# SPECIAL PROJECT FINAL REPORT

All the following mandatory information needs to be provided.

| | |
|---|---|
| **Project Title:** | Numerical studies for short-range forecast and nowcasting in complex terrain areas |
| **Computer Project Account:** | SPITGARB |
| **Start Year - End Year:** | August 2020 – December 2021 |
| **Principal Investigator(s)** | Valeria Garbero (mcy0), valeria.garbero@arpa.piemonte.it |
| **Affiliation/Address:** | Arpa Piemonte, Italy |
| **Other Researchers (Name/Affiliation):** | Massimo Milelli (mcy), massimo.milelli@arpa.piemonte.it |

The following should cover the entire project duration.

## Summary of project objectives

In the last years, the numerical weather prediction (NWP) models constantly improved, mostly due to major improvements in high-performance computing, which allows a finer horizontal grid resolution. Complex terrain, however, still poses a challenge for NWP models. In the framework of the EU Project RISK-GEST (PITEM RISK, Interreg 2014-2020 Alcotra IT-FR), high resolution numerical modelling has been used to study flood or flash flood events occurred in the area among Liguria and Piedmont, characterized by complex orography. The COSMO meteorological model at 2 km resolution was used to study such extreme events and different tests were performed in order to find the best configuration of the model. The impact of new physical parameterizations and different data assimilation methods on the forecast quality has been investigated using classic statistical indices and the innovative fuzzy verification. Simulations were also carried out with ICON model and compared with those of COSMO.

## Summary of problems encountered

Some problems were encountered in collecting and validating the observed data from the Italian meteorological network Dewetra, as there were some inconsistencies in the station database.

## Experience with the Special Project framework

My experience was very good, I did not encounter any problems regarding administrative aspects.

## Summary of results

1. The COSMO model was used at high resolution to study the November 1994 flood in Piemonte. The reforecast of the event was performed by COSMO at 2.2 km grid spacing (COSMO-2I) over Italy from 4$^{th}$ November up to +48h, nested in a domain over the Mediterranean area at 5 km grid spacing (COSMO-5M). The initial and boundary conditions for COSMO-5M were provided by a new set of experiments performed by IFS HRES at grid spacing of 9 km and driven by the new climate reanalysis dataset ERA5 HRES at 31 km horizontal resolution. We tested the 2 available parametrizations for the shallow convection, Tiedtke and Bechtold schemes, since the deep convection was resolved, and finally we tested the fully explicit convection. Some results are summarized in the paper "Reforecast of the November 1994 flood in Piedmont using ERA5 and COSMO model: an operational point of view".

2. Further simulations were performed by COSMO at 2.2 km resolution to study the VAIA storm on the 29-30 October 2018 and the extreme event among Liguria, Emilia-Romagna and Piemonte on 12-13 December 2017. Different model configurations have been evaluated to highlight a possible optimal configuration in reproducing the studied events. The impact on the model performance of new physics parameterizations recently developed in COSMO, i.e. the new "icon-like" physics, and hybrid initialization, which consists in the assimilation of observations from non-GTS stations or from radar to improve the determination of the initial conditions, were evaluated.
The different tested configurations are synthetized as follows:

a) OPE/RUN2: operational set-up and initial and boundary conditions (ICBC) provided by IFS at 9 km

b) ICON/CTRL: "icon-like" set-up and ICBC provided by IFS at 9 km

c) LHN/RUN4: "icon-like" set-up, ICBC provided by IFS at 9 km and latent heat nudging employed to assimilate precipitation rates observed by the Italian national radar composite up to +5h forecast

d) T2mN/RUN5: “icon-like” set-up, ICBC provided by IFS at 9 km and data assimilation by nudging for 2m temperature observed by non-GTS stations from Italian meteorological network up to +5h forecast

e) TR2mN/RUN6: “icon-like” set-up, ICBC provided by IFS at 9 km and data assimilation by nudging for 2m temperature and 2m relative humidity observed by non-GTS stations from Italian meteorological network up to +5h forecast

Different methods were used to assess which configuration is best in representing the events:

- the fuzzy technique (2D and 3D) to verify precipitation, by comparing the simulated maps with the observed data provided by the national radar mosaic and corrected by rain gauges
- the classic statistical indices (Mean Bias, Root Mean Square Error, correlation, etc.) to evaluate temperature and humidity at 2m and wind speed at 10m by comparing the simulation with the data provided by the meteorological stations of the Italian network.

December 11-12, 2017

The results of the 2D fuzzy verification for the various studied configurations are shown in Figure 1, while those concerning the 3D fuzzy are shown in Figure 2. Both the results are generally good and there are not significant differences between 2D and 3D, although of course 3D has slightly better scores: this means that the event is rather well reproduced, both in space and in time, by all the configurations. The impact of turbulence parameterization on precipitation pattern prediction is shown by comparing RUN2/OPE, a configuration implementing the configuration currently operational in COSMO-2I and COSMO-IT, with CTRL/ICON, a configuration implementing the new 'icon-like' physics introduced in COSMO. The results show that for this case study CTRL has significantly better scores for medium to high thresholds, reaching useful scale at 20 km for the 40 mm/3h threshold, while no useful scale is reached for 40 mm/3h by RUN2. Further improvement is observed from fuzzy 3D, where CTRL reaches the useful scale for 40 mm/3h as early as 11 km, while RUN2 still does not reach any useful scale demonstrating its difficulty in predicting, both spatially and temporally, the most intense peaks.

The impact of hybrid initialization on precipitation prediction is noted by comparing CTRL/ICON, a configuration with no assimilation, with RUN4/LHN, which assimilates latent heat data from radar for the first 4 hours of assimilation, RUN5/T2mN, which assimilates temperature data at 2m provided by non-GTS ground stations of the Dewetra national network, and RUN6/TR2mN, which assimilates both temperature and humidity data at 2m provided by non-GTS ground stations of the Dewetra national network. The 2D fuzzy shows no significant difference between the configurations, all of which are characterized by excellent scores. The 3D fuzzy further improves the scores, with useful scales achieved for all configurations at low to medium thresholds. The only configuration that reaches the useful scale at 6.6 km for the highest threshold of 40 mm/3h is RUN/LHN, which can therefore be elected as the best configuration.

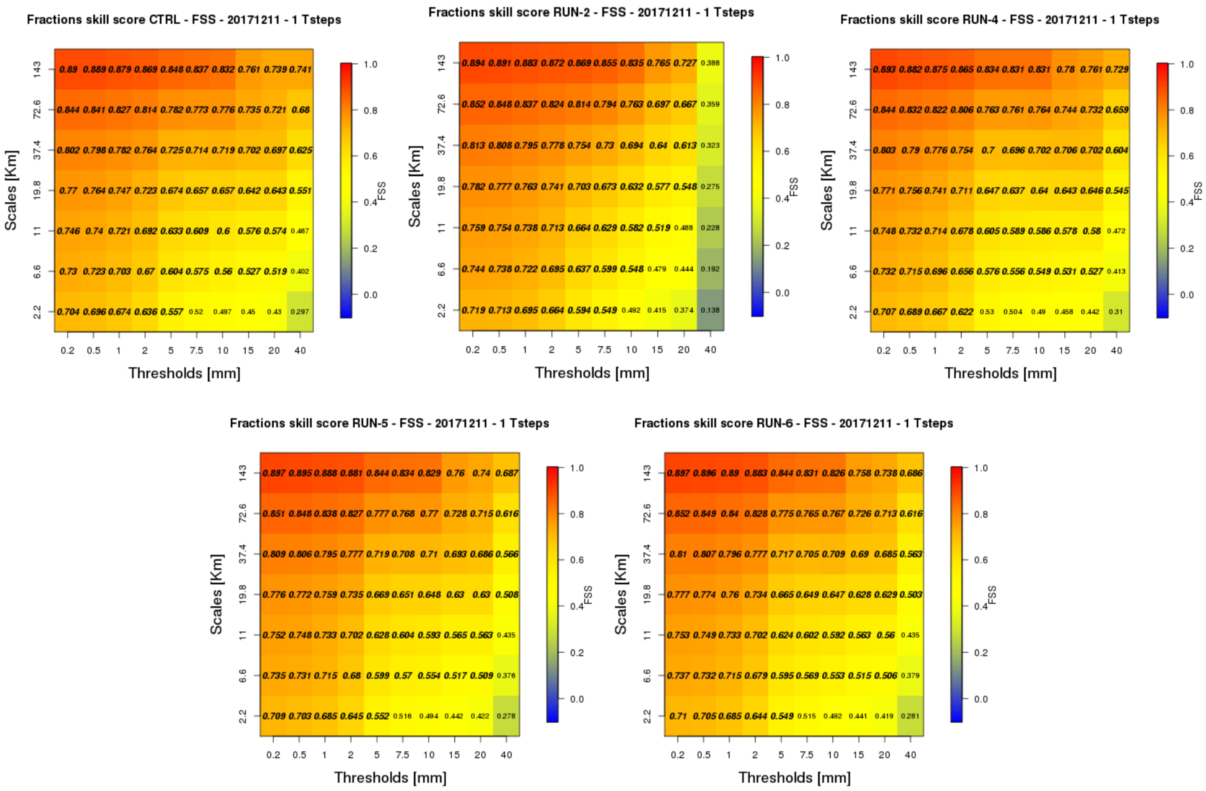


Fig. 1: FSS calculated for the first 24 hours of forecast (D0, 11/12/2017) – 2D fuzzy verification

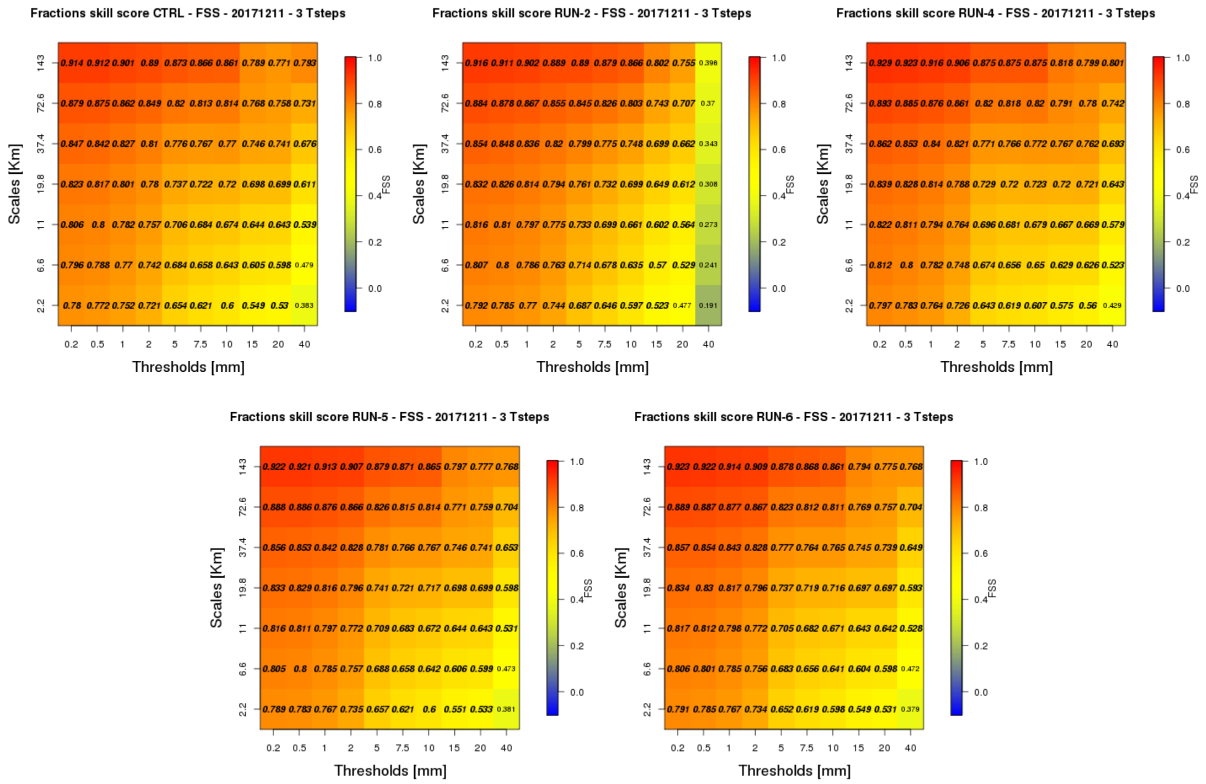


Fig. 2: FSS calculated for the first 24 hours of forecast (D0, 11/12/2017) – 3D fuzzy verification

The impact of new turbulence parameterization and hybrid initialization was also evaluated on the temperature and humidity variables at 2m in Figures 3-4. Regarding temperature, the hourly MB evolution shows a significant underestimation of the OPE operational configuration and an overestimation of the ICON 'icon-like' configuration. Latent heat assimilation significantly reduces the MB and RMSE, just as temperature assimilation improves the prediction up to 12-15 hours lead time. The Taylor diagram shows that the OPE configuration is characterized by a lower centered RMSE and a standard deviation that is closer to the observed one. Regarding humidity a slight overestimation characterizes the OPE run, while a definite underestimation characterizes those performed with the new 'icon-like' physics. Latent heat assimilation does not seem to have a good impact, increasing MB and RMSE, while temperature and humidity assimilation (TR2mN) improves the scores up to 8-10 hours lead time. The Taylor diagram shows how the 'icon-like' configurations better reproduce the variability of the observations, the standard deviation being similar to that observed.

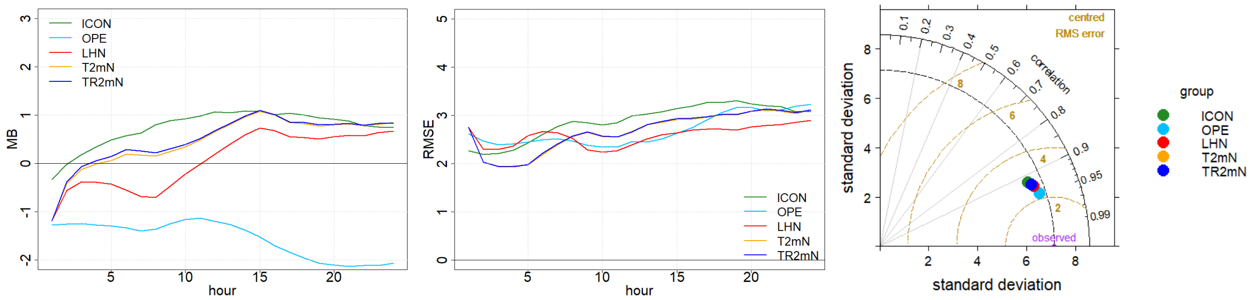


Fig. 3: Mean error (bias), Root Mean Square Error and Taylor Diagram for the T2m forecast (+24 hours)

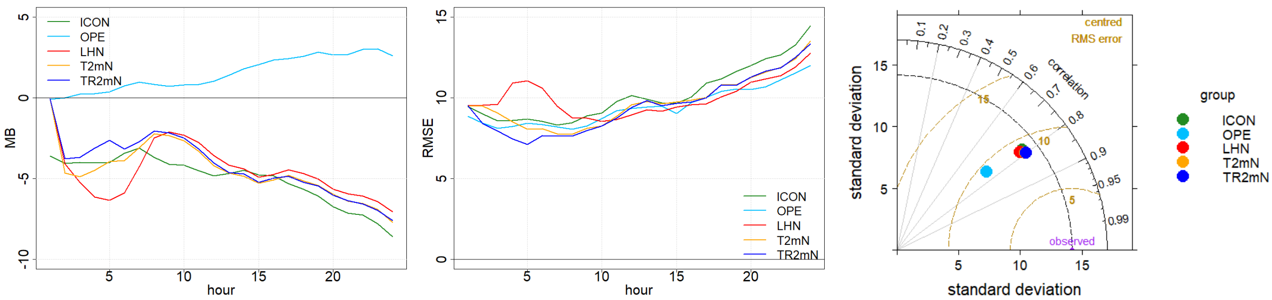


Fig. 4: Mean error (bias), Root Mean Square Error and Taylor Diagram for the RH2m forecast (+24 hours)

VAIA storm, October 29-30, 2018

The results of the 2D fuzzy verification for the various studied configurations are shown in Figure 5, while those concerning the 3D fuzzy are shown in Figure 6. The impact of turbulence parameterization on precipitation pattern prediction is shown by comparing RUN2/OPE, a configuration implementing the configuration currently operational in COSMO, with CTRL/ICON, a configuration implementing the new 'icon-like' physics introduced in COSMO. The fuzzy 2D results show that for this case study there are no significant differences between the two configurations, both of which are rather disappointing at all thresholds, particularly for mid to high thresholds: no useful scale is reached for thresholds greater than or equal to 20 mm/3h. However, it is noteworthy that the fact that precipitation affected a large part of Italy was penalizing for the medium-low thresholds, as it resulted in very high FSSuseful thresholds that were difficult to exceed. In absolute terms, in fact, FSS values for low thresholds are on the order of 0.7-0.8, a value that in case precipitation was not affecting large part of Italy would certainly have allowed FSSuseful to be exceeded and reach the useful scale. Fuzzy 3D shows a marked improvement in scores at all thresholds, although no useful scale is reached for the highest threshold of 40 mm/3h, confirming the difficulty of correctly predicting both in time and space the most intense peaks.

The impact of hybrid initialization on precipitation prediction is noted by comparing CTRL/ICON, a configuration with no assimilation, with RUN4/LHN, which assimilates latent heat data from radar for the first 4 h of assimilation, RUN5/T2mN, which assimilates temperature data at 2m provided by non-GTS ground stations of the Dewetra national network, and RUN6/TR2mN, which assimilates temperature and humidity data at 2m provided by non-GTS ground stations of the Dewetra national network. The 2D fuzzy designates the RUN4/LHN run as the best configuration, being the only one to reach the useful 20-km scale for low-to-medium average thresholds (2-7.5 mm/3h). The 3D fuzzy further improves the scores, with useful scales achieved for all low-medium thresholds for all configurations, demonstrating the difficulty of temporally localizing the event. The best configuration continues to be RUN/LHN, the only one to reach the useful scale at 20 km for 20 mm/3h thresholds; the useful scale for 40 mm/3h thresholds was not reached by any configuration.

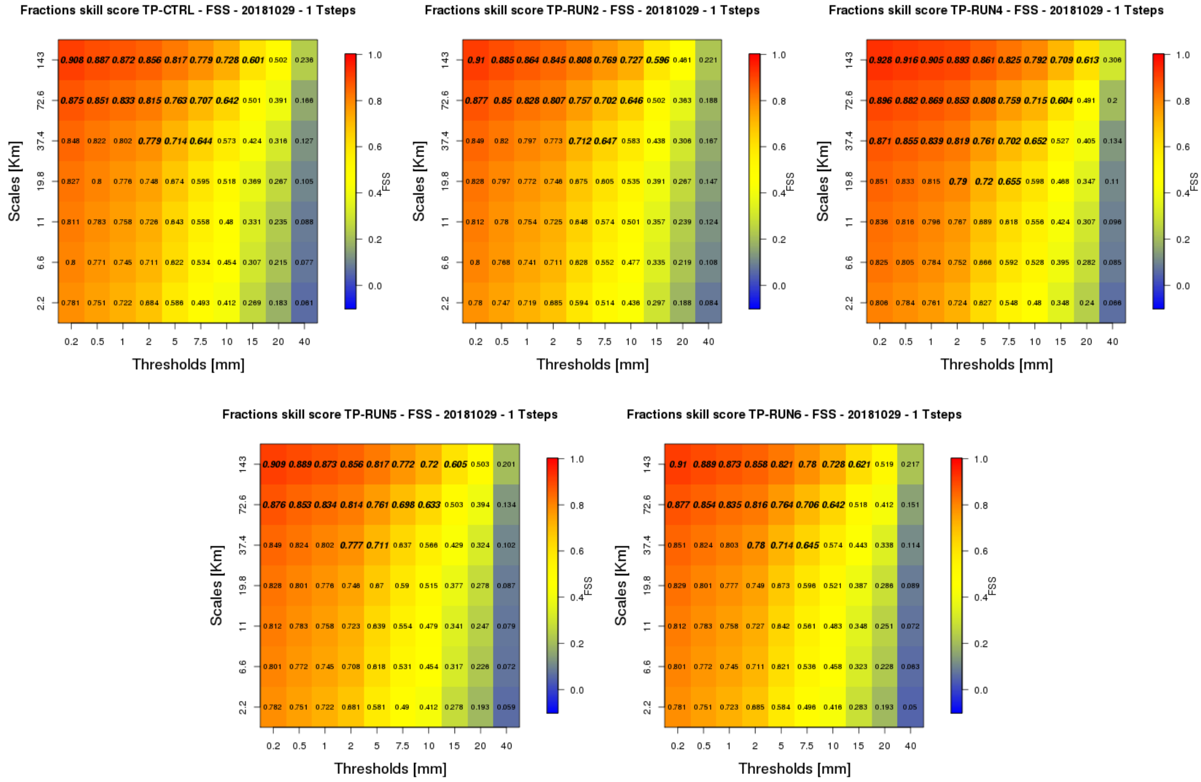


Fig. 5: FSS calculated for the first 24 hours of forecast (D0, 29/10/2018) – 2D fuzzy verification

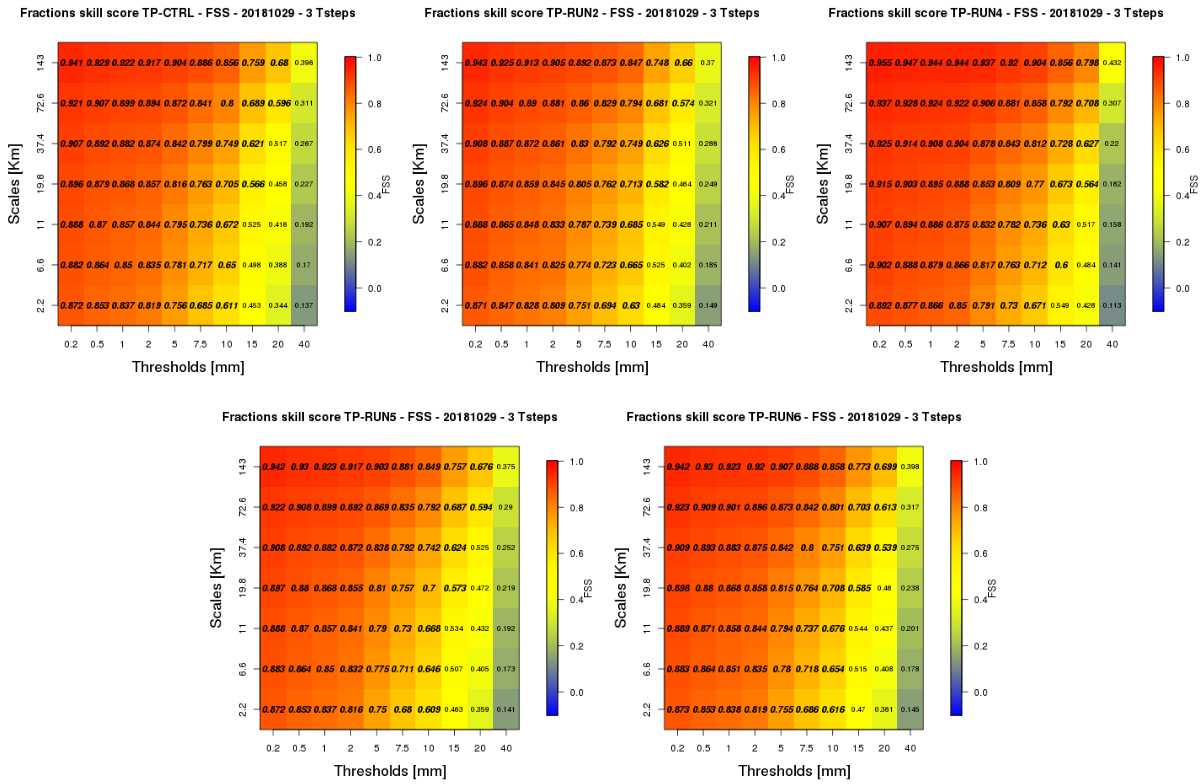


Fig. 6: FSS calculated for the first 24 hours of forecast (D0, 29/10/2018) – 3D fuzzy verification

The impact of the new turbulence parameterization and hybrid initialization was also evaluated on the variables of temperature and humidity at 2m and wind gust at 10m in Figures 7-9. For temperature, the hourly MB evolution shows an underestimation of the OPE operational configuration and a general overestimation of the 'icon-like' configurations for much of the lead-time. Latent heat assimilation significantly reduces the MB and RMSE, just as temperature assimilation slightly improves the prediction up to 12-15 h lead-time. The Taylor diagram shows no significant differences between the configurations. Regarding humidity a definite underestimation characterizes the runs performed with the new 'icon-like' physics and a significant deterioration is visible as lead-time increases for all configurations. Latent heat assimilation slightly improves the scores, but it is the assimilation of temperature and humidity (TR2mN) that produces the most significant improvement, effective up to 10-12 h lead-time. The Taylor diagram shows good model correlation and no significant differences between configurations. The impact of the various configurations on the wind gust prediction is analyzed by means of a 'conditional quantile plot,' which shows that all configurations largely overestimate the gusts, particularly in the higher values where all curves deviate from the 'perfect model.' The Taylor plot shows poor model correlation and highlights that the OPE run is characterized by a lower centered RMSE and a standard deviation closer to that observed than runs characterized by 'icon-like' physics.

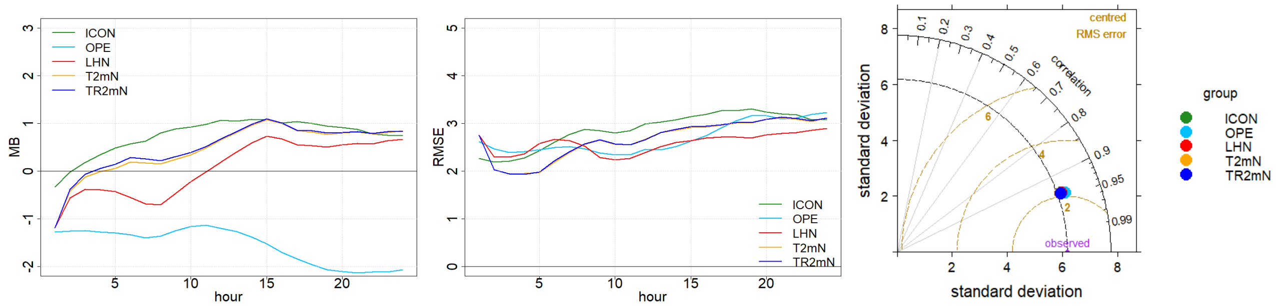

Fig. 7: Mean error (bias), Root Mean Square Error and Taylor Diagram for the T2m forecast (+24 hours)

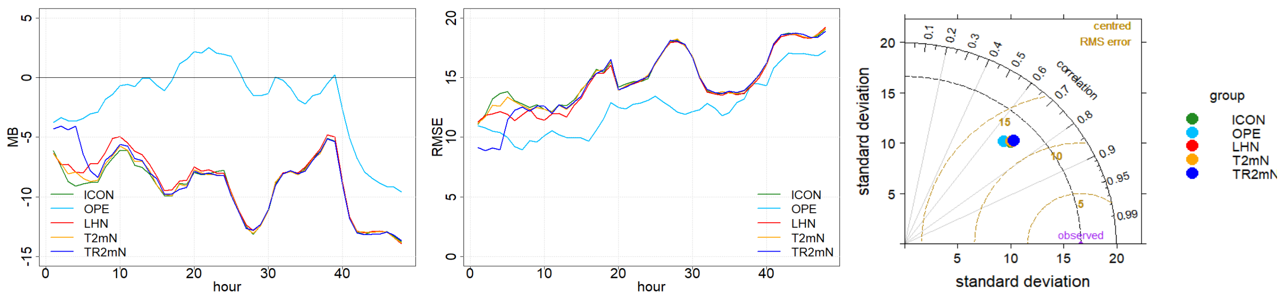

Fig. 9: Mean error (bias), Root Mean Square Error and Taylor Diagram for the RH2m forecast (+24 hours)

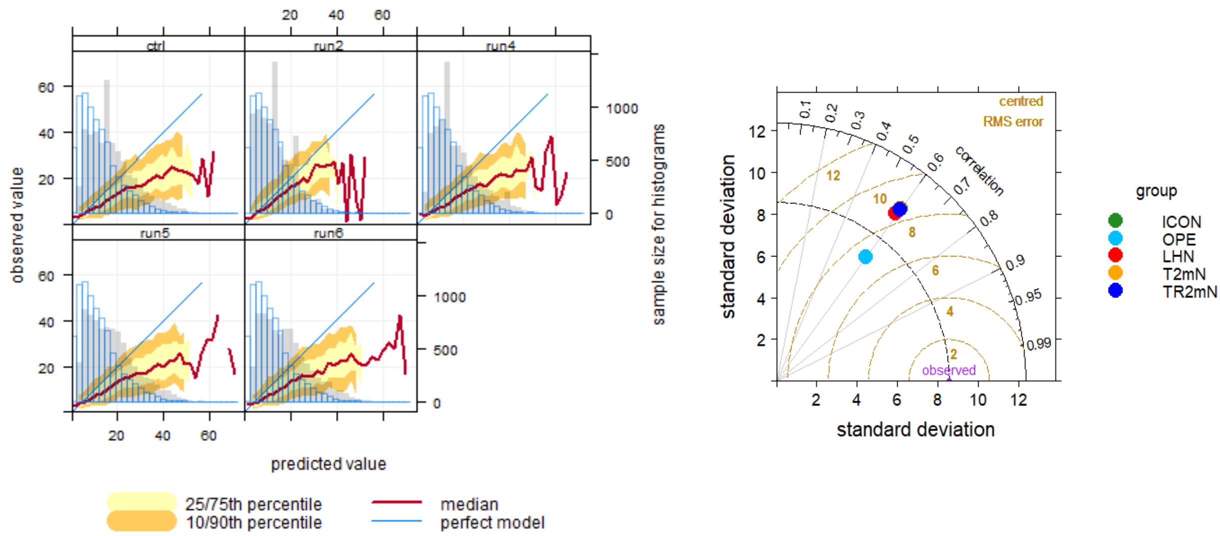

Fig. 9: Conditional quantile plot and Taylor Diagram for the 10 m wind gust forecast (+24 hours)

A preliminary study was conducted with ICON-LAM relative to the VAIA storm case-study so that the results could be compared with those of COSMO. Simulations at the resolution of 2.5 km were carried out on the Italian domain, nested onto a 5-km European domain whose initial and boundary conditions were provided by the ICON global model at 13 km every 3 hours, starting from October

29, 2018 at 00UTC until +48h. Verification of the ICON model and the comparison with COSMO model in his operational configuration (CTRL) were conducted with respect to wind gust and precipitation, the most significant variables of the event. 3D fuzzy verification in Figure 10 shows that the ICON model behaves worse for low to medium thresholds, as the $FSS_{useful}$ value at 2.2 km is reached for thresholds of 5 mm/3h by ICON and for thresholds of 10 mm/3h by COSMO. In contrast, the performance of the two models for high thresholds is poor and shows no significant difference, reaching no useful scale for the 40 mm/3h thresholds.
The 'conditional quantile' plot in Figure 11 shows how also ICON model largely overestimates wind gusts. The Taylor plot suggests a better performance of ICON, which shows a lower centered RMSE and a standard deviation that is closer to the observed standard deviation.

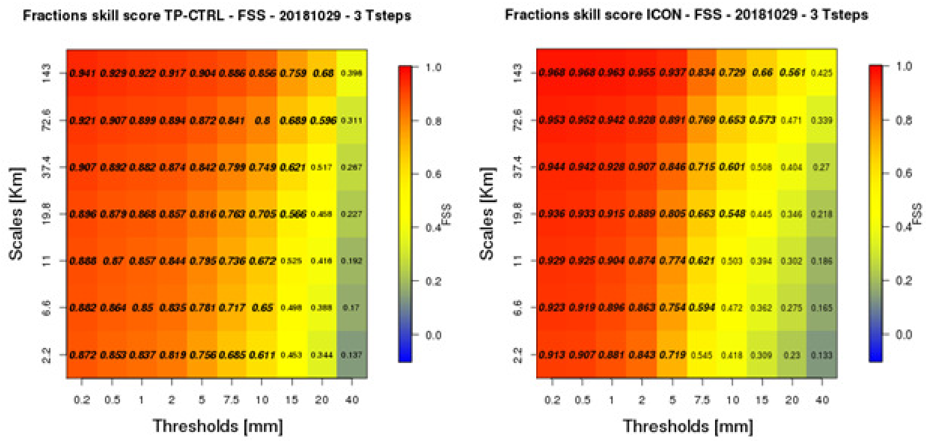


Fig. 10: FSS calculated for the first 24 hours forecast (D0) - 3D fuzzy: COSMO (ctrl) and ICON

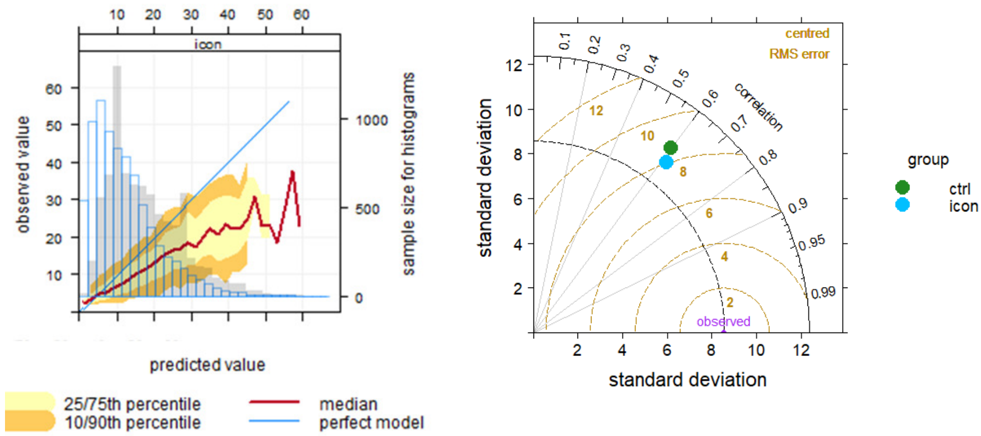


Fig. 11: conditional quantile plot for the 10 m wind gust ICON forecast (+24 hours) and Taylor Diagram for 10m wind gust comparing COSMO (CTRL) and ICON

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


V. Garbero, M. Milelli "Reforecast of the November 1994 flood in Piedmont using ERA5 and COSMO model: an operational point of view", Bulletin of Atmospheric Science and Technology volume 1, pages339–354 (2020)


## Future plans

These research activities will be continued in the Special Project 2022-2024 "Short-range forecast to investigate extreme weather events using COSMO and ICON models".