EGUsphere, referee comment RC1
https://doi.org/10.5194/egusphere-2022-646-RC1, 2022

Comment on egusphere-2022-646

Anonymous Referee #1

Referee comment on "Combining short-range dispersion simulations with fine-scale meteorological ensembles: probabilistic indicators and evaluation during a ^{85}Kr field campaign" by Youness El-Ouartassy et al., EGUsphere, https://doi.org/10.5194/egusphere-2022-646-RC1, 2022

General Comments

The authors demonstrate the value of ensemble meteorology by showing how it can be used to model the uncertainty in the dispersion of material released from a known source. Where previous studies have focussed on long-range dispersion and used meteorology from global NWP models this study examines the use of a high-resolution (2.5km) ensemble NWP model covering a limited area to provide meteorological input to a dispersion model. In addition, the study focuses on modelling the dispersion of material that is regularly discharged from a reprocessing plant and compares the model results to observations over a period of two months. This field campaign along with the meteorological model and the dispersion model are clearly described within the paper.

My main complaint about this paper is that, for me, it covers too many topics. This has two impacts, first I am distracted from the main results and second the secondary topics are not covered in great detail so I am left with too many questions, as can be seen by the length of the specific comments section. There are two main sections which take my attention away from the main results. The first is the work looking at different methods to compute stability. The second is the consideration of how to model dispersion over time period which are longer than a single meteorological forecast. Both of these are interesting topics but, I feel they would be better placed in separate papers where they can be discussed more fully. I have included my specific comments on both of these sections in the specific comments.

Specific Comments

In the Introduction many papers are mentioned and in some cases the work carried out is described. It would also be helpful to understand the results or outcomes of the work in those papers. For example, the authors note that evaluations of dispersion ensembles

were performed by Le et al, (2021) and De Meutter and Delcloo (2022) but they don't say whether the ensembles were found to perform well or whether the use of ensembles provided more information. Similarly the authors mention the works of Galmarini et al,. (2004a and b) in performing multi-model ensembles but do not say anything about the findings of those works.

Line 50: I'm not sure "coarse" is the appropriate word to use here as one of the studies referenced in the previous paragraph used meteorological data at a resolution of 2.5km which is not generally considered to be a coarse resolution.

Line 82: Is it possible to define "reasonable" in reference to the ^{85}Kr release? Is the error on the release rate known?

Line 115: Is it possible to provide an approximate activity concentration for the amount of ^{85}Kr naturally present in the environment or a ratio of the ^{85}Kr present in the environment to the amount of ^{85}Kr released by the reprocessing?

Line 117: Similar to line 82; is it possible to define "reasonable" in reference to the ^{85}Kr release? Is the error on the release rate known?

Line 125: In describing the terrain around La Hague as complex is it possible to provide values for the maximum and minimum elevations to provide meteorological readers with a reference point for how the terrain might affect the wind speed and direction?

Line 136 and 137: For me the availability of data at a 10-minute resolution doesn't, on its own, constitute an accurate and reliable source term. I would be interested to know the uncertainty on the measurements relative to the amount of material released.

Table 2: Would it possible to add the temporal resolution of the met data to the table? I think this is mentioned later on in the text but it would be helpful to include it in this table too.

Section 3.1: There are a large number of different skill scores which could be used for the verification of both deterministic and ensemble predictions. Would it be possible for the authors to include an explanation of why bias and spread-skill were chosen?

Figure 3: In the text the authors mention that there is a diurnal cycle in the bias, but I find this difficult to see because the bias shares the same axis as the mean values. Would

it be possible to place the bias on a separate axis to the mean values?

Section 4.1.1: I am very surprised that it is necessary to use more than one 24-hour forecast for this study. The furthest observation point is situated <20km from the release location and assuming direct transport it would only take more than 24 hours to travel this distance if the mean winds for the whole 24 hours were less than 0.8m/s. In addition, 9 further hours of the first forecast were also still available so it would be possible to carry out a 36-hour forecast without needing to combine meteorological data from different days.

Line 319: I am curious to understand why the first 8 forecast hours were skipped? Is this a recommendation of the developers of AROME-EPS or is it due to the location of the release relative to the edge of the meteorological model domain?

Line 320: In table 2, the authors state that AROME-EPS is run four times a day, so I was wondering why model runs which are 24 hours apart are stitched together to build a continuous time series rather than model runs which are only 6 hours apart. My feeling is that using model runs which are 6 hours apart would reduce jumps at the forecast joins.

Line 333: Were the comparisons made in section 3 carried out using the unprocessed or processed meteorological fields?

Line 337 and 338: Are the authors able to comment on the impact of setting the minimum ABL height to 200m and/or provide evidence that this is a reasonable minimum ABL for the study area?

In Figure 7 and 8 I find it difficult to determine where the peaks in the ensemble are as the grey lines overlap a lot. Would it be possible to plot the ensemble as a shaded area rather than individual grey lines?

Line 362. The authors state that the use of a stack height of 100m does not allow them to accurately predict concentrations at 2km from the source in stable situations. Please could they expand on why stable conditions are problematic for the dispersion model they are using.

Line 365: What does the word “this” refer to in the sentence which begins in this line?

Line 375, figure 9 and table 4: Reading table 4 I think that peaks 2, 3 and 4 are much

smaller than peaks 1 and 5. For me it would be helpful for this to mentioned in the text.

Line 381: I think, in this sentence, the authors are arguing that the peaks are small because they are located close to the edge of the plume where the concentration gradients are high. It would be helpful to see a figure showing this. In addition, the authors appear to be suggesting that the solution to the underprediction is simply to increase the width of the plume which could be done by changing the stability category. Firstly, I would be interested to see why the authors believe that the inability of the model to predict the peaks is due to the stability and not to the wind speeds and directions along the path the puff has taken from the source location. Secondly, increasing the spread of the plume may help the model to capture the peaks where they are located at the edge of the plume, but this will be at the expense of the magnitude of the peaks where they are located at or close to the centre of the plume. Finally, given the emphasis placed on the stability within the second half of the paper I would be interested to see comment in the first half of the paper on the meteorological variables which impact on the calculation of the stability.

Line 403: Within the literature there are a number of different techniques proposed for the assessment of the performance of ensembles. Would it be possible for the authors to briefly explain why they selected the method of Querel, 2022 which is designed for the assessment of deterministic simulations?

Line 462, 463: Can I just check that the statement made on these two lines refers to the assessment carried out with parameters $\Delta T=3h$, $\tau=2h$?

Lines 523-525: Suggest removing this paragraph or expanding it substantially. Clustering has been tried with dispersion ensembles (Klonner, 2013) and was not found to be useful with the boundary layer.

Technical Corrections

Line 13: Replace “As first step” with “As a first step”

Line19: Replace “than deterministic one” with “than the deterministic one”

Line 60: “demonstrate” rather than “examine”?

Line 60: “skillfully” rather than “skillful”

Line 146: Suggest replacing “which means it does not generate chemical or physical reactions” with “it is not chemically or physically reactive”.

Line 166: Suggest adding “(in the absence of deposition)” between “shown that” and “3-D wind field”.

Line 281: The range -0.2 to 1.75 m/s doesn’t appear to match the range in Figure 3.

Line 285: +10 and -15 don’t appear to match the minimum and maximum values in Figure 3.

Figure 4: In the y-axis labels what do the “dd” and “ff” mean?

Table 3: Please could the authors separate the bias and spread-skill columns and place separate wind direction and wind speed labels above them as I’m not sure what each column represents.

Line 399: I think this reference should be Wilks, 2006 not Daniel and Wilks, 2006.

Line 406: Replace “because the” with “because there”

Figure 10: Would it be possible to explain the meaning of “dd” and “ff” in the figure caption?

Line 419: For clarity I suggest using the same language here as in the definition of “TN” in line 421

Line 592: As mentioned previously I thin this reference should be Wilks, 2006 not Daniel and Wilks, 2006.

Line 653: Leadbetter, Jones and Hort has now been published and can be found here: https://acp.copernicus.org/articles/22/577/2022/