Assimilation of radar reflectivity volumes employing different observation error covariance matrices

Thomas Gastaldo[1,2], Virginia Poli[1], Chiara Marsigli[3], Pier Paolo Alberoni[1] and Tiziana Paccagnella[1] *[1] Arpae-SIMC Emilia-Romagna, Bologna, Italy*
[2] University of Bologna, Italy
[3] Deutscher Wetterdienst, Offenbach, Germany

1 Introduction

At Arpae-SIMC, the HydroMeteorological Service of Emilia-Romagna Region (Italy), the KENDA assimilation system [1] provides the analyses to the convection-permitting components of the operational modelling chain, consisting of one deterministic run and one ensemble system, both at 2.2 km horizontal resolution and with the same domain (greyscale in Figure 1). Currently, only conventional observations are assimilated, but tests are ongoing to include also reflectivity volumes [2] from the Italian radar network (solid lines in Figure 1).

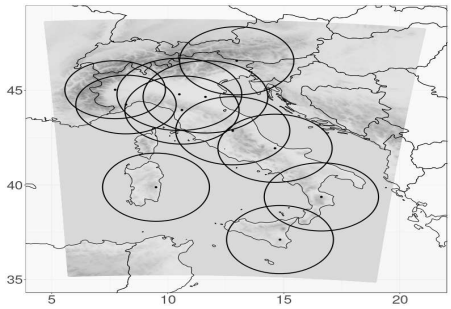

Figure 1: Integration domain (greyscale) of the COSMO model employed at Arpae-SIMC for high resolution model runs. The approximate coverage area for each radar at their lowest resolution of the Italian network is shown with solid lines.

The high spatial and temporal density of radar data demands a great care in setting the observation error covariance matrix $\mathbf{R}$. In fact, due to the great amount of data, small departures of the observation error from its actual value may lead to large errors in the analysis. Furthermore, reflectivity observations are spatially and temporally correlated and, therefore, the assumption made in most operational data assimilation systems of a diagonal $\mathbf{R}$ matrix is not realistic (see for example [2]).

The impact of using different estimations of the $\mathbf{R}$ matrix in the assimilation system is presented over two test periods. Results obtained when employing an unique observation error for all reflectivity volumes are compared to those obtained when a different value is specified for each observation, depending on the radar station and the distance from the station. The analyses, derived by each observation error matrix configuration, are used to initialize different forecasts. The comparison of the quantitative precipitation forecast (QPF) using the Fractions Skill Score (FSS [3]) allows to estimate the accuracy of the analysis itself. Finally, an estimation of spatial correlations between reflectivity observations is provided.

2 Observation error

The observation error ϵ_o has two components [4]: the measurement and the representation error. The former, also called instrument error, is the error associated with the measuring device alone, independently of how the measurements are used. The latter arises from 3 sources:

doi:10.5676/dwd_pub/nwv/cosmo-nl_19_03

- errors due to a mismatch between the scales represented in the observations and the model fields;
- errors introduced by the observation operator;
- errors due to pre-processing or quality control.

In data assimilation, an accurate estimation of the observation error is crucial since the observation error covariance $\mathbf{R} = E[\epsilon_o \epsilon_o^T]$ weights observations as $\mathbf{B} = E[\epsilon_b \epsilon_b^T]$ weights model background information (ϵ_b is the background error). While during the past decades a great effort has been done to improve the estimation of $\mathbf{B}$ (for example in the KENDA system it is fully flow dependent), small improvement have been done regarding the $\mathbf{R}$ matrix. In fact, $\mathbf{R}$ is fixed in time and generally assumed to be diagonal, that is observations are considered uncorrelated. Regarding the way to estimate it, one of the most used is the method proposed by Desroziers[5] which is based on the expect value of the product between observation-minus-analysis and observation-minus-background residuals.

3 Estimation of reflectivity errors

In order to estimate reflectivity error with a spatial dependence, we estimate the diagonal of $\mathbf{R}$ using Desroziers statistics and then we bin observations and the associated errors according to their horizontal and vertical distance from radar station. We use an horizontal step of 50 km and a vertical step 2 km. The estimation is performed for each radar of the Italian network over 3 periods,in order to have also a temporal dependence: from 31/08/18 at 00 UTC to 09/09/18 at 00 UTC (sept2018), from 30/09/18 at 15 UTC to 10/10/18 at 00 UTC (oct2018) and from 26/10/18 at 12 UTC to 11/11/18 at 00 UTC (nov2018).

Estimated values averaged over the three periods (sept2018, oct2018 and nov2018) and over all radars of the Italian network are shown in Figure 2. Values (y axis) are shown as a function of horizontal distance (x axis) and vertical distance (colours). As a general behaviour, we can notice that observation error increases with horizontal distance. This seems to be reasonable since the size of observed atmospheric volumes increases with the distance from the radar station. At the same time, we can notice that the observation error decreases with vertical distance up to the 4-6 km bin and then stabilizes. Also this behaviour seems to be reasonable since reflectivity observations close to the ground are more likely affected by non meteorological signals (i.e. clutter).

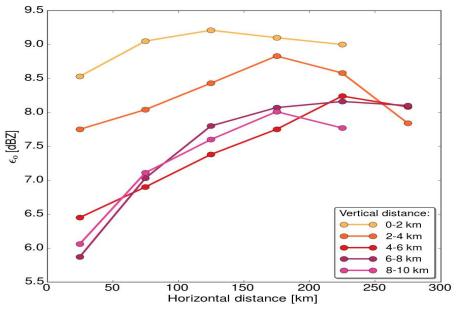

Figure 2: Estimated observation error for reflectivity volumes averaged over all periods and over all radars of the Italian network.

Due to the heterogeneity of our radar network and to the presence of different weather regimes in Italy, when the statistics is applied separately to each radar we can notice a certain variability. As an example, in Figure 3 estimated values of reflectivity errors are shown for Serano radar (left panel) in Central Italy and for Zoufplan radar (right) in North-Eastern Italy. Values are averaged over the 3 periods sept2018, oct2018 and nov2018. It can be noticed that the general behaviour described above is conserved but values and slopes of the curves vary quite significantly. A certain variability can be observed also when considering one radar but restricting the statistics to a single period. This is shown, for example, in Figure 4 for Zoufplan radar applying the Desroziers statistics only at the sept2018 period (left panel) and at nov2018 (right).

4 Use of estimated values of the observation error in KENDA

In order to evaluate the impact of using the estimated values of reflectivity error in the KENDA assimilation system, we perform 3 experiments. In *err_fix* experiment all reflectivity volumes have an error of 10 dBZ, as in our standard set-up for the assimilation of radar data. In *err_mean* experiment the observation error

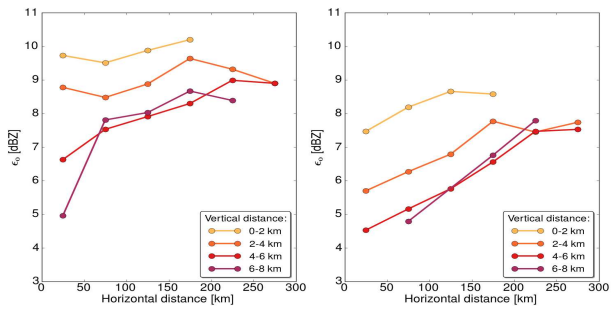

Figure 3: Same as Figure 2 but for computing the statistics only for Serano radar (left) in Central Italy and Zoufplan radar (right) in North-Eastern Italy.

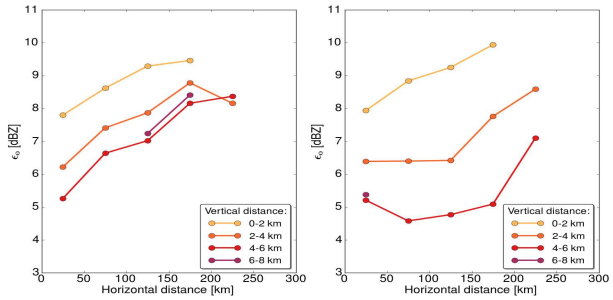

Figure 4: Same as Figure 2 but computing the statistics only for Zoufplan radar at two different periods: sept2018 (left) and nov2018 (right).

varies with radar station and with horizontal and vertical distance from station and it is averaged over all periods. Finally, in *err_period* experiment the observation error varies with radar station, with horizontal and vertical distance from station and with period.

The three experiments are performed for sept2018 and oct2018 periods. The KENDA system employs a 20 member ensemble plus a deterministic run and an assimilation window of 1 hour; Both conventional data and radar volumes (only the closest to analysis time for each radar) are assimilated. Finally, a deterministic forecast is initialized each 3 hours and forecast precipitation is verified by using the Fractions Skill Score (FSS). Regarding FSS, fixed spatial windows of 0.2 degrees are used and thresholds of 1 mm and 5 mm are considered. Observations are hourly rainfall fields from the Italian radar composite adjusted by rain-gauges.

Results are shown in Figure 5. Differences between the three experiments are small for both sept2018 (left panel) and oct2018 (right panel) periods. Regarding sept2018, FSS values for *err_mean* (red lines) are very close to those of *err_fix* (blue) for both the 1 mm (solid lines) and the 5 mm (dashed lines) threshold. In contrast, the performance of *err_period* (green) is generally slightly better than that of the other two experiments. However, when considering the oct2018 case, *err_mean* experiment is very slightly worse than *err_fix* and the worst results are obtained for *err_period*. In conclusion, due to the mixed results observed, we can state that the impact of employing a more accurate characterization of the observation error for reflectivity volumes in the assimilation system does not affect significantly the quality of forecast precipitation.

5 Estimated values of correlation between radar observations

Employing the Desroziers statistics, we also compute an estimation of spatial correlations for reflectivity errors. Similarly to the method described in Section 3, we bin pairs of radar observations according to their horizontal and vertical distance. We employ an horizontal step of 10 km and a vertical step of 1 km. Results obtained for the sept2018 case averaged over all radars of the Italian network are shown in Figure 6. As expected, errors are strongly correlated vertically and significant correlations can be seen up to 40 km in horizontal.

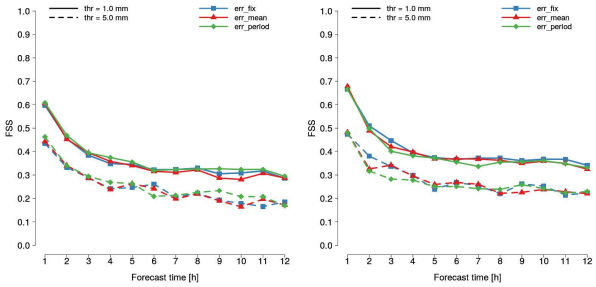

Figure 5: Fractions Skill Score for *err_fix* (blue lines), *err_mean* (red) and *err_period* (green) experiments employing a threshold of 1 mm (solid lines) and 5 mm (dashed lines). The verification is applied to sept2018 (left panel) and to oct2018 (right panel) periods.

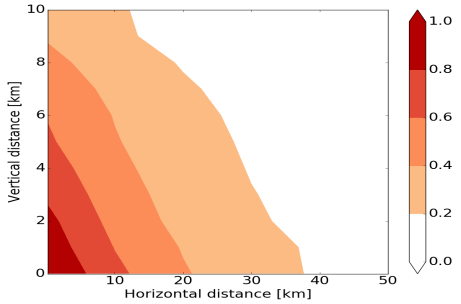

Figure 6: Spatial correlation between pair of reflectivity observations during sept2018 case.

6 Conclusions and future work

Even if reflectivity observation error varies quite significantly with time, radar station and distance from the radar, the use of more accurate values of errors in KENDA does not improve forecast accuracy. However, further tests are needed to confirm this result. The estimation of correlations between reflectivity errors shows that there is a strong correlation in space. Therefore, the exploitation of the correlation between pair of observations in the **R** matrix may be beneficial.

References

[1] Schraff, C., Reich, H., Rhodin, A., Schomburg, A., Stephan, K., Periáñez, A., and Potthast, R.: Kilometre-scale ensemble data assimilation for the COSMO model (KENDA), Q. J. Roy. Meteor. Soc., 142, 1453–1472, https://doi.org/10.1002/qj.2748, 2016.

[2] Gastaldo, T., Poli, V., Marsigli, C., Alberoni, P. P., and Paccagnella, T.: Data assimilation of radar reflectivity volumes in a LETKF scheme, Nonlin. Processes Geophys., 25, 747-764, https://doi.org/10.5194/npg-25-747-2018, 2018.

[3] Roberts, N.M. and H.W. Lean, 2008: Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events. Mon. Wea. Rev., 136, 78–97, https://doi.org/10.1175/2007MWR2123.1

[4] Janjić, T, Bormann, N, Bocquet, M, Carton, JA, Cohn, SE, Dance, SL, Losa, SN, Nichols, NK, Potthast, R, Waller, JA, Weston, P. On the representation error in data assimilation, *Q J R Meteorol Soc.* 2018; 144: 1257– 1278. https://doi.org/10.1002/qj.3130

[5] Desroziers, G. , Berre, L. , Chapnik, B. and Poli, P. (2005), Diagnosis of observation, background and analysiserror statistics in observation space. Q.J.R. Meteorol. Soc., 131: 3385-3396. doi:10.1256/qj.05.108