Ensemble activities in Arpae

discussion on a series of experiments aimed at improving COSMO-2I-EPS performance

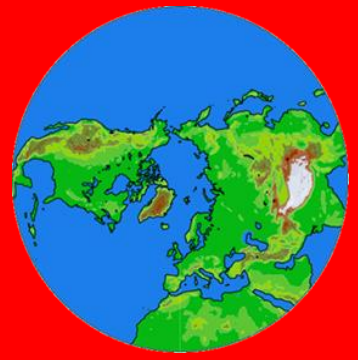
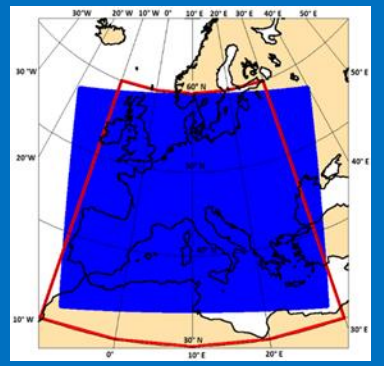
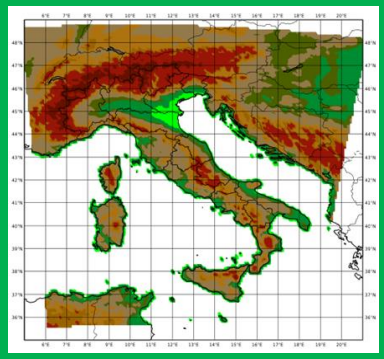
G. Pincini, I. Cerenzia, T. Gastaldo, D. Cesari, M.S. Tesini, E. Minguzzi

CONSORTIUM FOR SMALL SCALE MODELING
COSMO

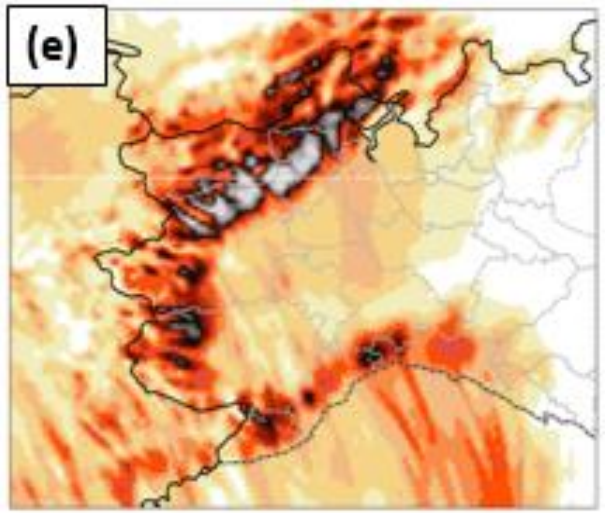
arpae
agenzia prevenzione ambiente energia emilia-romagna

ENSEMBLE PREDICTION SYSTEMS IN arpae

agenzia prevenzione ambiente energia emilia-romagna

MAIN TECHNICAL FEATURE \ ENSEMBLE SYSTEM	ECMWF ENS	COSMO-LEPS	COSMO-2I-EPS
Integration domain:			
Hor/Vert resolution	18km /91 lev	7km/40lev	2.2km/65lev
Initial Conditions	Hybrid-EnVar	ENS (+cluster analysis)	LETKF
Boundary Conditions	-	ENS (+cluster analysis)	AM ENS (nested on ENS)
Model Perturbation	Stochastic scheme	Parameter perturbation	-
Forecast range (hours)	240	132	51
Type of model	Hydrostatic model	Non-hydrostatic model	Non-hydrostatic model
Type of convection	Parameterized convection	Parameterized convection	Explicit convection
Ensemble size	51	20	20
Starting times (UTC)	00, 06, 12, 18	00, 12	21

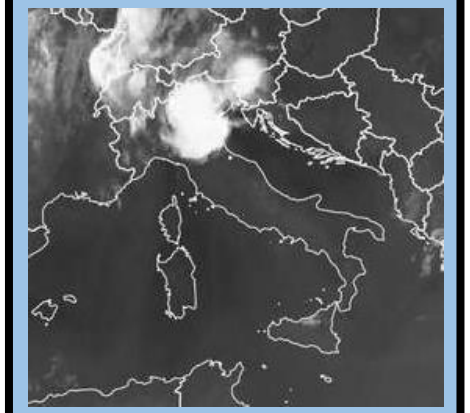
Ensemble activities



1994 Piedmont flood

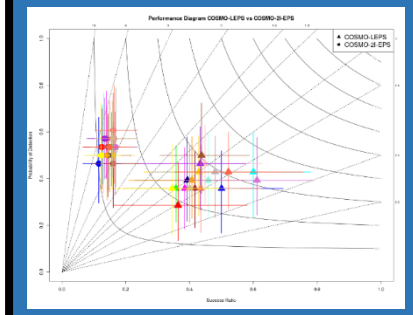
Friday 4 Set
12:00
PP-AWARE
I. Cerenzia

Flash flood in Palermo 15 July 2020

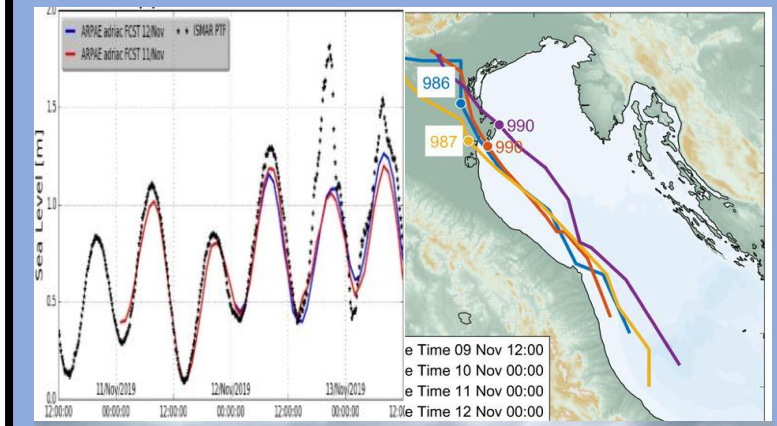


Severe thunderstoms summer 2020

SRNWP-EPS (in EUMETNET NWP Program)



Experiments on COSMO-2I-EPS



2019 flooding of Venice

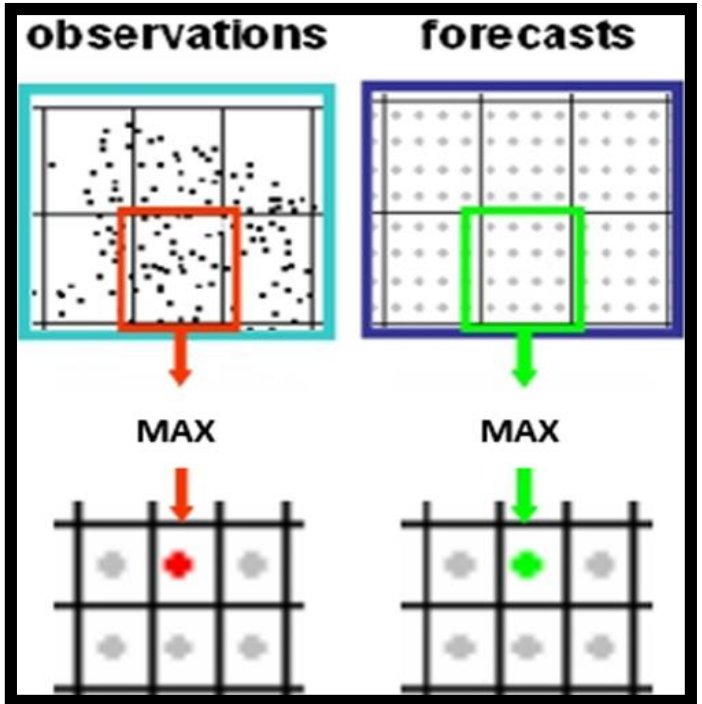
DESCRIPTION OF THE EXPERIMENTS

The intercomparison between the three ensemble systems is performed starting at **00 UTC** and with a forecast range of **48 hours** (post-processing frequency every 6 hours).

The systems are compared over the **Italian region**

Italian domain
Latitude: 35N – 48N
Longitude: 6E – 19E

DPCN Observational network
5524 stations



The domain is divided in squared area **(0.25° x 0.25°)**;
the precipitation values of all stations and all model grid points falling in the same box are aggregated and processed.
The **maximum** value for the precipitation field in each **box** has been performed.

Definitions of some probabilistic scores

❖ **Ranked Probability Scores (RPS)** is an extension of the RMSE to the probabilistic world and to the multi- category events. RPS ϵ (0,1); The lower the RPS, the better the ensemble system.

$$RPS = \frac{1}{J-1}\sum_{m=1}^{J}(\sum_{j=1}^{m} f_j - \sum_{j=1}^{m} o_j)^2$$

❖ **The percentage of outliers** of a probabilistic forecast system is defined as the probability of the observations lying outside the range spanned by the forecast values.

Here it is computed as the fraction of points of the domain where the observed value lies outside the minimum or maximum forecast value.

❖ **Brier Score (BS)** is the mean square error of the probability forecast. The BS averages the squared differences between pairs of forecast probabilities and the corresponding binary observations, representing the occurence of the event. BS ϵ (0,1); the perfect forecast has BS=0.

$$BS = \frac{1}{N}\sum_{i=1}^{N}(p_i - o_i)^2.$$

❖ **The Relative Operating Characteristic Curve Area (ROC Area)** is the area under the curve generated by plotting of the cumulative Hit Rate H against False Alarm Rate F.

$$H_k = \frac{a_k}{a_k + c_k} \qquad F_k = \frac{b_k}{b_k + d_k}$$

Following the contingency table

H_k is the proportion of events which were predicted by k ensemble members and actually happened

F_k is the proportion of events forecast by k members and did not occur

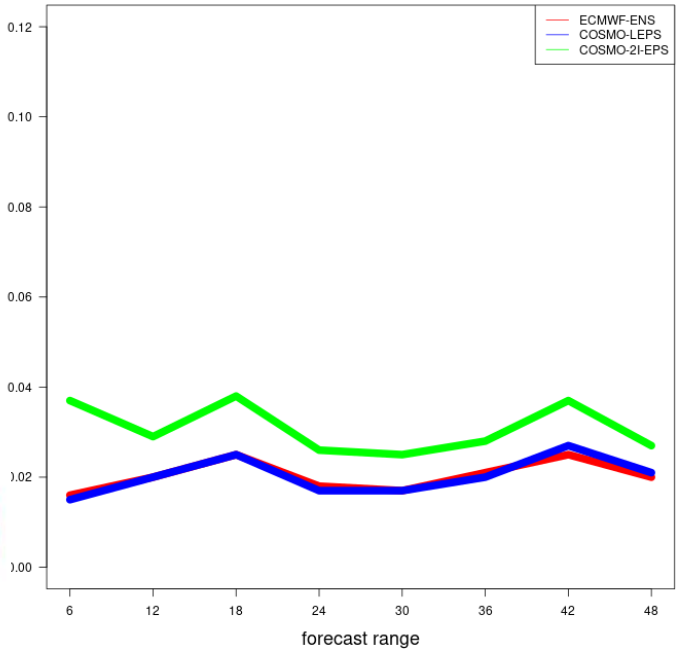
The maximum value is 1 and a value of 0.5 indicates a no-skill forecast system.

Contingency table	Observed YES	Observed NO
Forecast YES	a	b
Forecast NO	c	d

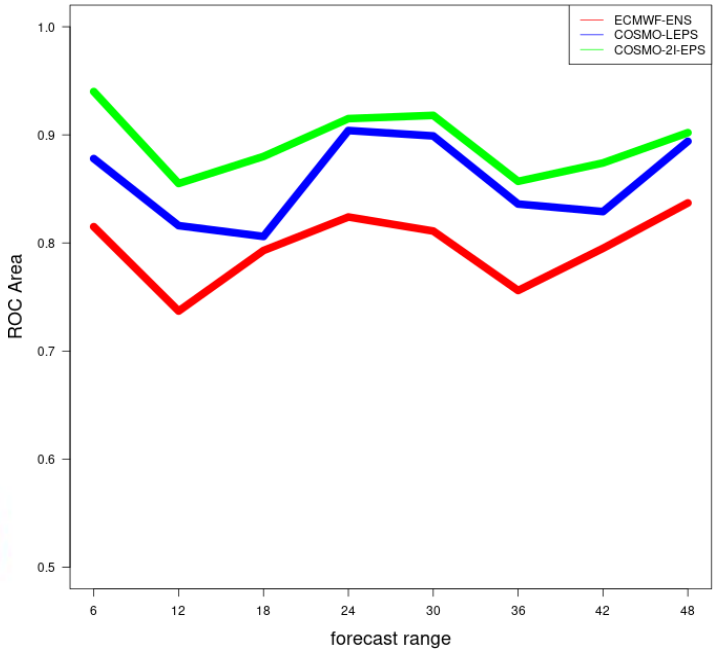
		OBS	
		YES	NO
FCST	YES	hits	false alarm
	NO	misses	correct rejection

Previous verification periods (January – April 2019)

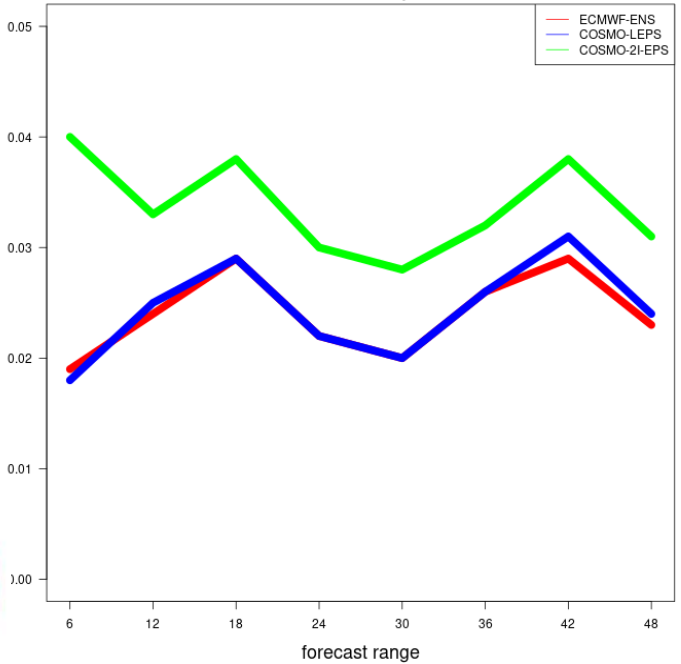
Brier Score (10mm)



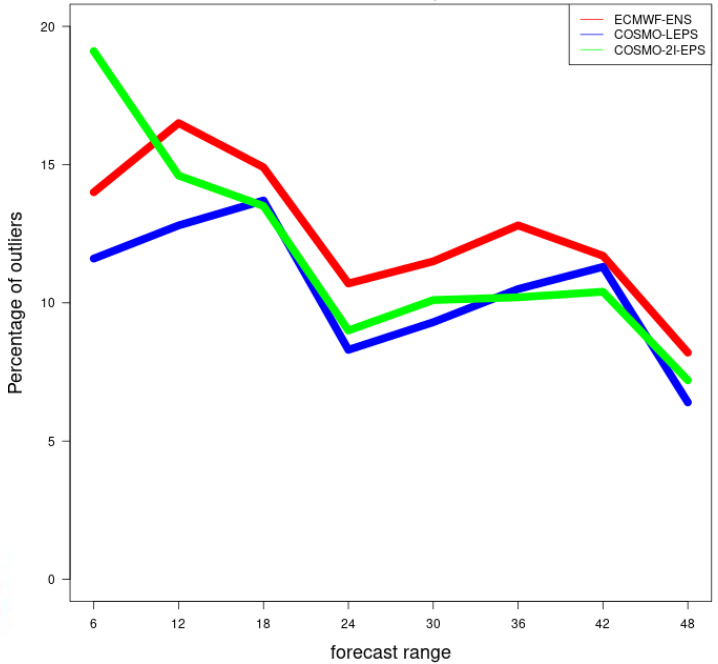
ROC Area (10mm)



Ranked Probability Score



Percentage of outliers



ECMWF ENS

COSMO-LEPS

COSMO-2I-EPS

COSMO-2I-EPS obtains positive results only for ROC Area and percentual of outliers, while the results are worse than COSMO-LEPS and ECMWF ENS for all other analyzed scores.

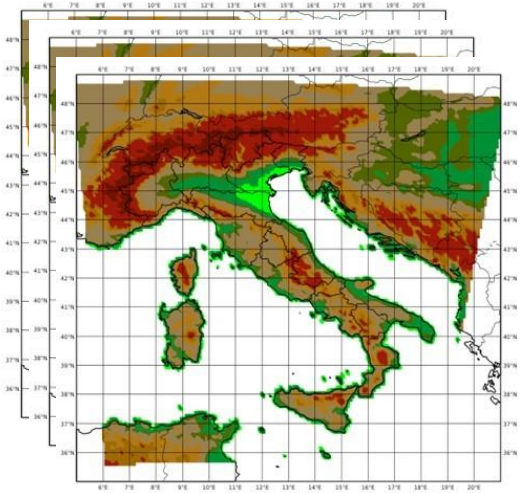
This led us to do other experiments on different periods of verification to try to understand the source of these negative performances.

ECMWF ENS
(18km)

COSMO-ME-EPS
(7km)

IC+BC

KENDA (2.2km)

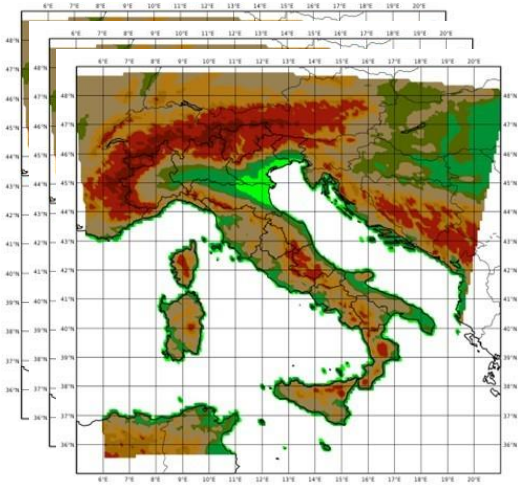


3h assimilation cycle

18 | 21 | 00 | 03 | 06

IC

BC
(from run initialized at 12UTC)



COSMO-2I-EPS
(2.2km)

- Run initialized at 21UTC once a day
- Lead time: +51h
- 20 ensemble members
- Initial perturbation: KENDA
- No model perturbations

POTENTIAL SOURCES OF ERRORS

EXPERIMENTS

A **Performance of COSMO-ME-EPS**

- **Run COSMO-2I-EPS using ENS in spite of COSMO-ME-EPS : moderate improvement**
- Compare performance of COSMO-ME-EPS with ENS, COSMO-LEPS and COSMO-2I-EPS: to be done

B **Performance of KENDA in producing the perturbed initial conditions**

- Remove KENDA (initialize from COSMO-ME-EPS directly): to be done
- Test KENDA performance (the deterministic forecast initialized by KENDA's deterministic analysis has the same or a better performance than the other models (ECMWF, COSMO-5M etc.): not easy

C **Unfair comparison between ensembles: ECMWF and COSMO-LEPS use IC and BC at 00 UTC while COSMO-2I-EPS employs IC (from KENDA, 2.2km) at 21UTC and BC (from COSMO-ME-EPS, 7km) at 12 UTC**

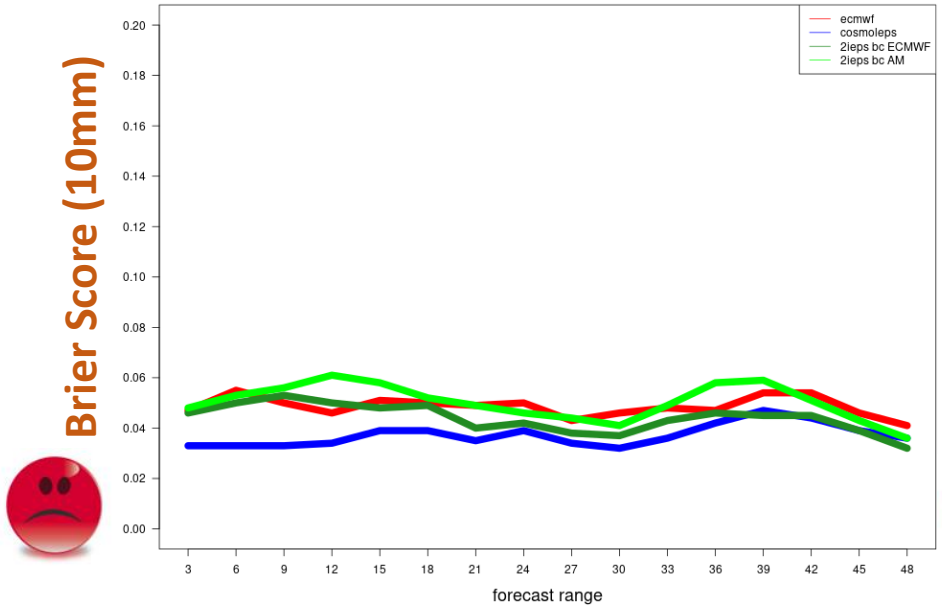
D **Incompatibility between IC (from KENDA at 21UTC) and BC (from COSMO-ME-EPS at 12UTC)**

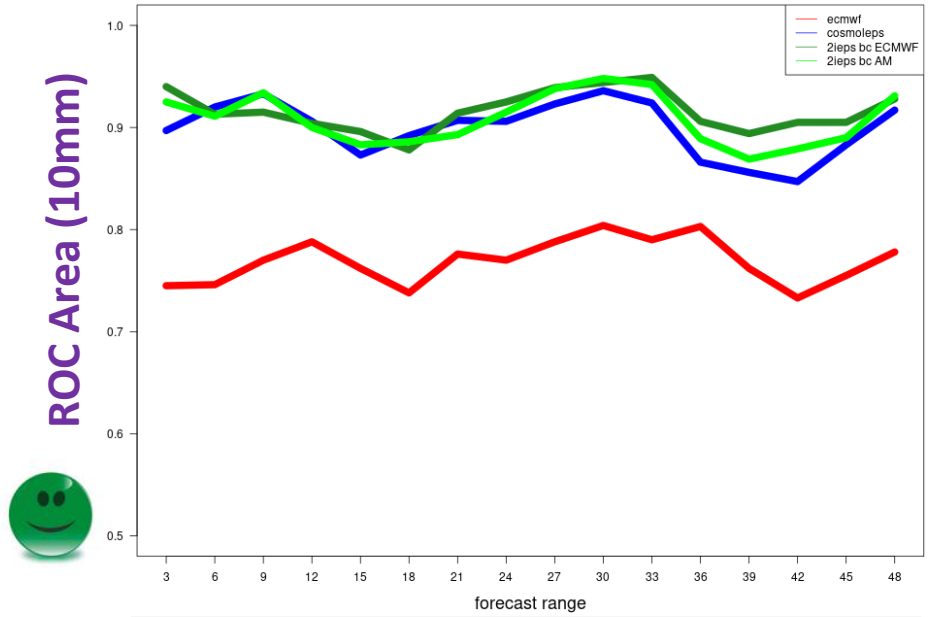
- **Initialize COSMO-2I-EPS at 00UTC (use IC from KENDA at 00UTC and BC from COSMO-ME-EPS at 00UTC): no variation or light degradation**

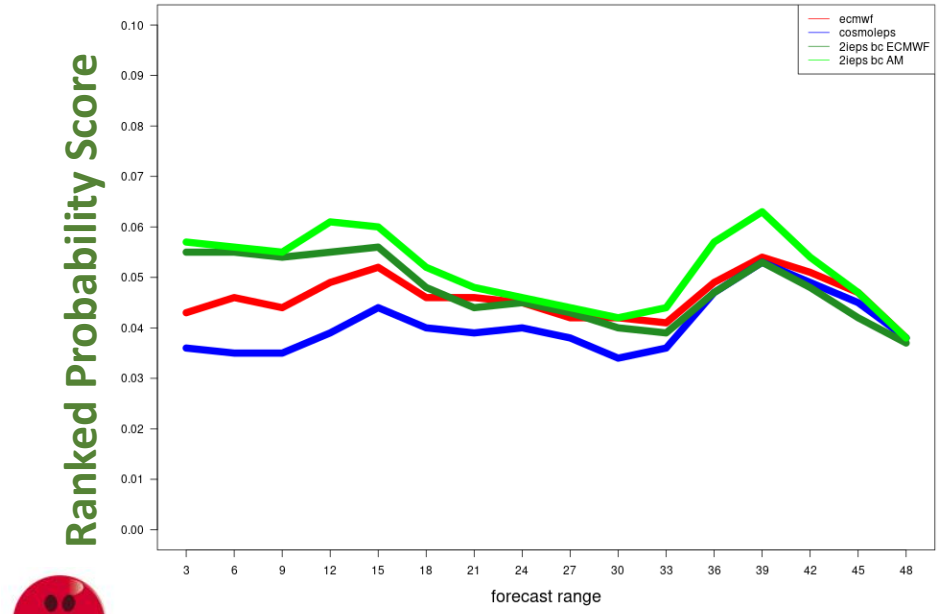
E **Lack of any model perturbation**

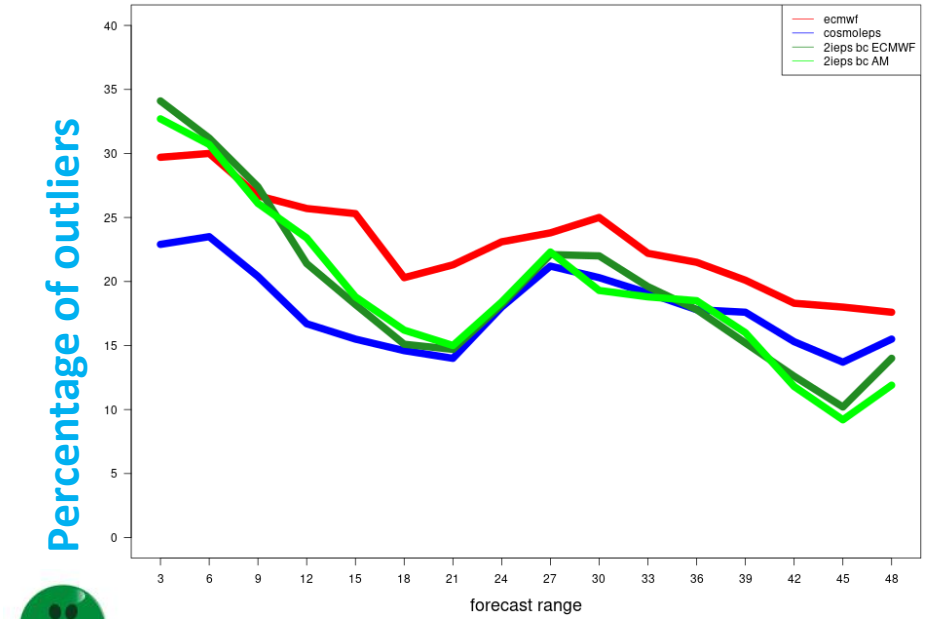
- Test SPPT: to be done

Run COSMO-2I-EPS using ENS in spite of COSMO-ME-EPS (from 22 Nov to 28 Nov 2019)









ECMWF ENS

COSMO-LEPS

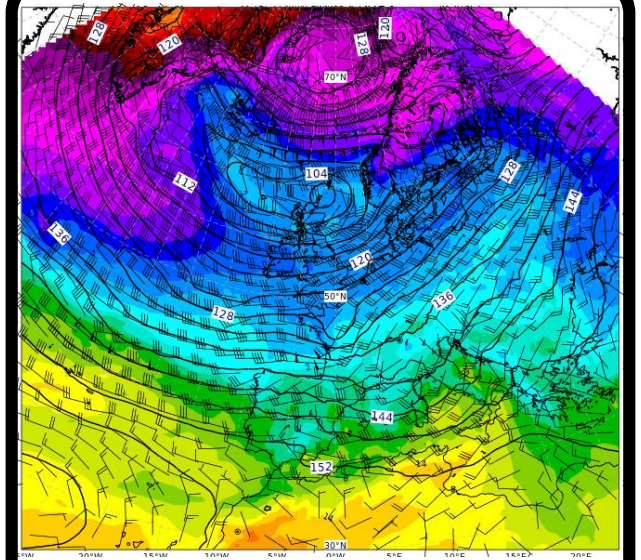
COSMO-2I-EPS ope

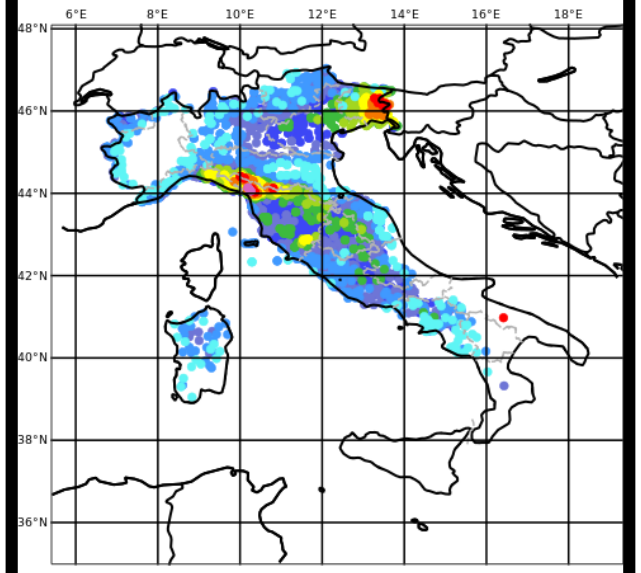
COSMO-2I-EPS bc ENS

Also in this experiment the good results of COSMO-2I-EPS are confirmed only for the percentage of outliers and for the ROC Area.

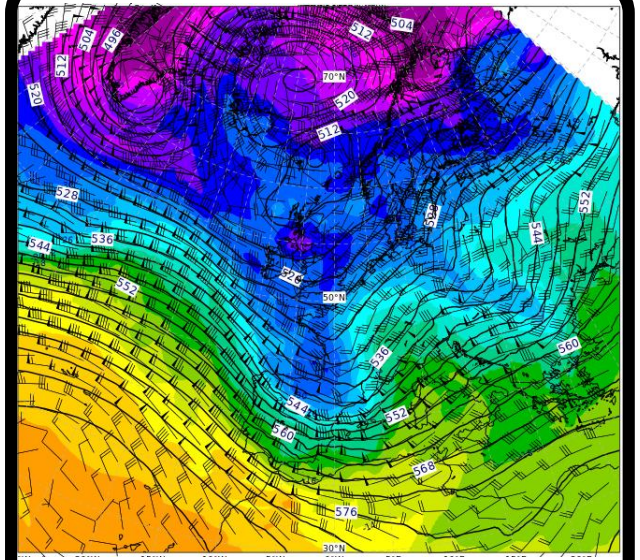
The use of ENS BCs leads to a slight improvement of most of the COSMO-2I-EPS scores (effect already known from the *Marsigli et al. (2014)*), even if often not enough to fill the existing gap with COSMO-LEPS.

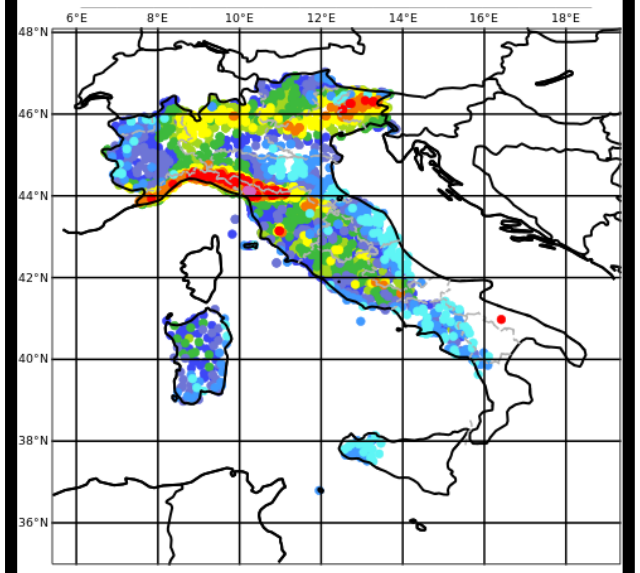
Initialize COSMO-2I-EPS at 00UTC (use IC from KENDA at 00UTC and BC from COSMO-ME-EPS at 00UTC)



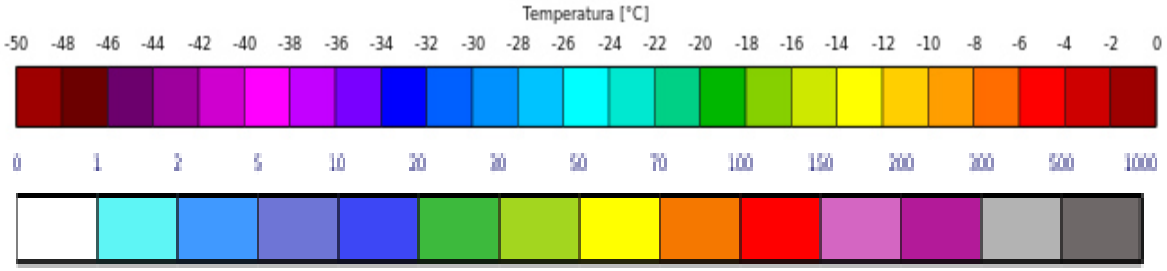


1 March

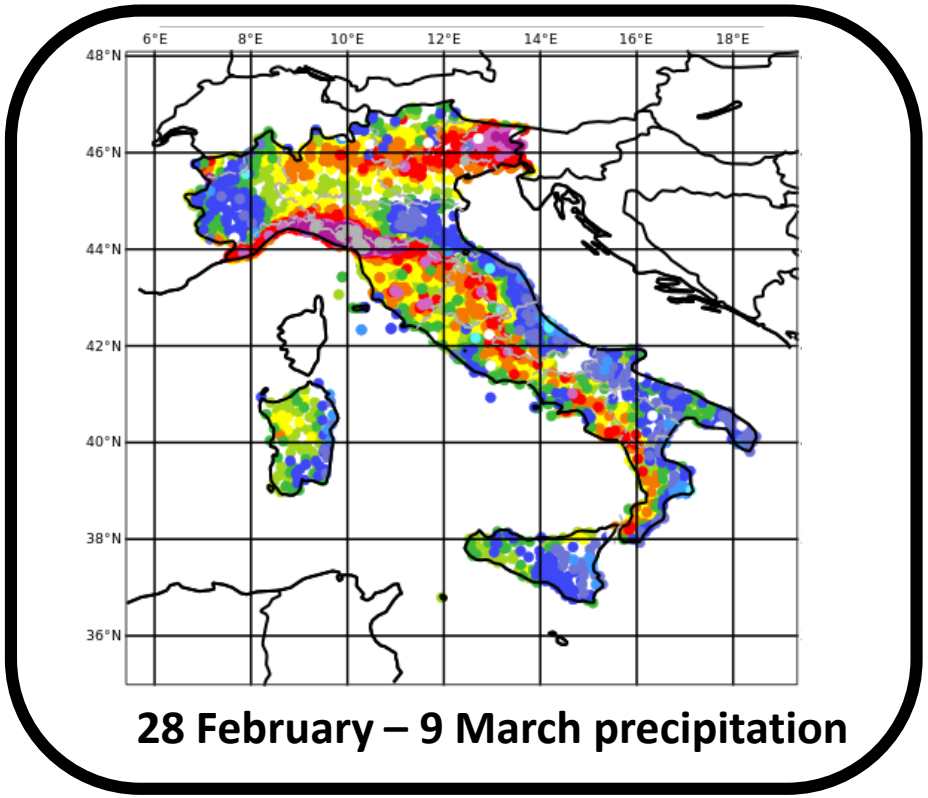




2 March

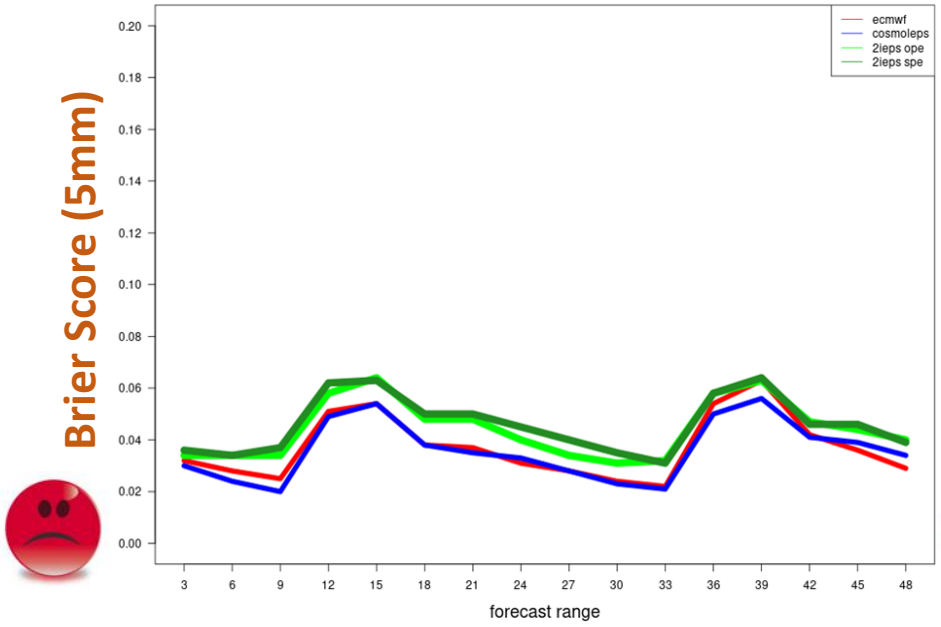


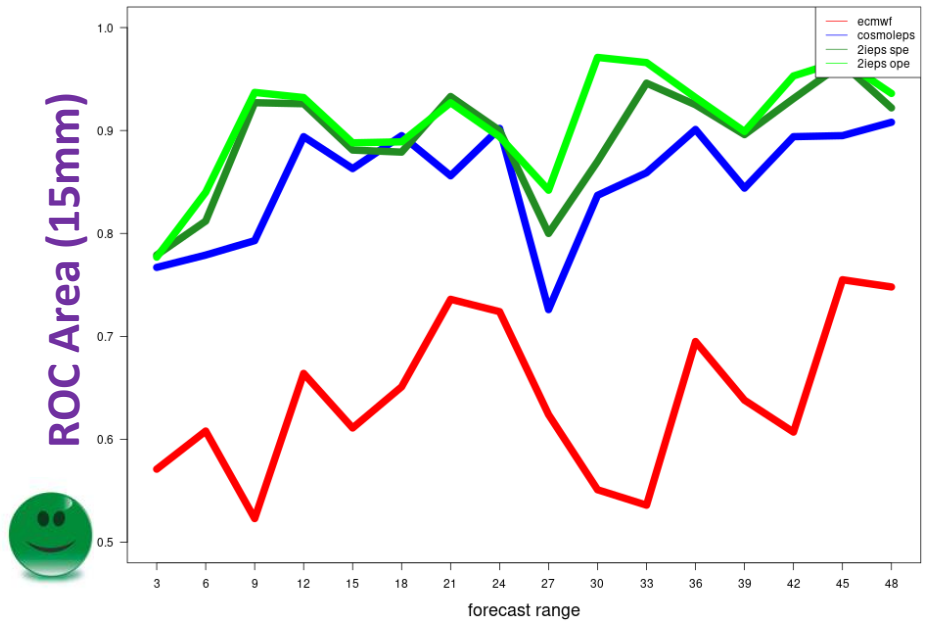
Synoptic situation and observed precipitation in the two most significant days of the verification period from 28 February to 9 March 2020

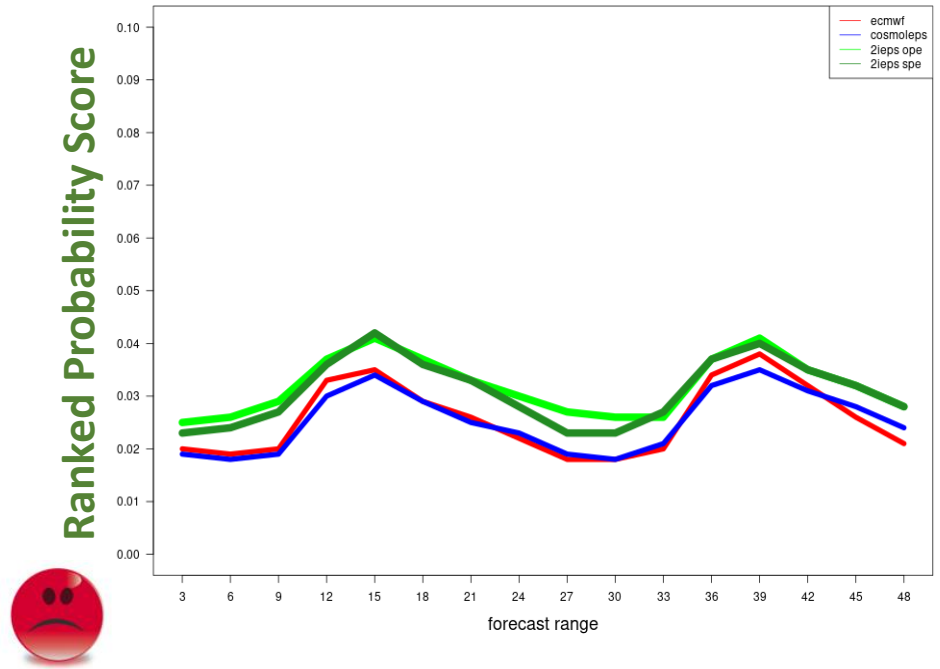


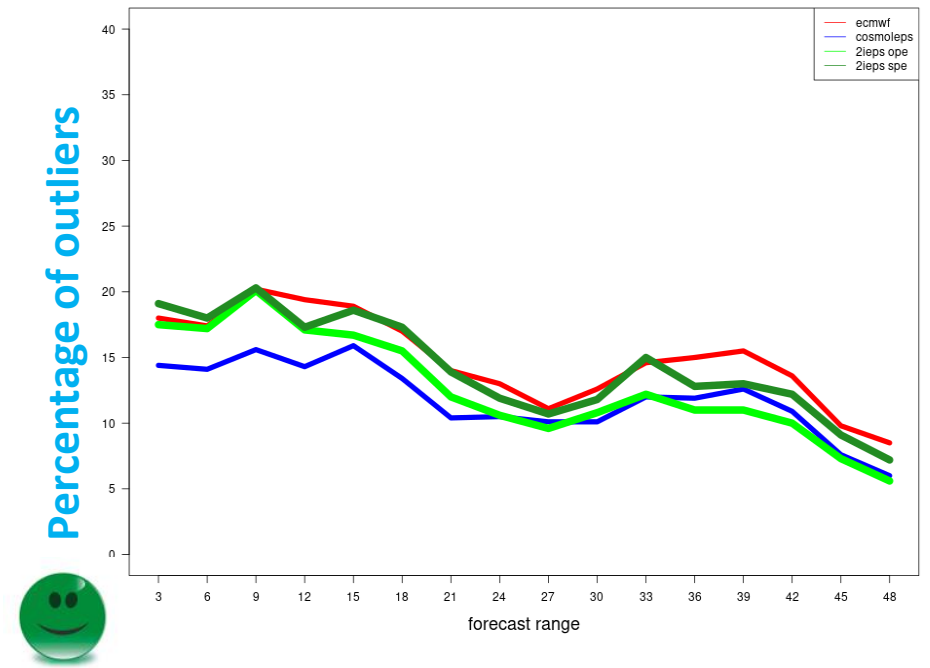
28 February – 9 March precipitation

Initialize COSMO-2I-EPS at 00UTC (use IC from KENDA at 00UTC and BC from COSMO-ME-EPS at 00UTC)









ECMWF ENS

COSMO-LEPS

COSMO-2I-EPS ope

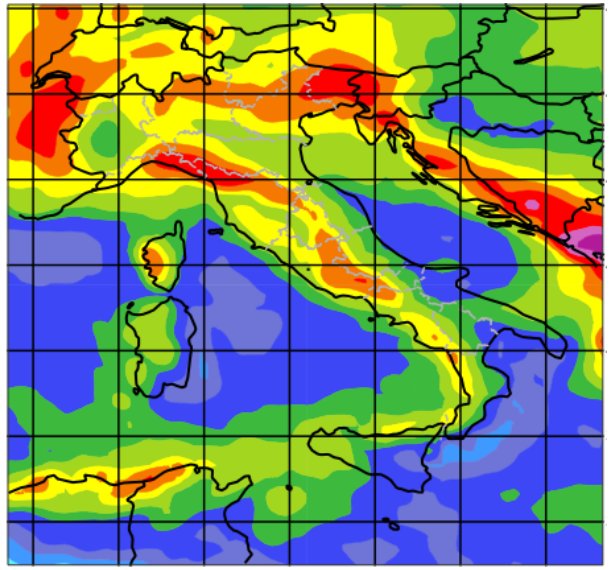
COSMO-2I-EPS init 00

Also in this experiment the good results of COSMO-2I-EPS are confirmed only for the percentage of outliers and for the ROC Area.

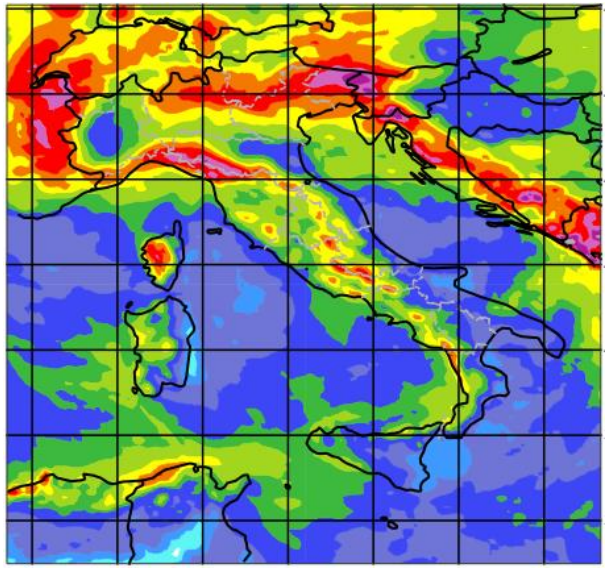
The use of init at 00 UTC leads to a slight deterioration or no variation of most of the COSMO-2I-EPS scores.

Often the gap with COSMO-LEPS continues to be marked.

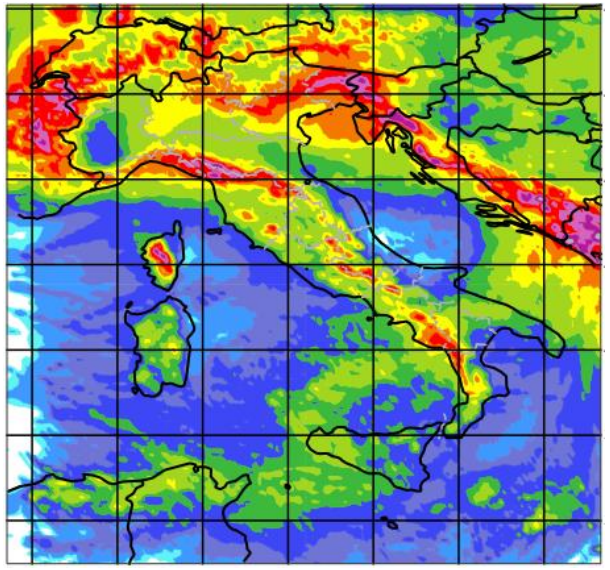
Initialize COSMO-2I-EPS at 00UTC (use IC from KENDA at 00UTC and BC from COSMO-ME-EPS at 00UTC)



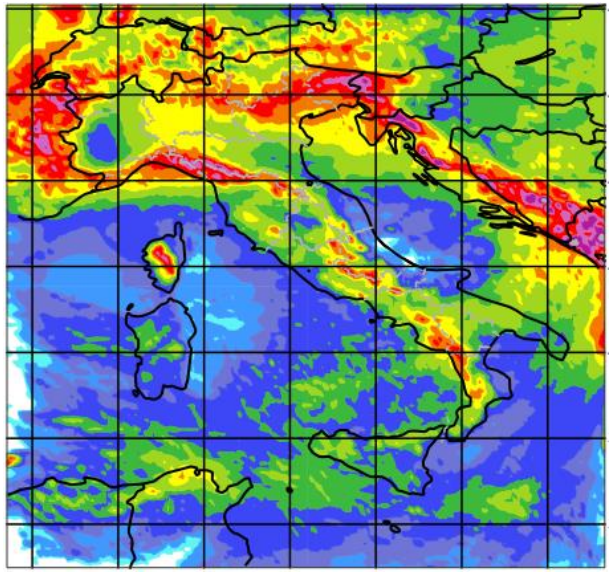
ECMWF



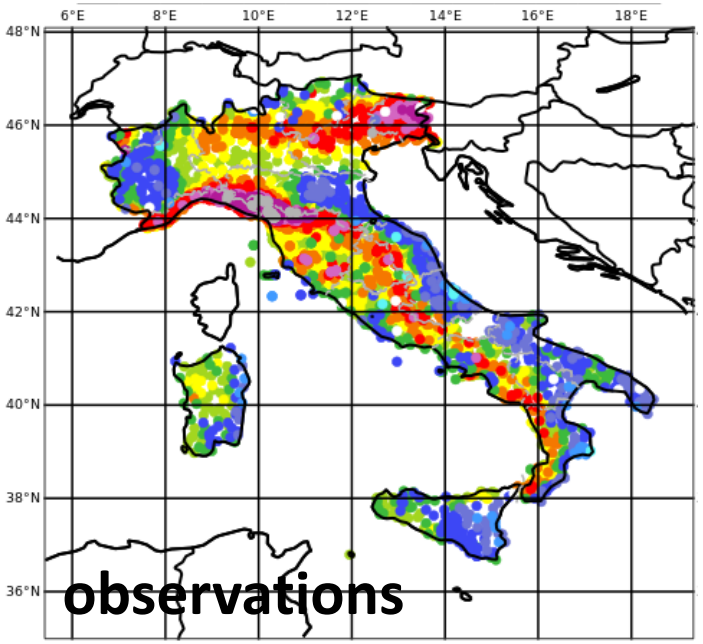
COSMO-LEPS



COSMO-2I-EPS ope

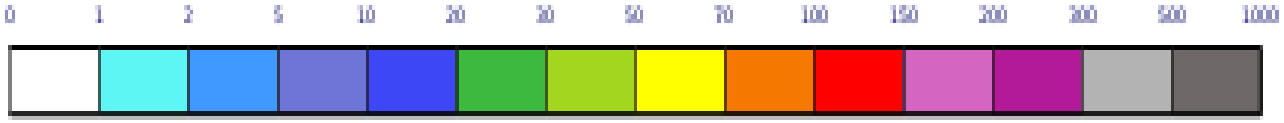


COSMO-2I-EPS init 00



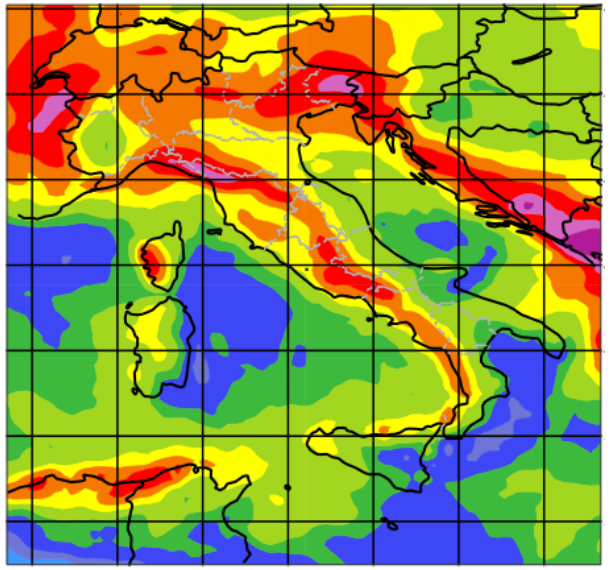
observations

whole period
Step 0-24
ensmean

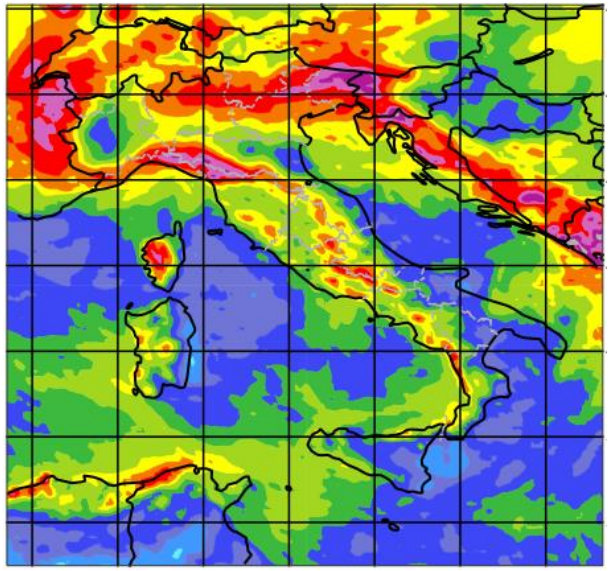


The ensemble mean of the forecasted precipitation over the entire verification period by COSMO-2I-EPS, and generally by the COSMO-based EPS, is in line with the values recorded by the rain-gauge network of the National Civil Protection Department.

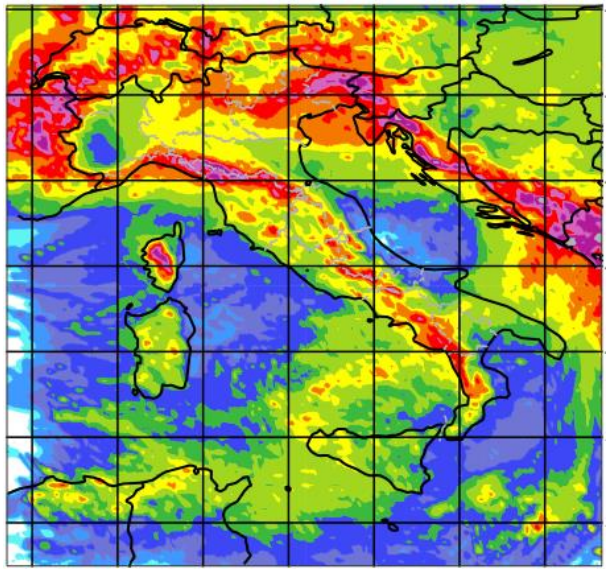
Initialize COSMO-2I-EPS at 00UTC (use IC from KENDA at 00UTC and BC from COSMO-ME-EPS at 00UTC)



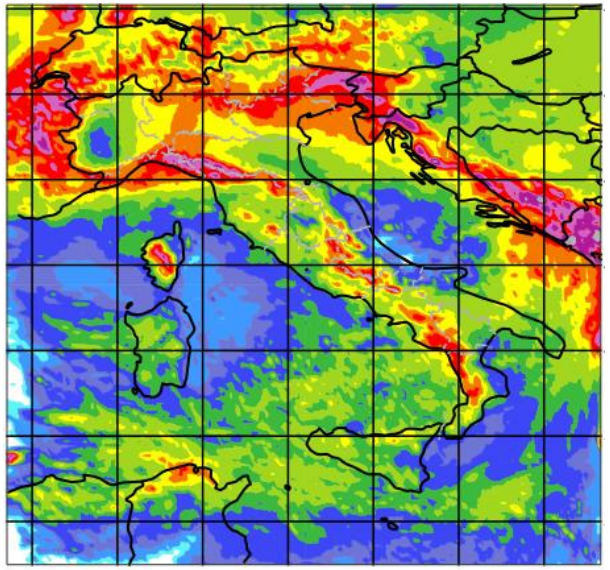
ECMWF



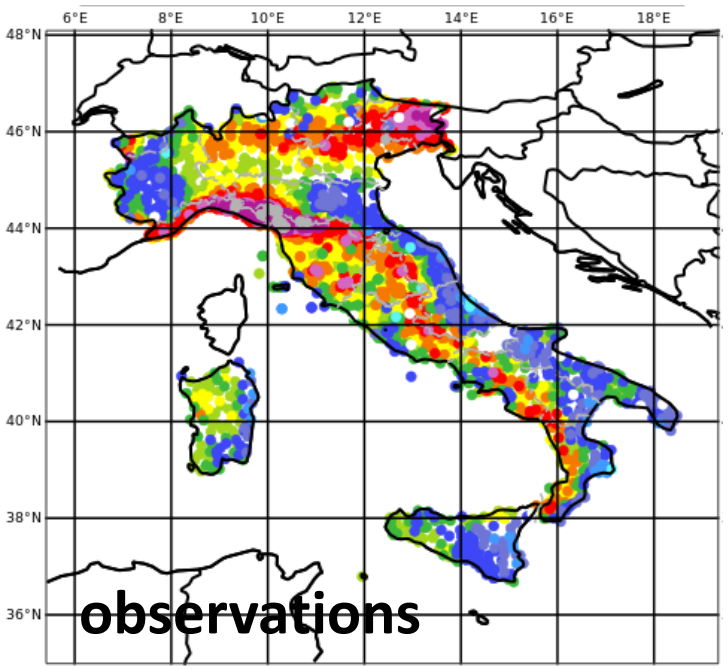
COSMO-LEPS



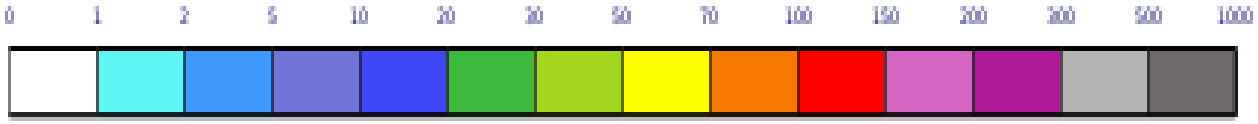
COSMO-2I-EPS ope



COSMO-2I-EPS init 00



whole period
Step 0-24
90° percentile



The 90° percentile of the forecasted precipitation over the entire verification period by COSMO-2I-EPS, and in general by the COSMO-based Ensemble Prediction System, is correctly represent the distribution and in part also the quantities of precipitation peaks above 150/200mm recorded by the rainfall network of the National Civil Protection Department.

Conclusions

Given the not encouraging results obtained so far we will do the other experiments listed to try to understand the cause of the bad performance of COSMO-2I-EPS

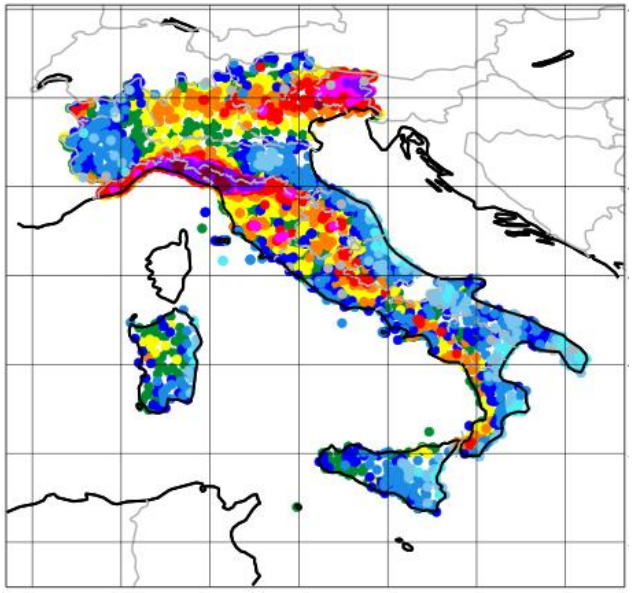
Based on the results shown so far we are not able to answer the following two questions:

1) **Every time the total precipitation is plotted on maps (both in case of single events and in case of longer periods) COSMO-2I-EPS seems to be always in good agreement with the observed data both in terms of intensity and distribution of the phenomena. In particular COSMO-2I-EPS seems to predict precipitation better than other Ensemble Prediction Systems. So why the scores of COSMO-2I-EPS are worse than COSMO-LEPS and ECMWF ENS?**

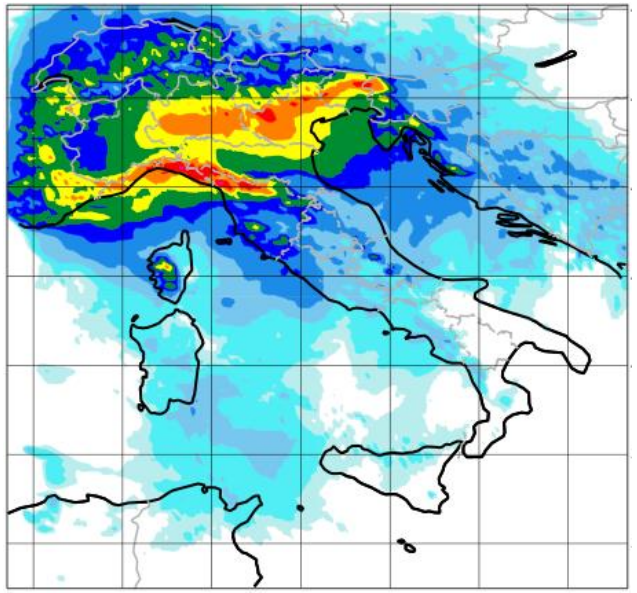
2) Why COSMO-2I-EPS always shows good results for ROC AREA and for the percentage of outliers, while in all the other scores it is the worst Ensemble Prediction System?

Initialize COSMO-2I-EPS at 00UTC (use IC from KENDA at 00UTC and BC from COSMO-ME-EPS at 00UTC): error on number of vertical levels

observations



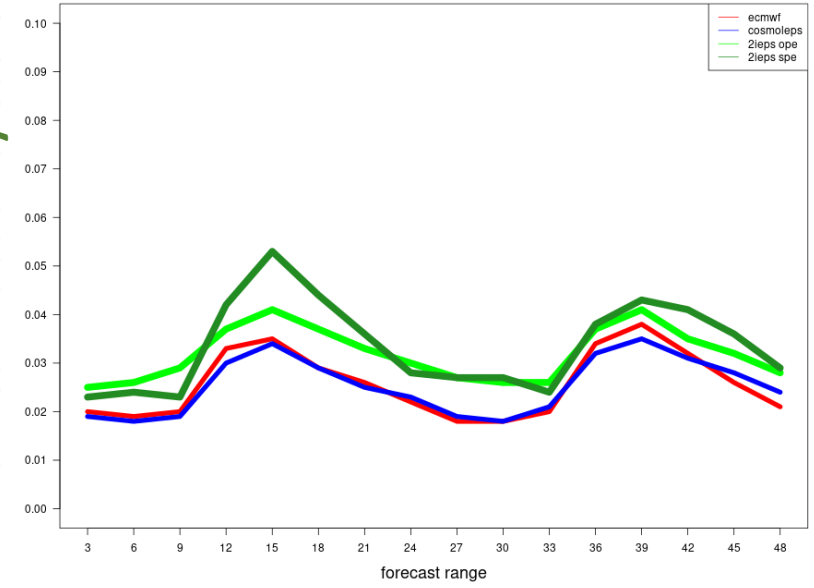
COSMO-2I-EPS spe



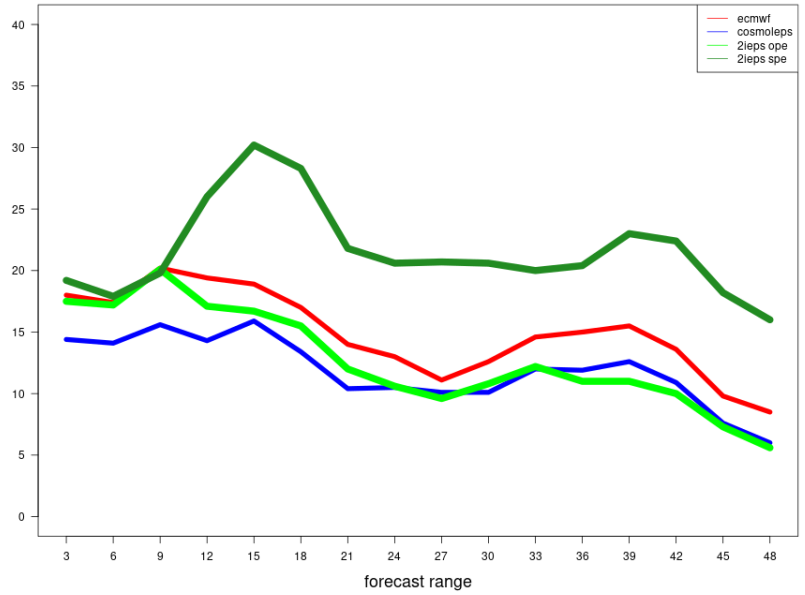
whole period
Step 24-48
ensmean

ECMWF ENS
COSMO-LEPS
2I-EPS ope
2I-EPS init 00

Ranked Probability Score



Percentage of outliers



In doing this type of verification initially an error had been made with the number of vertical levels of COSMO-2I-EPS, so that the grib used had failed.

In fact from the total precipitation map it is evident that there is something wrong.

The same can be said by observing the percentage of outliers and the ROC Area (not shown), while all other scores do not show significant degradation.

- Why only percentage of outliers and ROC Area show that there is an error?

So these two scores are the only ones that reward COSMO-2I-EPS gribs when they contain no errors and get worse when the same gribs are not prepared correctly.

- Are all the other scores really effective in evaluating the performance of an Ensemble Prediction Systems?

Thanks for the attention