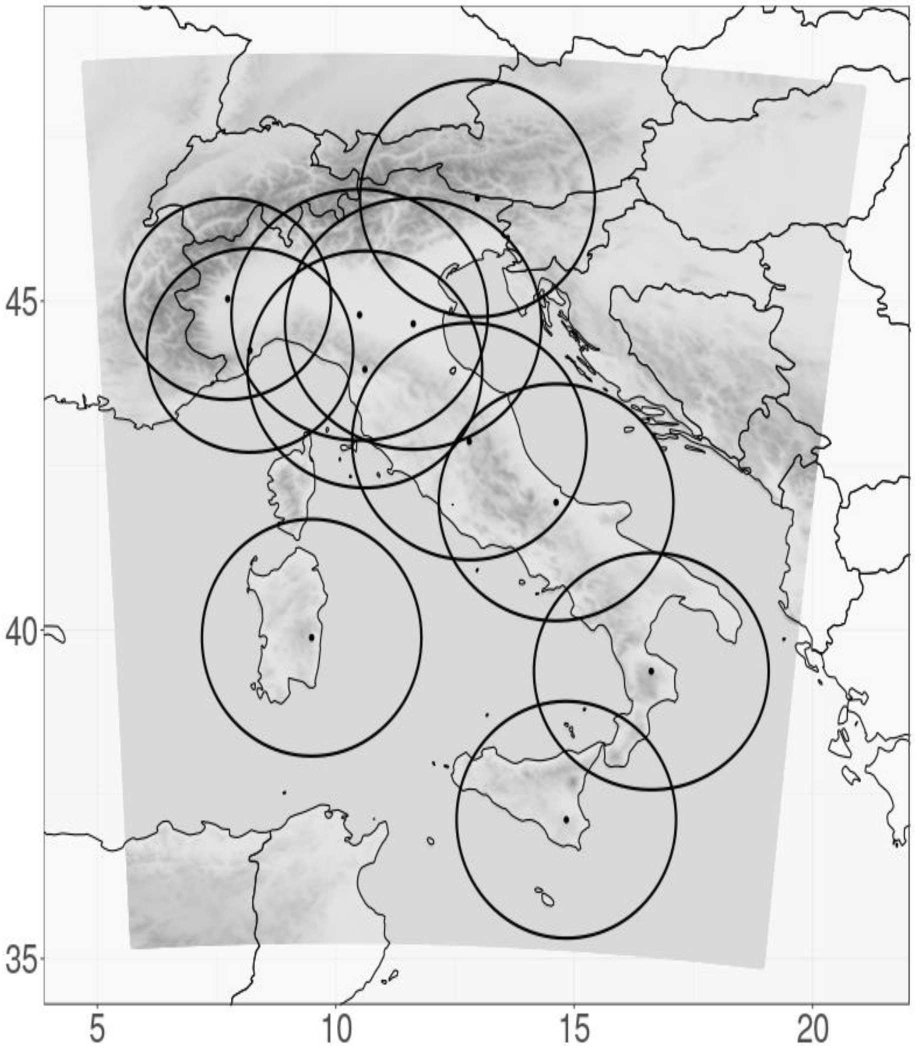
Assimilation of radar reflectivity volumes employing different observation error covariance matrices

Thomas Gastaldo[1,2], Virginia Poli[2], Chiara Marsigli[3], Pier Paolo Alberoni[2], Tiziana Paccagnella[2]

(1) University of Bologna
(2) Arpae Emilia-Romagna
(3) Deutscher Wetterdienst

Introduction



At Arpae, the **COSMO-2I** model provides operationally high resolution forecasts (2.2 km horizontal resolution) over Italy.

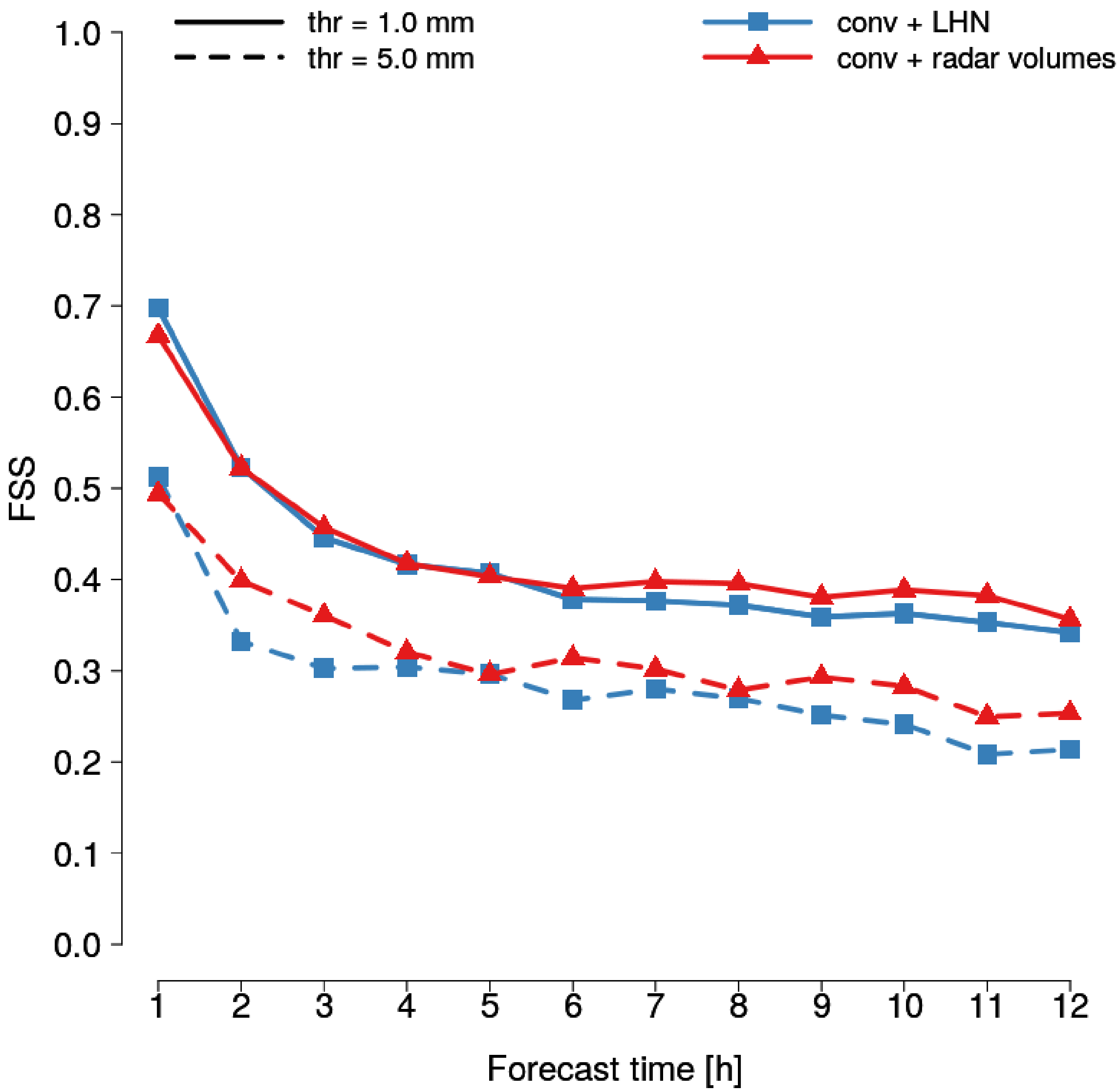
Initial conditions are generated by the **KENDA** data assimilation system, based on a LETKF scheme. At present, only conventional data are employed and LHN is performed during assimilation cycles.

Currently, tests are ongoing in order to replace LHN with the direct assimilation of radar reflectivity volumes from the Italian radar network.

Aim of the work

Several tests performed at Arpae showed that the assimilation of reflectivity volumes **improves forecast accuracy** compared to the use of LHN. However, the impact is not always very strong.

To further exploit radar data we decided to better characterize their **observation errors** and, as a future plan, to introduce spatial correlations in the KENDA system.



Observation error

Observational error ε_O has 2 components:

$$\varepsilon_O = \mathbf{y} - \mathbf{y}^* = \varepsilon_I + \varepsilon_R$$

ε_I = instrumental error

ε_R = representation error

In turns, the representation error arises from 3 sources:

$$\varepsilon_R = \varepsilon_U + \varepsilon_H + \varepsilon_P$$

ε_U = error due to unresolved scales and processes

ε_H = error introduced by the operator *H*

ε_P = error due to pre-processing

In data assimilation, the observation error covariance matrix $\mathbf{R} = E[\varepsilon_o \varepsilon_o^T]$ "weights" the observations, as $\mathbf{B} = E[\varepsilon_b \varepsilon_b^T]$ "weights" model background information (ε_b is the background error).

Estimation of R

In the LETKF scheme, $\mathbf{B}$ is estimated by using an ensemble and, therefore, it is flow-dependent. On the contrary, $\mathbf{R}$ is fixed in time and generally assumed to be diagonal. It should be bore in mind that:

- diagonal elements of $\mathbf{R}$ are the variances ε_O^2 for each observation,
- off-diagonal elements are the covariances between pair of observations.

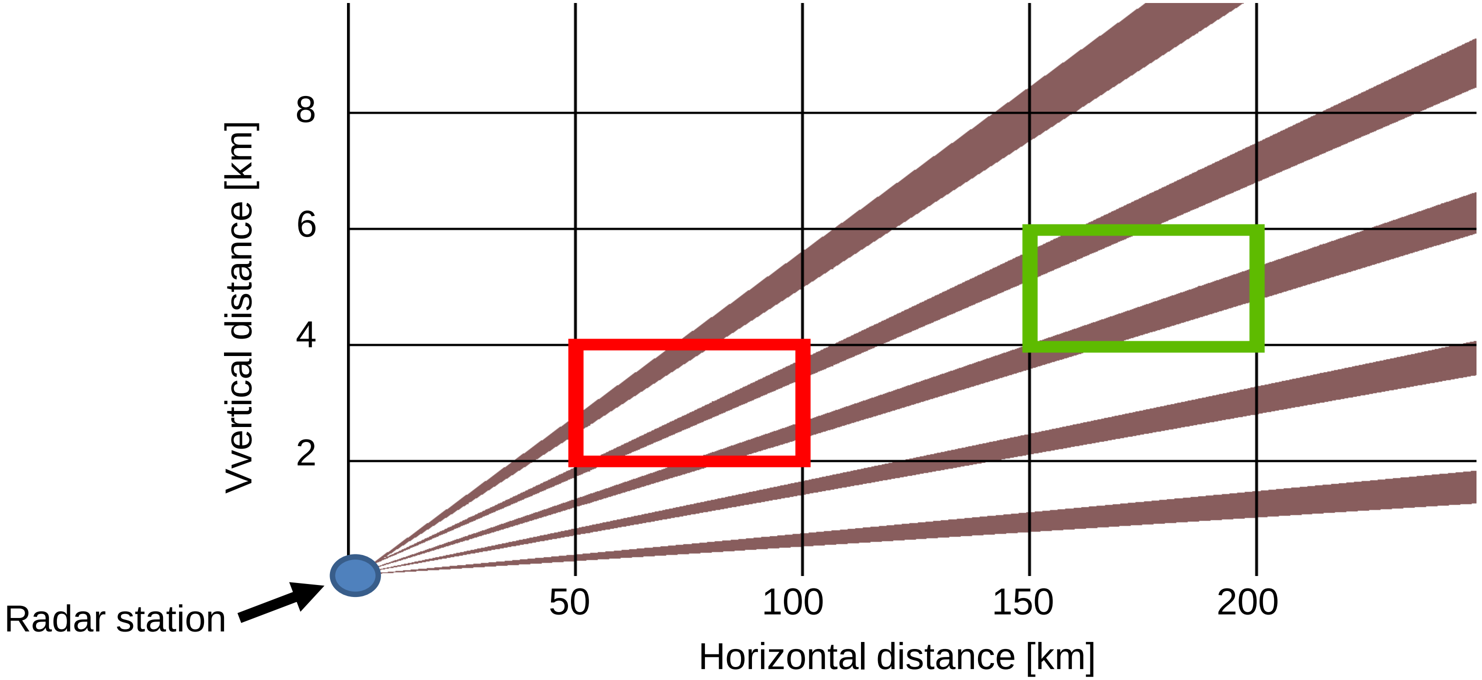
An estimation of $\mathbf{R}$ can be obtained by employing the statistics of Desroziers et al., 2005:

$$E[\boldsymbol{d}_a^o(\boldsymbol{d}_b^o)^{\mathrm{T}}] \cong \mathbf{R}$$

$$\boldsymbol{d}_a^o = \boldsymbol{y} - H(\boldsymbol{x_a}) \qquad \boldsymbol{d}_b^o = \boldsymbol{y} - H(\boldsymbol{x_b})$$

Estimate of ε_O with spatial/temporal dependence

In order to estimate reflectivity error ε_O with a **spatial dependence**, we estimate the diagonal of $\mathbf{R}$ using Desroziers statistics and then we **bin** observations according to their horizontal and vertical distance from radar station. Horizontal step is 50 km, vertical step 2 km.

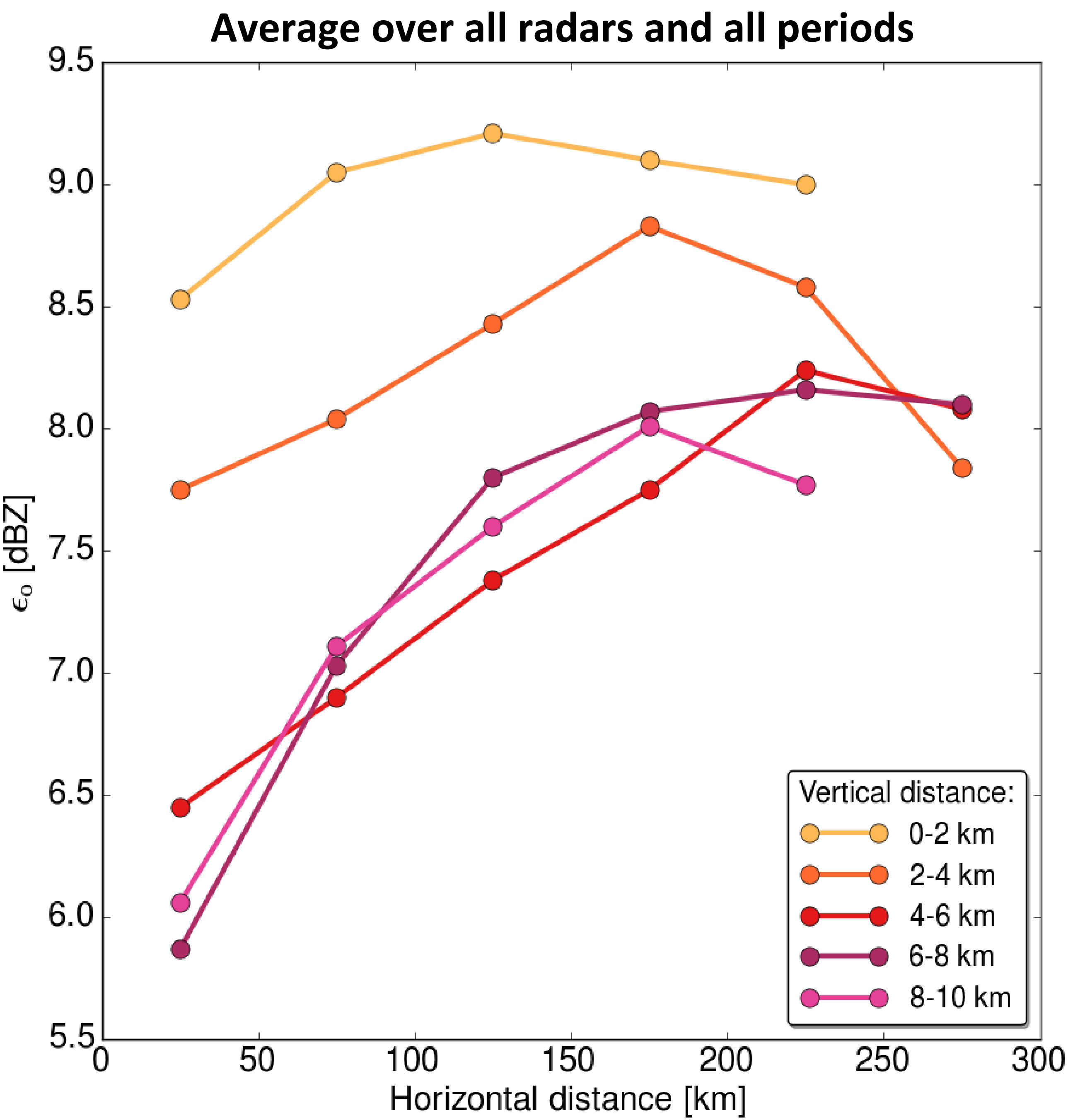


The estimation is performed for each radar of the Italian network over 3 periods, in order to have a **temporal dependence**:

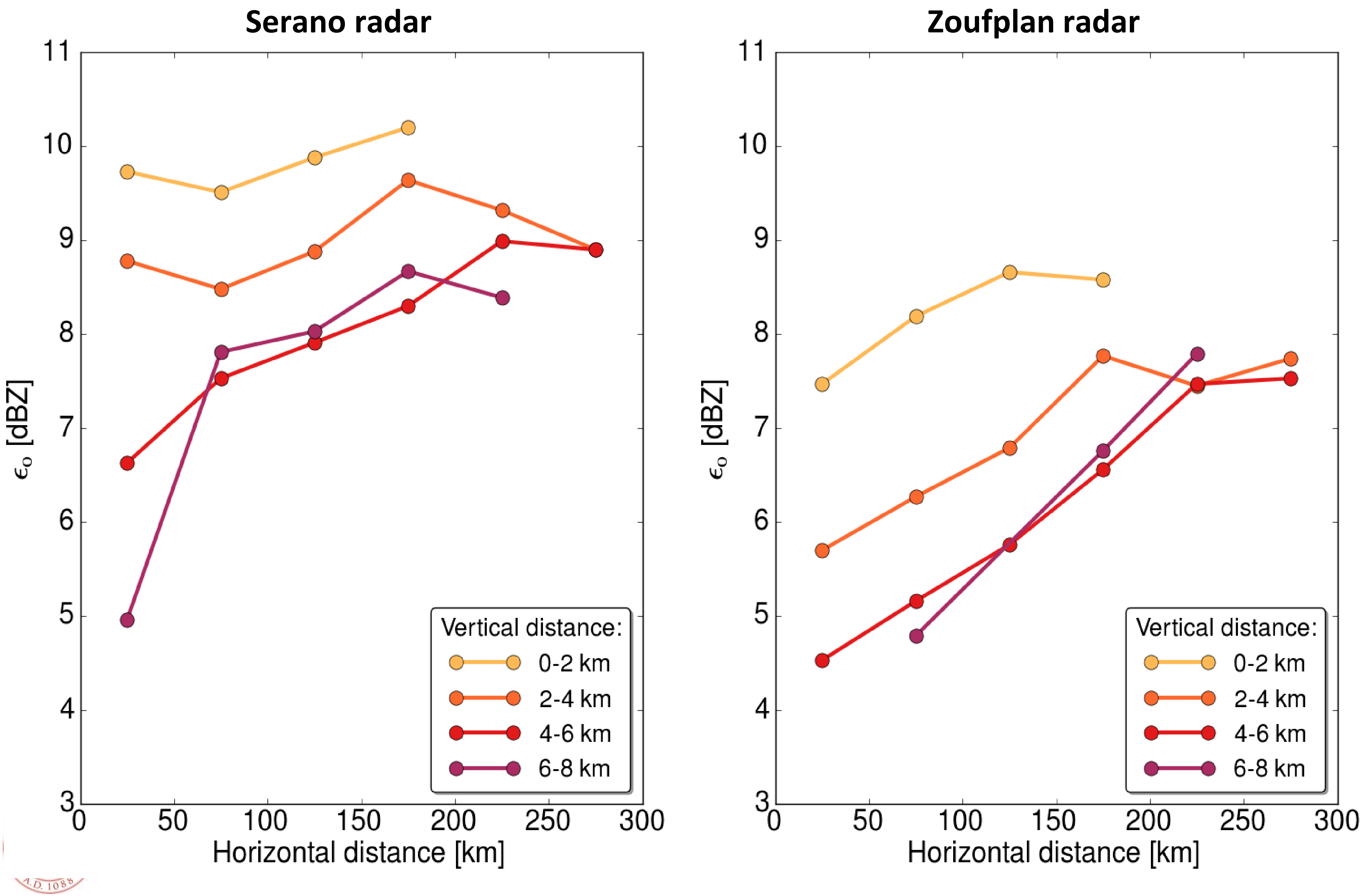
- From 31/08/18 at 00 UTC to 09/09/18 at 00 UTC (**sept2018**)
- From 30/09/18 at 15 UTC to 10/10/18 at 00 UTC (**oct2018**)
- From 26/10/18 at 12 UTC to 11/11/18 at 00 UTC (**nov2018**)

arpae emilia-romagna

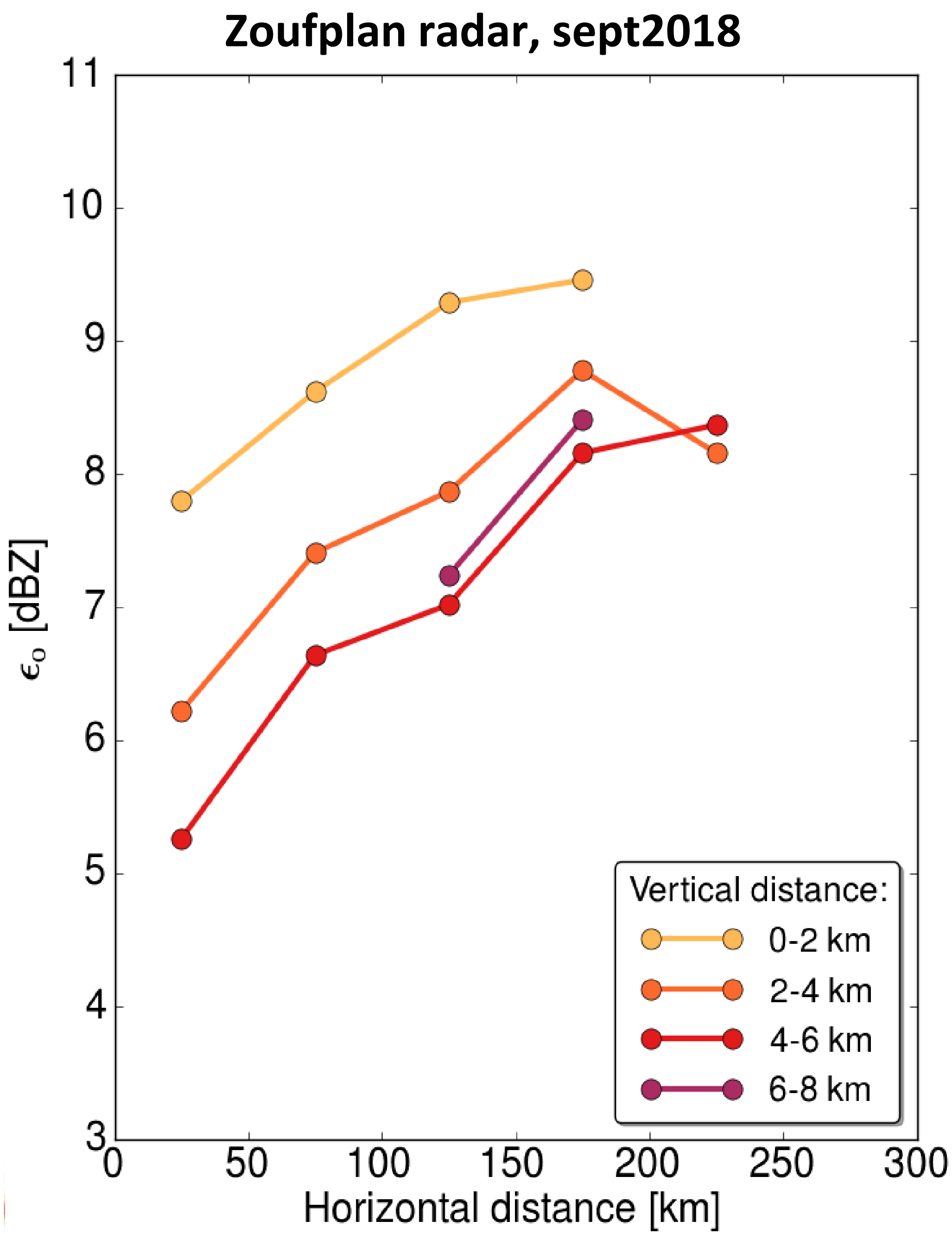
Estimate of ε_0 with spatial/temporal dep.: average

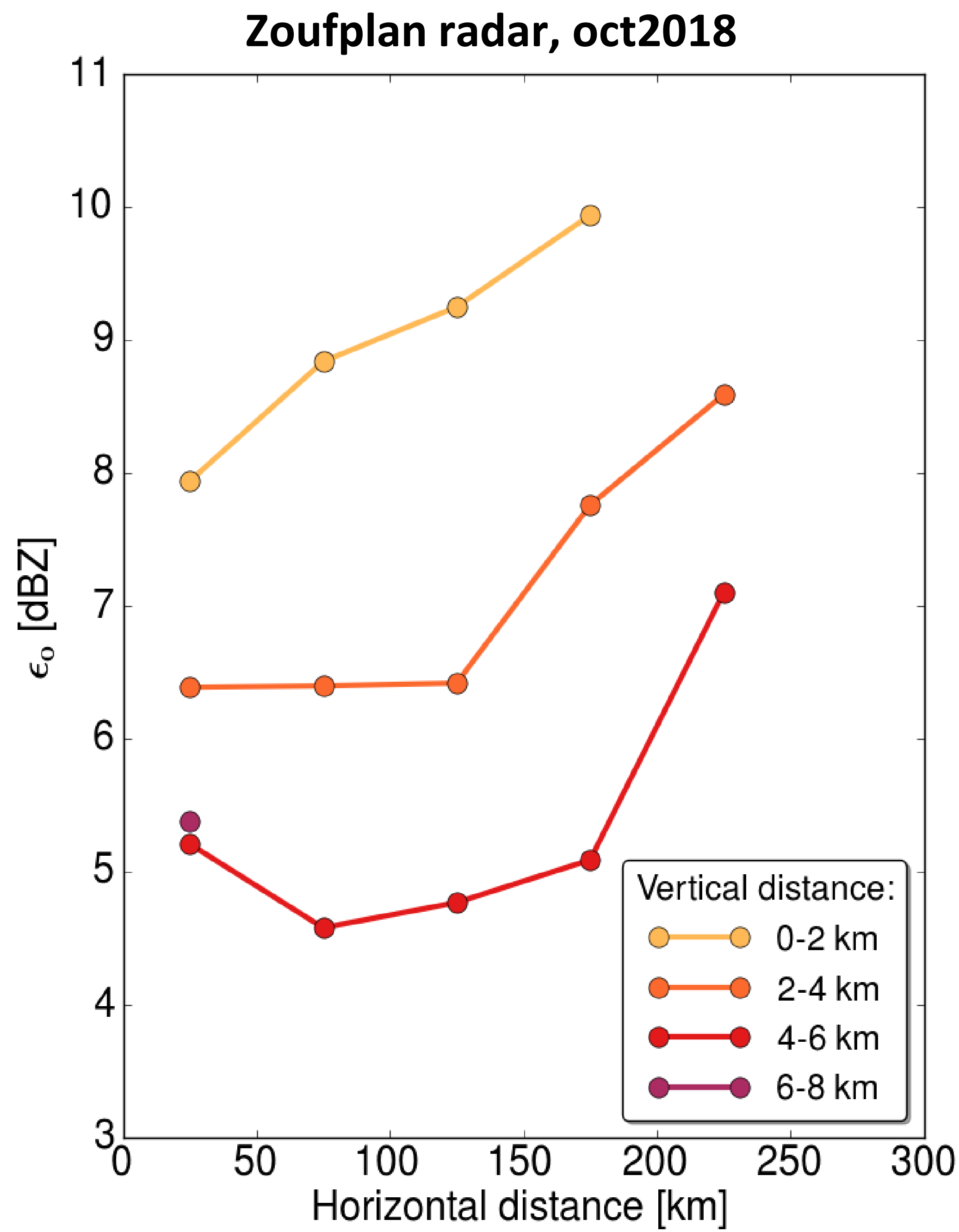


Estimate of ε_O with spatial/temporal dep.: different radar



Estimate of ε_0 with spatial/temporal dep.: different period





Use of estimated values of ε_O in KENDA

For sept2018 and oct2018 periods, 3 different experiments were run:

- **err_fix**: all reflectivity volumes have an error of 10 dBZ;
- **err_mean**: error ε_O varies with radar station and with horizontal and vertical distance from station, average over all periods
- **err_period**: error ε_O varies with radar station, with horizontal and vertical distance from station and with period.

General **set-up** of all these experiments:

- KENDA employs a 20 member ensemble plus a deterministic run and an assimilation window of 1 hour;
- assimilation of conventional data and radar volumes (only the closest to analysis time for each radar);
- a deterministic forecast is initialized each 3 hours and forecast precipitation is verified by using the Fractions Skill Score (**FSS**).

Verification with FSS

The Fractions Skill Score is a spatial verification measure which compares the forecast and observed fractional coverage of grid-box events in spatial windows of increasing size.

$$FSS = 1 - \frac{\frac{1}{N}\sum_N (P_f - P_o)^2}{\frac{1}{N}\left(\sum_N P_f^2 + \sum_N P_o^2\right)}$$

P_f = forecast fraction

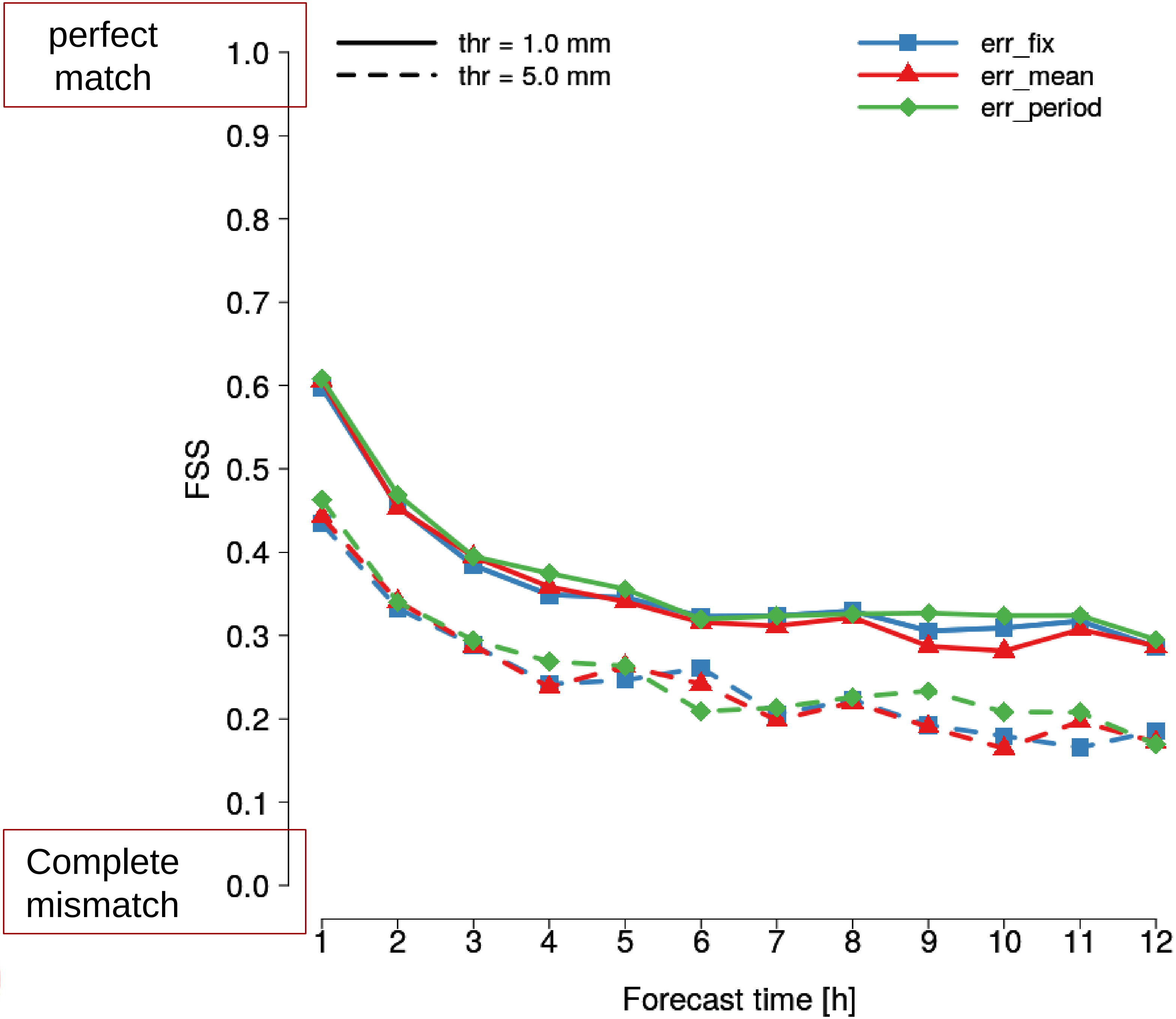
P_o = observed fraction

N = no. of spatial windows in the domain

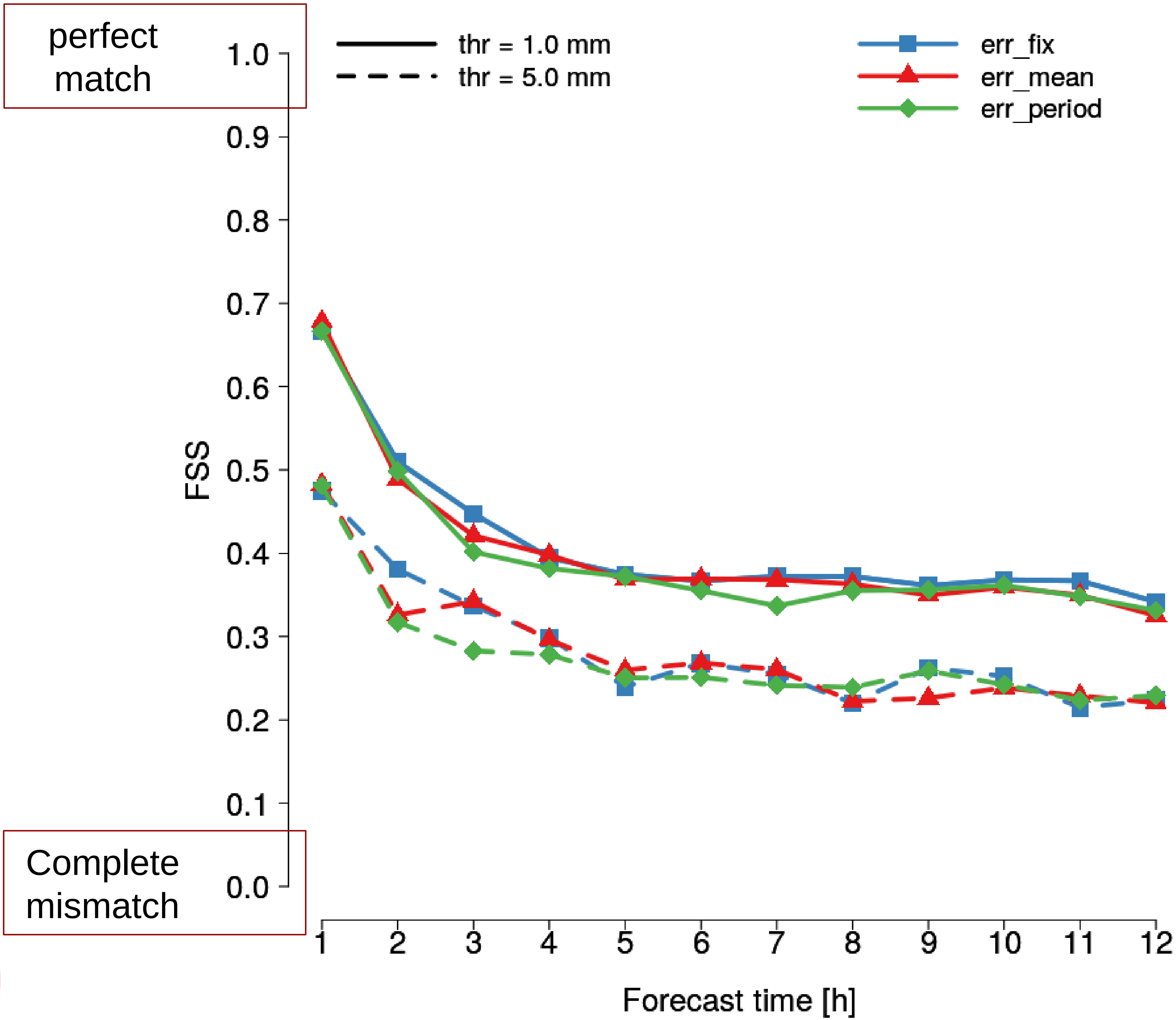
Specifics of our implementation:

- Domain covered with boxes: 0.2 X 0.2 degrees
- Verification of hourly precipitations
- Verification of all the forecasts at the same forecast time
- Observations are hourly rainfall fields from the Italian radar composite adjusted by rain-gauges
- Events were defined by different precipitation thresholds

Results: sept2018



Results: oct2018

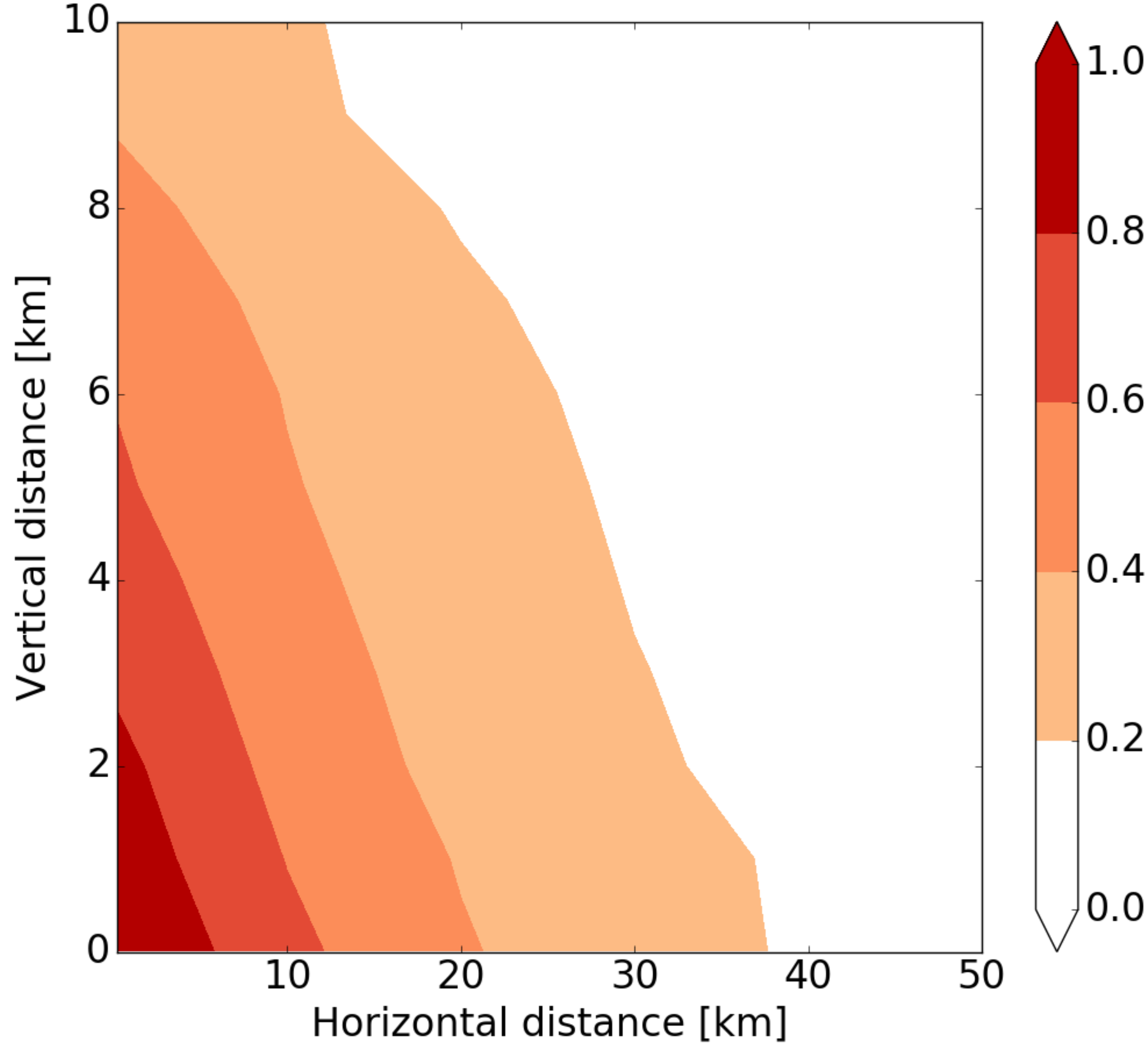


Correlation between radar observations

Pair of radar observations are binned according to their horizontal and vertical distance:

- Horizontal step = 10 km;
- Vertical step = 1 km.

The estimation is performed over the sept2018 period.



Conclusions and future plans

Even if reflectivity observation error ε_O varies quite significantly with time, radar station and distance from the radar, the use of more accurate values of ε_O in KENDA does not improve forecast accuracy. However, further tests are needed to confirm this result.

As expected, radar data are strongly correlated in space. The exploitation of the correlation between pair of observations in the $\mathbf{R}$ matrix may be beneficial. To implement the use of spatial correlations some modifications are needed in KENDA. This will be tested in the next months.

Thank you!