Review of "**High-Resolution Data Assimilation for Two Maritime Extreme Weather Events: A comparison between 3DVar and EnKF**"

by Diego Saúl Carrió, Vincenzo Mazzarella, and Rossella Ferretti.

The authors investigate potential methods for increasing forecasts and forecast lead times of mesoscale cyclones originating over the Mediterranean sea. Specifically, they compare 3D-Var and EnKF, and evaluate the added value of radar reflectivity and atmospheric motion vectors.

I am pleased with the revisions, which fully addressed the points raised and makes the outcomes of the study clear. Now the manuscript is in good shape and I recommend minor textual revisions, with no need for another review round.

Minor comments

It could be noted that some predictability was already present in the noDA experiment since ECMWF performed assimilation at a somewhat larger spatial scale?

L143: Wording: "solve" should probably mean "resolve"?

L148: Rephrase "In this context, which DA method is more suitable?".
For example: "Given limited computational resources, it is unclear which DA method is more accurate." However, as you know, the answer might depend on how big the resources are.

L427-428: "The assimilation of each observation results in a reduction of the ensemble spread, attributed to using a reduced-moderate ensemble size"
Confusing. Assimilating an observation reduces the analysis variance in variational and Kalman filter assimilation methods. If you want to motivate adaptive inflation, you could say that the small/finite ensemble sizes shrink the ensemble spread more than it should.

L429ff: Which technique was applied: spatially varying or homogeneous? According to the citation it was spatially varying?

Fig 8(i): Data for 3DVar CNTRL is missing.
Lower panels: RMSE unit is mm/h?

Fig 13: Caption: "Probability of cyclone center occurrence" add for example "(within 20 km)"
Check if the values on the colorbars are correct: are all values below 1% or is 0.16 actually 16%?
Tick values should appear once on the colorbar. I guess there is a rounding in place.

L980-981: Ensemble members are deterministic forecasts, right? If so, replace "deterministic numerical weather models" by "the NODA forecast".

L1060: "much less" ... can you quantify that approximately?

L1062-1034: "it does not need either to simulate model trajectories between the assimilation of a set of observations at time t1 and the subsequent set of observations valid at t2"
Confusing. Do you mean that 3D-Var simulates one model trajectory, while the EnKF needed to simulate 36 trajectories?

L900-903: "very different" feels vague. Better to be specific.
"some members could completely fail in the prediction of the weather event": Does it mean that some members did not predict the existence of a cyclone but just unorganized convection?

L903-904: " In this situation, our small-to-moderate ensemble will probably produce a poor flow-dependent background error covariance matrix": If it is somewhat probable that there is no cyclone, then this information should be in the background error covariance, I would say. However, you might mean that large uncertainty/spread leads to substantial nonlinearity, which is detrimental to the analysis accuracy of the EnKF.

L905-907: " for which the ensemble mean will be smoothed out significantly". The ensemble mean should not be expected to be a good forecast of the true state, in case the distribution is non-Gaussian, which it will be for extreme precipitation and the cyclone's pressure field.

L909-911: it is important to note that **although** the ensemble mean of the EnKF_SYN is not correctly reproducing the intensification of Qendresa, some of the ensemble members very well reproduce the observed MSLP both in deepening and timing"
In the light of the above comment, it should not be unexpected that averaging removes extreme values.

L1069: "the 3DVar performs better than the EnKF ensemble mean"
Yes, but then again, this is likely for extreme events because the ensemble mean is averaging over the skewed probability density function. I suggest rephrasing, since obviously in probabilistic metrics, like ROC/AUC, the situation is reversed.