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	2018
Project Title:	Investigation of case studies using COSMO-based deterministic and ensemble systems
Computer Project Account:	spcoleps
Principal Investigator(s):	Montani Andrea (Italy)
Affiliation:	Arpae-SIMC Sercvizio IdroMeteoCLiam, viale Silvani 6 40122 Bologna (Italy)
Name of ECMWF scientist(s) collaborating to the project (if applicable)	…………………………………………………………… ……………………………………………………………
Start date of the project:	2018
Expected end date:	2020

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)			2.100.000	500.000
Data storage capacity	(Gbytes)			600	200

Summary of project objectives

(10 lines max)

To assess the sensitivity of COSMO forecast skill to the use of the newly implemented Bechtold convection scheme; experiments will be performed in both deterministic and ensemble mode.

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 run COSMO model with the Bechtold schemes for a number of past cases, so as to perform a systematic intercomparison between the operational set-up and the experimental one.

This is described in the attached report.

List of publications/reports from the project with complete references

Vasconi M., Montani A., Paccagnella T., 2018. Sensitivity of forecast skill to the parameterisation of moist convection in a limited-area ensemble forecast system, Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2018-21, in review.

Summary of plans for the continuation of the project

(10 lines max)

It is planned to perform intercomparison during the winter season between COSMO-T (COSMO model run with the operational Tiedtke scheme) and COSMO-B (COSMO model run with Bechtold scheme). Modifications to the current implementation of the Bechtold scheme within COSMO will be also tested.

Sensitivity of precipitation forecast skill to the parameterisation of moist convection in COSMO-based ensemble systems

Matteo Vasconi

Department of Physics and Astronomy, University of Bologna, Italy and Arpae-SIMC, Bologna, Italy

Andrea Montani, Tiziana Paccagnella

Arpae-SIMC, Bologna, Italy

1 Introduction

The parameterisation of convection in limited-area models is an important source of uncertainty as regards the spatio-temporal forecast of precipitation. As for the limited-area model COSMO, hitherto, only the Tiedtke convection scheme (Tiedtke, 1989) was available for the operational runs of the model in convection-parameterised mode. In addition to this the Bechtold scheme, implemented in ECMWF global model, has recently been adapted for COSMO applications. The development and implementation of ensemble systems in which different convection schemes are used, provides an opportunity to upgrade state-of-the-art probabilistic systems at the convection-parameterised scale. The sensitivity of the COSMO model forecast skill to the use of either the Tietdke or the Bechtold (Bechtold et al., 2008; 2014) schemes is assessed by performing different sets of experiments. This study is part of the CIAO COSMO Priority Task.

The performance of COSMO model run with the different schemes is investigated in ensemble mode with particular attention to the types of forecast errors (e.g. location, timing, intensity) provided by the different convection schemes in terms of total precipitation.

A 10-member ensemble has been run for approximately 2 months with the Bechtold scheme, using the same initial and boundary conditions as members 1-10 of the operational COSMO-LEPS ensemble system (which has 20 members, all run with the Tiedtke scheme). The performance of these members is assessed and compared to that of the system made of members 1-10 of COSMO-LEPS in terms of total precipitation prediction.

Finally, the performance of an experimental 20-member ensemble system (which has 10 members run with the Bechtold plus 10 members run with the Tiedtke scheme) is compared to that of operational COSMO-LEPS over the 2-month period. The new system turned out to have higher skill in terms of precipitation forecast with respect to COSMO-LEPS over the period. In this approach the use of the Bechtold scheme is proposed as a perturbation for the COSMO-LEPS ensemble, relatively to how uncertainties in the model representation of the cumulus convection can be described and quantified.

2 System description and methodology of analysis

Some experiments have been performed, in order to evaluate the COSMO model performance in ensemble mode when it is run either with the Tiedtke or the Bechtold scheme, so as to assess overall abilities and shortcomings of the system (Vasconi, 2017). Firstly, we have built a test suite to run a 10-member ensemble with the Bechtold scheme (referred to as

Acronym	Ensemble size	Convection scheme	ICs-BCs
COSMO-LEPS	20	Tiedtke	from ECMWF-ENS
Cleps-10B	10	Bechtold	the same as 1-10 of COSMO-LEPS
Cleps-10T	10	Tiedtke	the same as 1-10 of COSMO-LEPS
Cleps-20bt	20	Bechtold + Tiedtke	the same as COSMO-LEPS

Table 1: Main features of the ensemble systems of Section 2

Cleps-10B), which uses the same initial and boundary conditions as members 1-10 of the operational COSMO-LEPS (which has 20 members, all run with the Tiedtke scheme). This suite has been run from 28$^{\text{th}}$ March to 31$^{\text{th}}$ May 2017 with an integration domain covering Central-Southern Europe and Italy (shown in Fig. 1), at the horizontal resolution of about 7 km and 40 vertical layers, and with a 132-hours forecast range, always starting at 00 UTC. In particular, the sensitivity of the ensemble system to the different parameterisation schemes has been assessed by comparing the performance of Cleps-10B to that of Cleps-10T, which is the 10-member ensemble provided by members 1-10 of COSMO-LEPS, the operational ensemble system of the COSMO consortium, over the verification period.

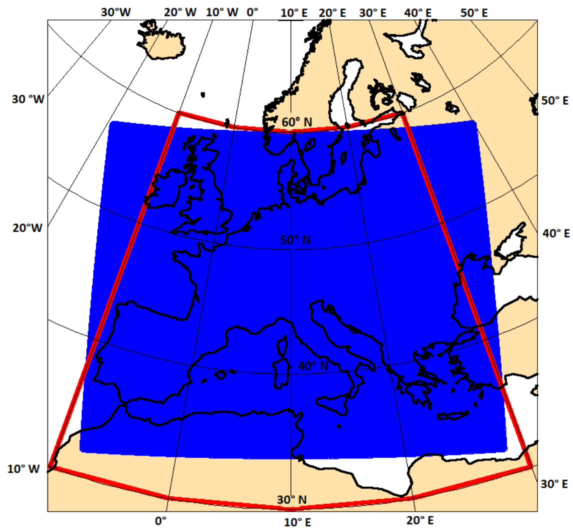

Figure 1: COSMO-LEPS integration domain (blue area) and clustering area (inside the red line).

A further step in the study of COSMO ensemble system sensitivity to different formulation of moist convection is the implementation of a new probabilistic system, hereafter Cleps20bt, in which a multi-physics approach in the model representation of the cumulus convection is followed. This system is generated by adding the members of Cleps-10B to members 11-20 of COSMO-LEPS. Therefore, Cleps20bt has 10 members run with the Bechtold scheme plus 10 members run with the Tiedtke scheme and no duplication of initial and boundary conditions. The basic idea of the Cleps20bt implementation is that certain closure parameters used in model formulation (as for the moist convective processes) may be based on approximate physical knowledge. As a consequence their values may be somewhat arbitrary, or they may have been tuned to give optimal results for test cases that are not necessarily representative of more general applications and/or for applications at high resolution. A summary of the ensembles features is presented in Table 1.

The performance of the ensemble systems was analysed by considering the probabilistic prediction of 6-h cumulated precipitation exceeding a number of thresholds for forecast up to 132 hours over the 2-month period. Since precipitation has a high-spatial variability, a high-

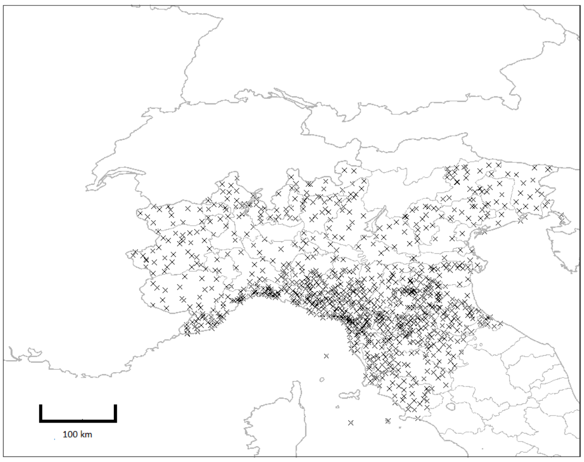

Figure 2: Observation network used for verification.

density network, made of about 1000 stations over Northern Italy (Fig. 2), has been adopted in order to assess the predictive skill of the ensemble systems. For the comparison of the model forecasts against station reports the grid point closest to the observation one is selected. In particular the performance of the different ensemble systems of Table 2 is examined for six different 6-h cumulated precipitation thresholds: 1, 5, 10, 15, 25, 50 mm/6-h. Several thousands of events were reported for the first two thresholds, and several hundreds for the 15 mm/6-h threshold. On the other hand it is immediately worth pointing out that, when considering the highest thresholds (25, 50 mm/6-h), a low number of occurrences, even below 10 for the 50 mm/6-h, was found over the verification period. As a consequence this does not allow any solid statistical conclusion on the effective performance of the system for these events over the period.

For each forecast range, the model performance has been evaluated by computing the following "traditional" probabilistic scores (Wilks, 1995): the Brier Skill Score (BSS), the Ranked Probability Skill Score (RPSS), and the Percentage of Outliers (Buizza, 1997). A summary table of the verification features is reported in Table 2.

Verification features	
variable:	6-h cumulated precipitation (00-06, 06-12,..UTC);
Period:	from 28th March to 31th May 2017 (about 60 days);
region:	Northern Italy;
method:	nearest grid-point; no-weighted fcst;
obs:	non-GTS network, no obs error;
fcst ranges:	0-6 h, 6-12 h,..., 126-132 h;
thresholds:	1, 5, 10, 15, 25, 50 mm/6 h;
systems:	Cleps-10B vs Cleps-10T, Cleps20bt vs COSMO-LEPS;
scores:	BSS, RPSS, Percentage of Outliers.

Table 2: Main features of the verification configuration for the ensembles

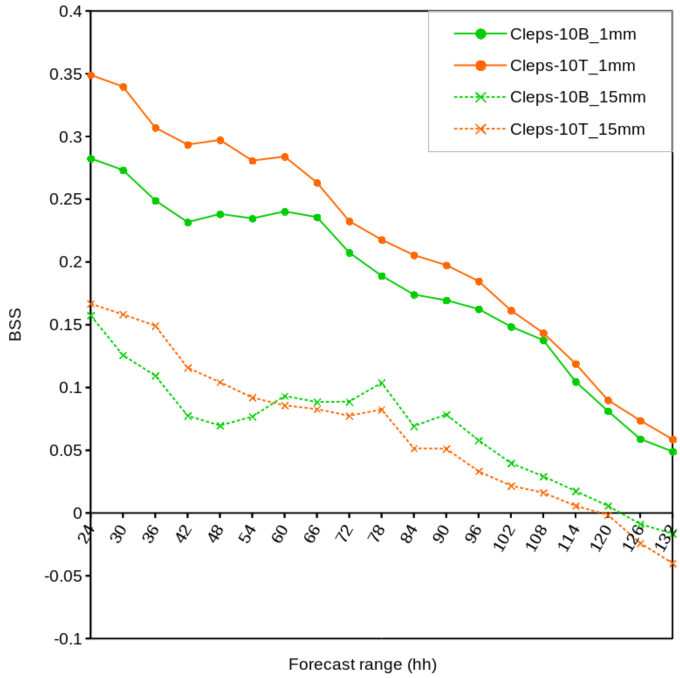

Figure 3: 24-h running mean of BSS in Cleps-10T and Cleps-10B (orange and green line respectively) for 1 mm/6-h and 15 mm/6-h (solid and dashed line respectively) thresholds.

3 Comparison of 10-member ensemble system run with different schemes

The BSS (Brier Skill Score) for the Cleps-10T and Cleps-10B is presented in Fig.3. A 24-h running mean is here applied to "smooth" the diurnal cycle in model performance, improving the readability of the plot. This score tries to represent a quantitative estimate of the added value detectable in precipitation prediction by using the model forecast rather than a reference one (in this case, climatology of the observed sample over the verification period). The attention has been focused on two thresholds (1 mm/6-h and 15 mm/6-h), which have a quite large number of occurences (higher than 1000 for the former, some hundreds for the latter) over the verification period.

It is worth noticing that the BSS shows clearly the loss of predictability with increasing forecast range for both systems. The model forecast has added value with respect to the reference climatology up to +120 hours. However the plot shows a different skill of the 2 systems when different thresholds and forecast ranges are considered. Over the verification period, Cleps-10T performs generally better than Cleps-10B for the lower threshold (1 mm/6-h), while the opposite is true in high precipitation rates prediction for forecast ranges from 3 days onwards. In other words, the ensemble systems seem to describe different types of forecast errors, possibly related to the different convection schemes (Vasconi, 2017).

In addition to this, the RPSS (Ranked Probability Skill Score) of this system has been computed for different forecast ranges and compared to that of COSMO-LEPS during the same period. The plot in Fig. 4 shows a better performance of Cleps-10T for the forecast ranges up to +48 hours.

These results can be seen consistent with the theory according to which the ensemble systems which are run using either convection schemes can describe a larger variety of uncertainty and errors in precipitation prediction.

Finally, the skill of the two systems has been assessed in terms of Percentage of Outliers (that

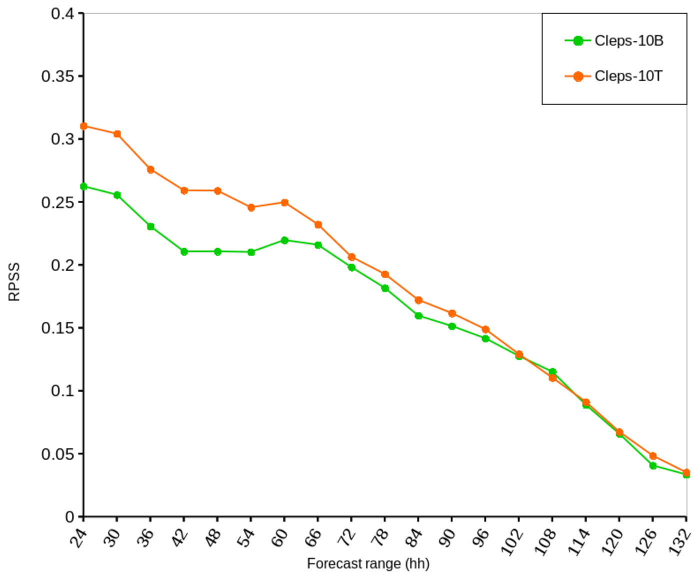

Figure 4: 24-h running mean of RPSS in Cleps-10T (orange line) and Cleps-10B (green line).

is the cases in which observed rainfall value is not inside the ranges of possible values predicted by the ensemble members, Fig. 5). Firstly it is worth pointing out that the total percentage of outliers (left panel) for both systems tends to decrease with increasing forecast range because of the increasing spread with time between the ensemble members. A better performance of Cleps-10T, which has a lower number of outliers than Cleps-10B, can be noticed, in particular for the earlier forecast ranges. The right panel of Fig. 5 represents respectively the fraction of points in which observations lie above/below the range of predicted values by the ensemble system.

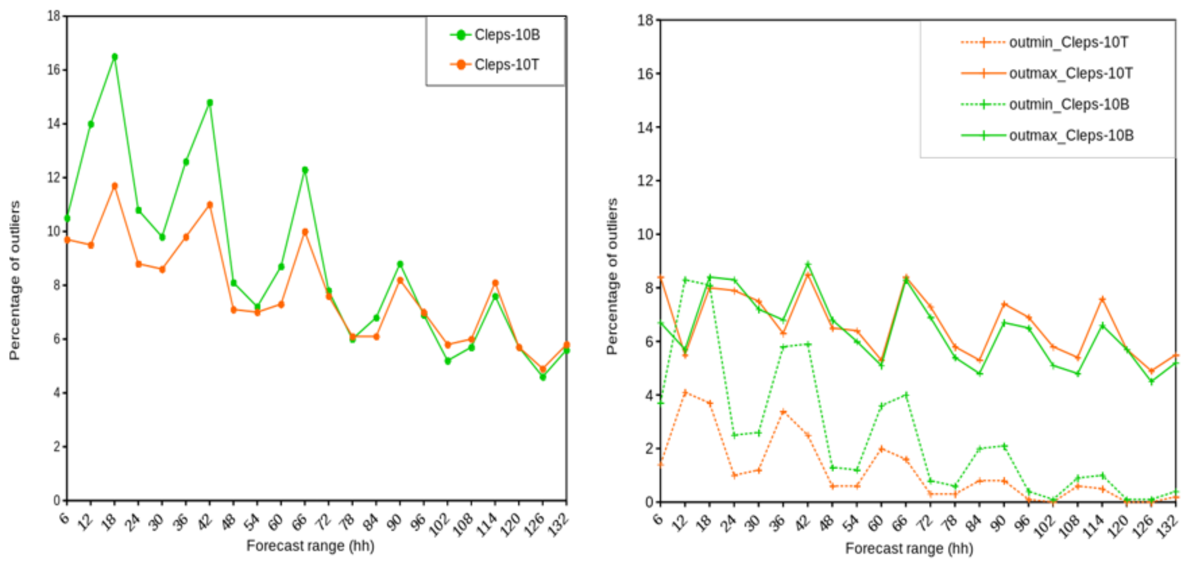

Figure 5: Left panel: Percentage of outliers for different forecast ranges in Cleps-10T and Cleps-10B (orange and green line respectively). Right panel: Percentage of outliers above/-below maximum/minimum predicted values

A large amount of outliers below the minimum forecast value, indicative of an overestimation of minima of precipitation amount by Cleps-10B runs, can be seen. In particular the percentage of outliers lying below the minimum predicted values is higher for Cleps-10B than for

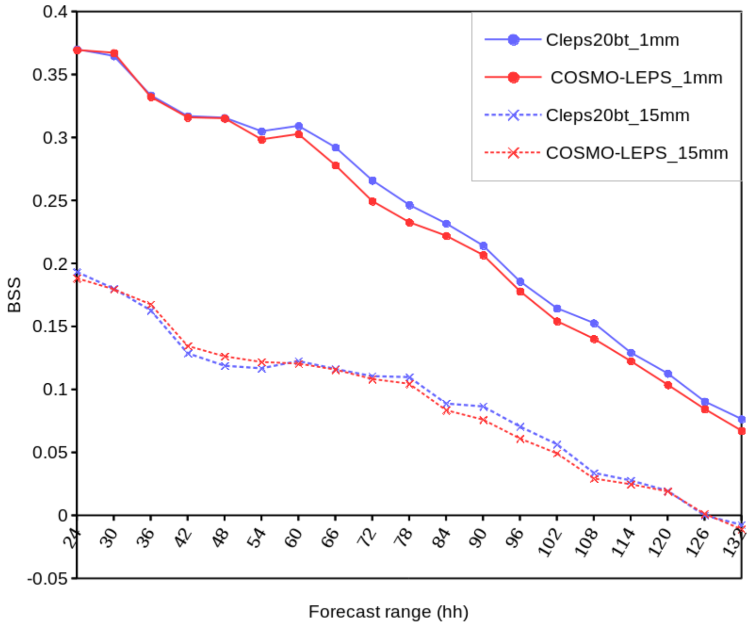

Figure 6: 24-h running mean of BSS values for 6-h accumulated precipitation exceeding 1 mm and 15 mm (solid and dashed line respectively) for different forecast ranges in COSMO-LEPS (red line) and Cleps20bt (blue line).

Cleps-10T for all the forecast ranges studied. This seems to indicate that members with the Bechtold scheme tend to produce some light prepitation also when it is not observed. On the other hand, the fraction of analysis point above the maximum tends to be similar or slightly lower for Cleps-10B. This excessive drizzle effect could be due to the shallow convection treatment adopted by the Bechtold scheme. This scheme in fact allows "shallow convection" to produce precipitation, whereas the Tiedtke scheme does not. It is possible that further tuning of the Bechtold scheme, when adopted at high resolution, is necessary to address this “drizzle” issue.

4 Performance of Cleps20bt and comparison with that of COSMO-LEPS

A quantitative evaluation of Cleps20bt skill in terms of precipitation forecast over the the same period is then presented. The basic idea of this study is that ensemble systems which are run using either convection schemes can describe a larger variety of uncertainty and errors in precipitation prediction (Vasconi, 2017). Thus the implementation of ensemble systems in which the two schemes are "mixed" seems to be a reasonable issue to deal with uncertanties due to the ambiguity linked to the use of a scheme or the other. It is worth pointing out that the implementation of this experimental system is consistent only because the average skill of the model when it is run in ensemble mode with the Bechtold scheme turned out to be roughly indistinguishable, from a statistical point of view, from that provided by running the model with the Tiedtke scheme, as shown in the previous Section. In fact, in a well-constructed ensemble, the skill of each individual member, averaged over a large number of events, should be approximately identical not to introduced biases and/or systematic errors in the ensemble members distribution.

The forecast skill in terms of precipitation of Cleps20bt is then assessed and compared to that of COSMO-LEPS. The main results of this study are presented in the following plots.

In Fig. 6 BSS (Brier Skill Score) is presented for different forecast ranges by considering several thresholds. In particular the focus is on the same threshold as for the 10-member case,

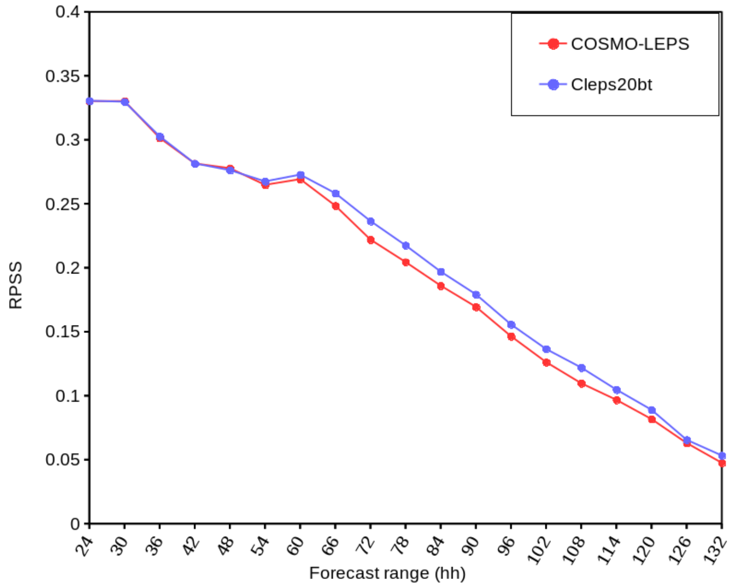

Figure 7: 24-h running mean of RPSS values for 6-h accumulated precipitation for different forecast ranges in COSMO-LEPS (red line) and Cleps20bt (blue line).

for which a relative large number of events has been reported (1 mm/6-h and 15 mm/6-h). In order to provide an overall description of the model system performance for the different precipitation thresholds, the values reported in the plot are obtained, once again, by computing the running mean of the 6-h precipitation forecast skill over 24 hours. The plot shows that Cleps20bt has higher values of BSS than COSMO-LEPS for the thresholds reported, especially for forecast ranges from 42 hours onwards (blue and red lines respectively).

In addition to this, the RPSS (Ranked Probability Skill Score) of this system has been computed for different forecast ranges and compared to that of COSMO-LEPS during the same period. The comparison between the 24-h running mean of RPSS for the two systems is presented in Fig. 7. Also in this case a better performance of Cleps20bt than that of COSMO-LEPS is evident for forecast ranges from 2 days onwards: for example RPSS in the forecast range +60-66 hours is about 5% higher in Cleps20bt than in COSMO-LEPS; it is about 10% higher in the new system for +90-96 h, +96-102 h ranges.

A similar behaviour can be detectable also in other scores (Brier Score and ROC Area), which are not presented here.

Finally the performance of the systems is evaluated in terms of the percentage of outliers (left panel in Fig. 8). In addition to this, similarly to the 10-member ensembles case, the percentage of outliers are discriminated between the fractions of points in which observed values lay outside the forecast range over the full verification period (right panel in Fig. 8). The percentage of outliers is reduced in Cleps20bt over most of the forecast ranges with respect to COSMO-LEPS, especially from 3 days (+72 hours) onwards.

The right panel in Fig. 8 shows that the total percentage of outliers is reduced in Cleps20bt as a consequence of a decrease in the number of points wherethe total precipitation maxima are underestimated compared to COSMO-LEPS. In fact the fraction of observations found above the maximum forecast value is lower in Cleps20bt than in COSMO-LEPS, for most of forecast ranges, especially in the medium range (from +72 hours onwards). This is a quite encouraging result because Cleps-20bt turns out to perform better than the operational COSMO-LEPS in forecasting the possible peaks in cumulated precipitation over the 2-month period. It is worth underlining that the probabilistic forecast of these values is one of the most important

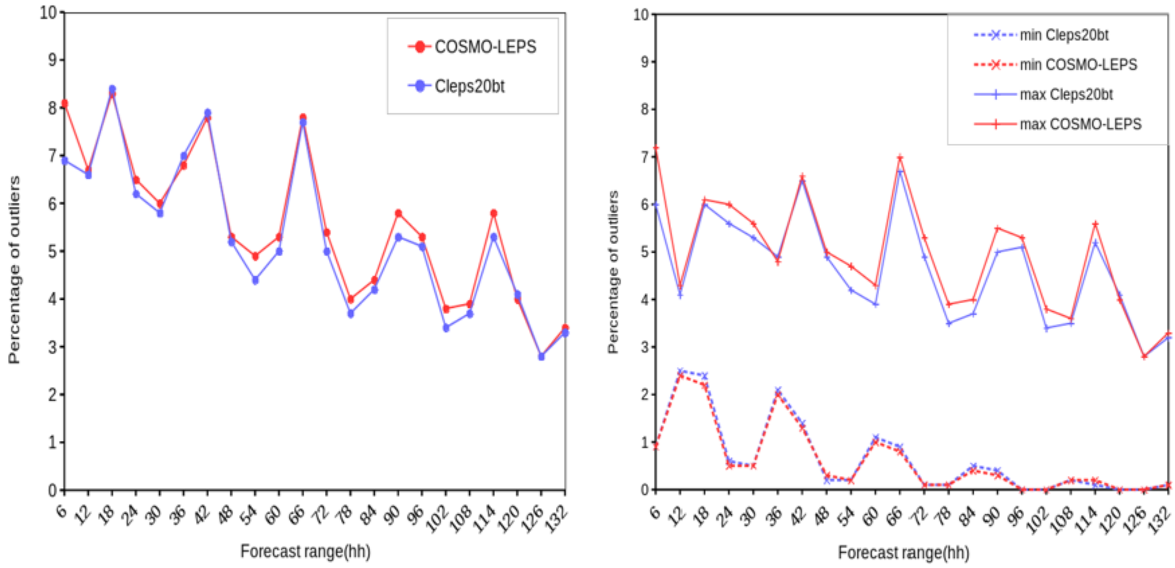

Figure 8: Left panel: Percentage of outliers for different forecast ranges in COSMO-LEPS (red line) and Cleps20bt (blue line). Right panel: Percentage of outliers above/below maximum/minimum predicted values.

issue of operational systems, because it regards the correct prediction of heavy rainfall events, which may have a high impact on the society.

This result, together with those presented in this section, substantially agrees with the idea that, by adding a physical perturbation to the system (like what we have done in this work using an ensemble system in which two different moist convective schemes are used), we can obtain a more appropriate description of the phase-space of all possible future atmospheric states which are compatible with the uncertain model formulation of the moist convection sub-grid processes. Thus, according to this experimentation, the generation of a multi-physics ensemble system provides a positive impact on the forecast capability at high resolution. This is especially true in early-medium range, when model errors start playing an important role and it is crucial for an ensemble system to provide an accurate description of the different sources of forecast deficiency (Vasconi, 2017).

4 Summary and Outlook

The impact of the use of two moist convection schemes (the Tiedtke and Bechtold schemes) has been studied in ensemble mode. Firstly a 10-member ensemble with the Becthtold scheme (Cleps-10B), which uses the same initial and boundary conditions as members 1-10 of the operational COSMO-LEPS, has been run has been run for approximately 2 months. The performance of these members has been assessed and compared again to that of Cleps-10T, the 10-member ensemble made of members 1-10 of COSMO-LEPS; in particular the spread-/skill relation of the two 10-member ensemble in terms of total precipitation is evaluated. Verification has been performed for precipitation events occurred over Northern Italy (using the forecast at the gridpoints nearest to about 1000 stations) from 28th March to 31th May 2017. The average skill of the Cleps-10B runs turned out to be substantially indistinguishable, from a statistical point of view, from that provided by the Cleps-10T ones. However a deeper analysis suggests that the two ensemble systems are characterised by different types of forecast errors. Therefore a new 20-member ensemble system (Cleps20bt, which has 10 members run with Bechtold plus 10 members run with Tiedtke and no duplication of boundary conditions) has been implemented. In this system the Bechtold scheme is used as a perturbation for the COSMO-LEPS ensemble, so as to provide a quantitative description of uncertainties linked to the model representation of the cumulus convection. Cleps20bt has been shown to have higher skill than COSMO-LEPS over the verification period. In addition to this, the comparison of the Percentage of Outliers in the two systems shows a reduction in the fraction of observed points lying outside the maximum or minimum forecast value in Cleps20bt. These results suggest that the use of a probabilistic system in which a multiple moist convection formulation is used, provides the opportunity to have a more comprehensive description of the uncertainties in total precipitation forecast linked to the sub-grid cumulus representation.

However, further work is necessary on this topic. Firstly the sensitivity of model forecast skill in terms of other variables (2-m temperature, humidity, 10m- wind speed) has to be assessed. In fact the use of different schemes is expected to have a great impact also on these variables at high resolution scales. In addition to this, we plan to perform runs in ensemble mode for other seasons and at 5 km of horizontal resolution.

References

[19] Arakawa, A., 2004. The cumulus parameterization problem: Past, present, and future. *J. Cli.*, **17**, 2493-2525.

[2] Bechtold, P., M. Köhler, T. Jung, F. Doblas-Reyes, M. Leutbecher, M. Rodwell, F. Vitart and G. Balsamo, 2008b. Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales. *Quart. J. Roy. Meteor. Soc.*, **134**, 1337-1351.

[19] Bechtold, P., Semane, N., Lopez, P., Chaboureau, J., Beljaars, A. and Bormann, N. 2014. Representing Equilibrium and Nonequilibrium Convection in Large-Scale Models. *J. Atmosph. Sc.*, **71**, 734-753.

[19] Bechtold, P., 2017. Atmospheric moist convection. Meteorological Training Course Lecture Series, ECMWF.

[19] Buizza, R., 1997. Potential Forecast Skill of Ensemble Prediction and Spread and Skill Distributions of the ECMWF Ensemble Prediction System. *Mon. Wea. Rev.*, **125**, 99-119.

[19] Doms. G. and Baldauf, M., 2015. A Description of the Nonhydrostatic Regional COSMO-Model. Part I: Dynamics and Numerics. (www.cosmo-model.org).

[19] Kuo, HL and Raymond, WH, 1980. A Quasi-one-Dimensional Cumulus Cloud Model and Parameterization of Cumulus Heating and Mixing Effects. *Mon. Wea. Rev.*, **108**, 991-1009.

[19] Marsigli, C., Montani, A., Nerozzi, F., Paccagnella, T., Tibaldi. S., 2001. A strategy for high-resolution ensemble prediction. Part II: limited-area experiments in four Alpine flood events. *Quart. J. Roy. Meteor. Soc.*, **127**, 2095-2115.

[19] Montani, A., Capaldo, M., Cesari, D., Marsigli, C., Modigliani, U., 2003a. Operational limited-area ensemble forecasts based on the Lokal Modell. ECMWF Newsletter 98, 2-7. Available at: http://www.ecmwf.int/publications/.

[19] Montani, A., Marsigli, C., Nerozzi, F., Paccagnella, T., Tibaldi, S., and Buizza, R., 2003b. The Soverato flood in Southern Italy: performance of global and limited-area ensemble forecasts. *Nonlin. Proc. Geophys.*, **10**, 261-274.

[19] Montani, A., Cesari, D., Marsigli, C., and 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*, **63**, 605-624.

[19] Mullen, SL. and Buizza, R., 2001. Quantitative Precipitation Forecasts over the United States by the ECMWF Ensemble Prediction System. *Mon. Wea. Rev.*, **129**, 638-663.

[19] Randall, D., Arakawa, A., Khairoutdinov, M., and Grabowski, W. 2003. Breaking the Cloud Parameterization Deadlock. *Bull. Amer. Met. Soc.*, **115**, 1547-1564.

[19] Steppeler, J., Doms, G., Schättler, U., Bitzer, W., Gassmann, A., Damrath, U., Gregoric, G., 2003. Meso-gamma scale forecasts using the nonhydrostatic model LM. *Meteor. Atmos. Phys.*, **82**, 75-96.

[19] Tennekes, H., 1986. Forecasting forecast skill. *Proceedings of the ECMWF Workshop on Predictability in the Medium and Extended range*, Reading, England.

[19] Tibaldi, S., Paccagnella, T., Marsigli, C., Montani, A. and Nerozzi, F., 2006. Limited area ensemble forecasting: the COSMO model. Predictability of Weather and Climate. Cambridge University Press, 489-513.

[19] Tietdke, M., 1989. A Comprehensive Mass Flux Scheme for Cumulus Parameterization in Large-Scale Models. *Mon. Wea. Rev.*, **117**, 1779-1800.

[19] Vasconi, M., 2017. Sensitivity of forecast skill to the parameterisation of moist convection in the COSMO model. *Master's Thesis in Physics of the Earth System*, University of Bologna, Italy. Available at http://amslaurea.unibo.it/14566/1/Tesi_Magistrale_Vasconi.pdf.

[19] Wilks, D., 1995. Statistical methods in the atmospheric sciences. International Geophysics Series, Vol 59. Academic Press.