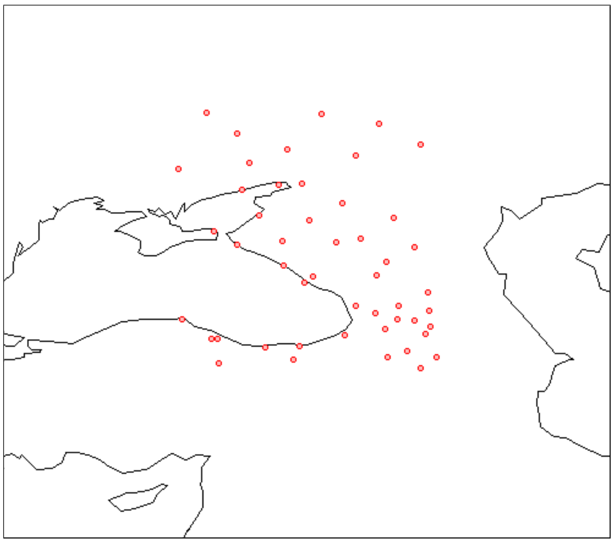

Figure 1: Location of the stations used for the intercomparison exercise between COSMO-S14-EPS, ALADIN-LAEF and "combined".

The skill of the systems was studied for 6 different thresholds: 1, 5, 10, 15, 25 and 50 mm/12h. The following probabilistic scores were computed: the Brier Skill Score (BSS), the Ranked Probability Skill Score (RPSS), the Relative Operating Characteristic Curve (ROC) area, the Rank Hystograms (RK) and the Percentage of Outliers (OUTL). For a description of these scores, the reader is referred to Wilks (1995). The main features of the intercomparison exercise are also summarised in Table 2.

3 Results

The performances of COSMO-S14-EPS, ALADIN-LAEF and of the "combined" ensemble are presented in Fig. 2, where we evaluate their ability to predict two different weather events: 12-hour precipitation exceeding 5 mm (left panel) and 10 mm (right panel). The values of the ROC area are plotted against the forecast range for each event.

It can be noticed that the ROC area values are above 0.8 for all systems and for both thresholds, indicating that both COSMO-S14-EPS and ALADIN-LAEF manage to discriminate these events. The skill of ALADIN-LAEF (blue lines) is quite constant with the forecast ranges, while the performance of COSMO-S14-EPS (red lines) varies. The latter ensemble

Table 2: Main features of the intercomparison configuration.

variable:	12-hour precipitation (18-06, 06-18 UTC)
starting time:	12 UTC;
period:	from 15 January to 15 March 2014;
region:	40-50N, 35–45 E;
method:	nearest grid–point;
observations:	SYNOP reports (about 73 stations/day);
fcst ranges:	6-18h, 18-30h, 30-42h, 42-54h, 54-66h;
thresholds:	1, 5, 10, 15, 25, 50 mm/12h;
scores:	ROC area, BSS, RPSS, OUTL;
systems:	ALADIN-LAEF, COSMO-S14-EPS, **combined**

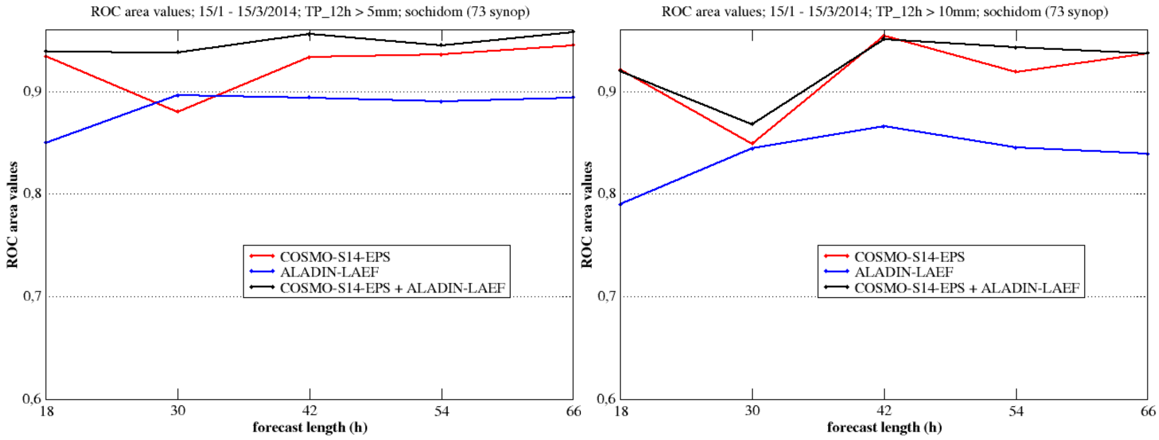

Figure 2: ROC area values as a function of the forecast range for two different weather events: 12-hour precipitation exceeding 5 mm (left) and 10 mm (right). The scores are calculated over the period January-March 2014. Red lines refer to COSMO-S14-EPS, blue lines to ALADIN-LAEF, black lines to the "combined" ensemble.

seems to be sligtly superior for most forecast ranges. The combined ensemble (black lines) provides the best scores especially for the lower threshold (left panel of Fig. 2), suggesting the added value of a multi-model approach with respect to the single-model one. The higher skill of the combined ensemble can also be noticed if either lower or higher thresholds are considered (not shown).

The above results are confirmed and even strengthened if the performance of the systems is analysed in terms of "integrated" scores, that is not depending on a particular threshold value. Fig. 3 shows the RPSS and the percentage of outliers (left and right panels, respectively) for the systems under investigation. As for the RPSS, the performance of the two single-model ensembles is very similar. It can be noticed that COSMO-S14-EPS and ALADIN-LAEF exhibit opposite cycles of the score: the former (latter) system has better skill in predicting precipitation which occurs during night-time (day-time). The combined ensemble (black line) takes the best of both components and provides higher scores with a reduced daily cycle. As for the outliers (right panel in Fig. 3, the added value of the combined ensemble is extremely clear for all forecast ranges, with a 50% reduction of the number of times the analysis is out of the forecast interval spanned by the ensemble members.

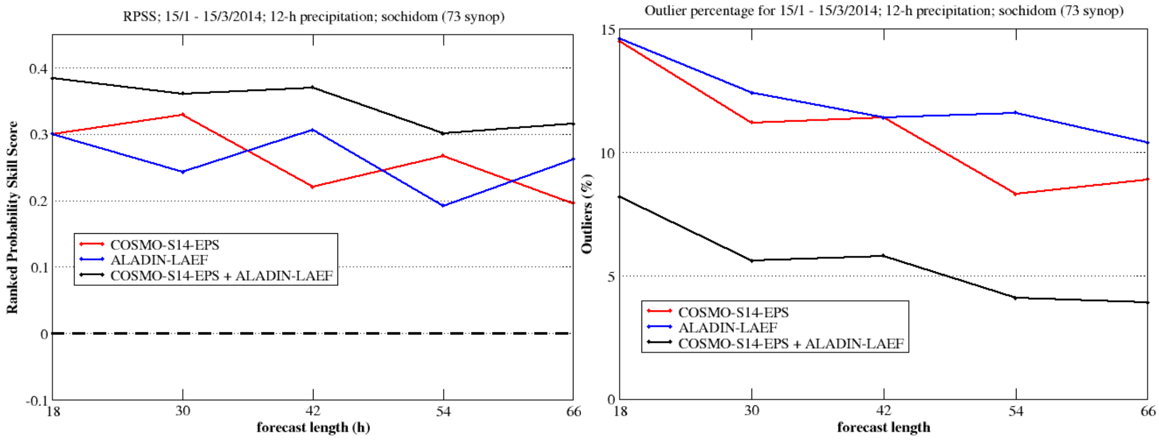

Figure 3: Ranked Probability Skill Score (left panel) and percentage of Ouliers (right panel) as a function of the forecast range. The scores are calculated over the period January-March 2014. Red lines refer to COSMO-S14-EPS, blue lines to ALADIN-LAEF, black lines to the "combined" ensemble.

4 Summary and Outlook

Although the results presented in the previous section are based on the combination of only two systems, there is the potential for a remarkable gain in predictability by blending different ensemble systems.

As for the future, it is planned to consolidate the verification results, by considering the performance of all participating system for more variables (in particular, 2-metre temperature) and by using the extra-observations (non GTS) which were taken during the Olympics.

References

Kiktev D., 2011. Forecast and Research: the Olympic Sochi Testbed (FROST-2014). Concept paper. Available at http://frost2014.meteoinfo.ru

Montani A., Cesari D., Marsigli C., Paccagnella T., 2011. Seven years of activity in the field of mesoscale ensemble forecasting by the COSMO-LEPS system: main achievements and open challenges. *Tellus*, **63A**, 605-624. DOI: 10.1111/j.1600-0870.2010.00499.x

Montani A., Marsigli C., Paccagnella T., 2013. Development of a COSMO-based limited-area ensemble system for the 2014 Winter Olympic Games. *Cosmo Newsletter No. 13*, 93-99. Available at http://www.cosmo-model.org

Wilks, D.S., 1995. Statistical Methods in the Atmospheric Sciences. Academic Press, New York, 467.