Dear Dr. Brioude, referee #1, and referee #2,

We would like to thank you for the time and effort dedicated to providing feedback on our manuscript. We believed that the revised submission has undergone a considerable improvement in terms of explaining the study necessity and achievements thanks to the constructive comments and suggestions made by referees #1 and #2. All page numbers in our responses refer to the revised manuscript file without tracked changes. In the following, we list our responses to all comments and highlight the changes implemented in the revised manuscript.

Our responses to all comments made by referee #1 are as follows:

- I do not see any novelty neither in methods nor in results or discussion. I do not mean to degrade authors' great work, but I do not see this manuscript of relevance for publication in ACP.

  - ✓ **The novelty of this paper lies in the application of the FLEXPART Langrangian model for radionuclide dispersion simulations to inter-compare the variability and quantify uncertainties using a broad ensemble of (re-)analyses, as well as forecasted meteorological input datasets.**
  - ✓ **There are several works previously published in ACP using FLEXPART with a single meteorological dataset (Zhu et al., ACP 2020)(Sauvage et al., ACP 2017). In this study, we have designed a four-member ensemble that allows for the first time to inter-compare simulations produced by FLEXPART and FLEXPART coupled with the Weather Research and Forecasting model (FLEXPART-WRF) and capture the effect of downscaling on FLEXPART dispersion modelling. We evaluate the relative performance of forecast runs against three re-analysis runs (Table 1), both complementing and extending the approach (taking a reanalysis run as a reference run) by Leadbetter et al. (ACP, 2022).**
  - ✓ **Our paper also expands upon previous studies published in ACP that focus on a fictitious case to quantify the impact risks. For instance, the study by Salminen-Paatero et al. (ACP, 2020) uses dispersion modeling results with the SILAM model to study due to hypothetical reactor accidents in Finland. Our methodology of**

continuous release allows us to uniquely estimate the probability of occurrence over each hour-of-day and month-of-the-year.

- ✓ **In summary, our paper is not devoted to the model performance analysis in a specific real-world accident, but rather we aim to highlight and quantify the strong variability due to diurnal and seasonal meteorological variations stemming from the choice of re-analysis used for dispersion modelling. We capture the range of uncertainty through the iterative multi-day simulations, starting each day of the year, and the analysis of the resulting age spectrum of pollutants. Thus, our results will benefit the future development of early warning systems for both aerosol and gaseous pollutants and toxic substances that are subject to transport processes.**
- ✓ **Finally, we illustrate for the first time how the use of different meteorological inputs causes differences due to planetary boundary layer height (PBLH) representation in Langrangian models. We also demonstrate how the low spatial resolution of meteorological inputs causes the omission of the sea and land breeze circulation effects on the PBLH and, as a result, the variability in radioactive tracer concentrations.**

- The authors write "Using an ensemble of meteorological inputs, this study primarily aims to investigate the seasonal and diurnal changes in the transport and surface concentration and deposition magnitude of radionuclides in the event of a potentially possible nuclear accident". I am very sensitive with radiological issues and I think they should be handled very carefully, because then can have a negative pshycholgical impact to the public. What is "potentially possible nuclear accident" supposed to mean? There is no explanation that could justify this. Why did the authors study this particular hypothetical release? Why did they not study, for example, a hypothetical release from an older reactor? For instance, several Balkan reactors (which I do not want to name, but one can easily google) from the Soviet-era have shown functionality problems during the last 10 years and could affect a more significant area (central Europe) where a larger population lives and reproduces.

  - ✓ **While few new nuclear power plants are licensed in the Western world, and most Soviet-era stations are nearing the end-of-life decommissioning, several nuclear facilities are planned or proposed, and in the last few years are under construction or becoming operational in the Middle East/North Africa (MENA) region. The Barakah station is the latest NPP to become operational in a region**

with unique climatological conditions that were previously void of such developments and where the risk from radionuclide dispersion received little coverage in the literature, as opposed to Europe and the US.

- ✓ **In that regard, we also address what levels of radionuclide concentrations and deposition may affect the populated areas of Qatar in the event of a nuclear accident, a matter of significant concern due to the geopolitical situation in the region. The particular location has been selected because Barakah is the first nuclear power plant in the region, and additional ones have been planned.**
- ✓ **It is beyond the scope of this study to designate the causes or estimate the probability of a nuclear accident. Information about the risks of nuclear accidents is not shared by the industry and governments but needs to be taken seriously. We simulate a fictitious accident at the severity level of the Fukushima disaster but note that our results are indicative and can be scaled for any magnitude of emission from a small leak or release of radionuclides from an INES7 accident.**
- ✓ **The following explanation will be added to the revised manuscript: We simulate a fictitious accident at the severity level of the Fukushima disaster to compare the simulations with those produced for this accident.**

- Usually, for the assessment of transport of radionuclides and the impact of meteorological fields in transport modelling, more sophisticated state-of-the-art databases are used. I would encourage the authors to use the ETEX (Nodop, K., Connolly, R., and Girardi, F.: The field campaigns of the European Tracer Experiment (ETEX): Overview and results, Atmos. Environ., 32, 4095–4108, 1998) and ETEX-2 (https://doi.org/10.1016/j.atmosenv.2008.07.027) experiments and repeat their assessment rather that a hypothetical release that may never happen or cause the aforementioned problems (see previous comment).

  - ✓ **We would like to thank you for your comment and the proposed references. Indeed, there are several studies assessing Langrangian dispersion models using controlled release experiments. As stated above, our aim is different and not directly comparable with ETEX. Repeating ETEX would not provide any information about the transport of nuclear tracers or other toxic substances in the Middle East. The current study is a contribution to the establishment of an early**

warning system in Qatar, being the first country in the region that is planning such a system. We feel that it is important for scientists in our field (and ACP) to reach out to this region, and not only focus on Europe.

- ✓ **The analyses presented in this study are based on the median of numerous simulations (1460 simulation days or 35040 simulation hours at each point) to capture diurnal and seasonal variations throughout the year and uniquely capture the uncertainty from the input meteorology. Hence, we believe that our findings related to the seasonal and diurnal changes in the transport efficiency and the concentration and deposition of radionuclides and their spatial distribution are both timely for the region of interest and relevant for scientists and decision-makers for designing early warning systems and the preparedness for potential nuclear accidents.**

- - An alternative solution for publication might be to focus on the model developments they have done, correct the manuscript and submit to GMD. This would require a detailed validation of the results, which lacks here.

  - ✓ **Other than section 3.3 which concerns developments and the performance analysis of FLEXPART/FLEXPART-WRF, the major part of this study is devoted to the topics outlined above. The main focus is on the seasonal and diurnal changes in the transport and deposition of radionuclides to the region of interest (and in Qatar, subsection 3.1). We further analyze the temporal and spatial distribution of radioactive materials, the distribution of radionuclides in relation to the population density, the synoptic patterns leading to the transport of dense radioactive plumes, and the sensitivity to atmospheric turbulence. We feel that these topics are suitable and aimed toward the subject matter and audience of ACP rather than GMD.**

- In line 175, the authors are talking about a nuclear accident, but then release particles for only 24h? During the 2 worst nuclear accidents (Chernobyl and Fukushima), emissions lasted much longer, which makesa the study completely un realistic.

  - ✓ **Our study is not replicating previous accidents, rather we simulate emissions over 24-hours for each day over a full year period. This, along with aggregating statistically the median output, amounts to a continuous emission over a full year**

and allows us to gather meaningful representation of the seasonal and diurnal median changes in the distribution of radioactive materials basis (in total we emit over 365 days and simulate 1460 days). The diurnal variation in the radionuclide dispersion is also considered by stratifying the simulated concentrations (Figure 2) and deposition (Figure 3) corresponding to the hourly age of the lagrangian particles. We designed this analysis (along with those shown in Figures 4 and 5) to determine what time of the day and year is associated with the higher probability of the transport of dense radionuclide plumes from a hypothetical release in the Barakah nuclear plant to the study area. To the best of our knowledge, this is the first time that such a method is implemented, and we believe that it can be used to provide important information and guide the formulation of preparedness plans.

✓ **We note that in terms of emission magnitude, to translate our findings for a case in which an event with different intensities would occur, one can simply linearly scale the reported concentration/deposition risk. A realistic accident could be simulated, when it occurs, by applying our methodology in an early warning context and scaling the source strength based on available information about the accident. Similarly, by essentially simulating continuous release over a full calendar year, we can probabilistically capture the eventualities irrespective of the length of the release. Our methodology follows other studies that did not set out to determine the source term but to investigate the spatio-temporal distribution of pollutants (due to the effect of atmospheric/modeling conditions). For instance, Leadbetter et al. (2022) used a hypothetical release of 1 PBq Cs137 equivalent over 6 h at an elevation of 50 m.**

- Same paragraph later mentions that "... particles are initially distributed at height levels between 100 and 300 m above the ground level over the emission point". Since we have a nuclear accident and given our previous experience with nuclear accidents, one may expect emissions at higher altitudes (see paper from Stohl's group) depending of course if there was a thermal explosion (such as in Chernobyl) or a hydrogen explosion (such as in Fukushima). Hence, one understands that a sensitivity study is also required to examine what the impact of injection altitude would be on transport. I would expect large differences on transport between emissions that occurred at 300 m and at 3 km (such as those that were calculated for the 2 major nuclear accidents in 1986 and 2011).

- ✓ **In model sensitivity studies of the emission altitude (Evangeliou et al., ACP 2013; Table 1) we note that in the case of Chernobyl, other than the first few days when the graphite core was on fire (a deprecated design), the bulk of the emissions occurred at lower altitudes. Our study is indeed based on the paper by Stohl et al. (Figs. 4, 5) for the more recent and relevant example of Fukushima. We note that in that paper the inversion over three emission layers in altitude shows that for all practical purposes, the emissions were predominantly (almost 100% for Cs137) within the 0–50 m, and 50-300 m layers.**
- ✓ **Besides, carrying out sensitivity studies (in addition to what we have done for the turbulence schemes) causes a significant increase in the calculation load.**

- - Line 281: "Using conversion factors from Spiegelberg-Planer (2013), 131Iconc_seas_max (in a unit of Bq m-3) are converted to the maximum hourly doses from inhalation (in a unit of µSv)". This is not a proper dose-rate calculation. I would encourage the authors to calculate inhalation doses using the models presented in the WHO report for Fukushima that is the most recently updated: https://www.who.int/publications/i/item/9789241503662
- **Thank you for your suggestion. We agree with your comment and recalculated the inhalation dose based on the suggested reference and add it to the revised manuscript.**

  - ✓ **Please see the lines 292 to 300:**

    Using the model proposed by WHO (2012) for internal dose from inhalation, $^{131}I^{intg\_conc\_seas}$ (in a unit of Bq m^{-3}) is converted to the effective dose from inhalation of ^{131}I (in a unit of µSv). The model inputs are specified for three age groups of 1-year-old infants, 10-year-old children, and adults (11 years old and up). Considering that about 90% of Qatar's population is in the adult age group (UNStats, 2020), the inhalation doses computed for this age group are discussed here (Fig. 5) and those for two others are available in the supplement (Figures S7 and S8).

LEADBETTER, S. J., JONES, A. R. & HORT, M. C. 2022. Assessing the value meteorological ensembles add to dispersion modelling using hypothetical releases. *Atmospheric Chemistry and Physics*, 22, 577-596.

SAUVAGE, B., FONTAINE, A., ECKHARDT, S., AUBY, A., BOULANGER, D., PETETIN, H., PAUGAM, R., ATHIER, G., COUSIN, J.-M. & DARRAS, S. 2017. Source attribution using FLEXPART and carbon monoxide emission inventories: SOFT-IO version 1.0. *Atmospheric Chemistry and Physics*, 17**,** 15271-15292.

ZHU, C., KANAYA, Y., TAKIGAWA, M., IKEDA, K., TANIMOTO, H., TAKETANI, F., MIYAKAWA, T., KOBAYASHI, H. & PISSO, I. 2020. FLEXPART v10. 1 simulation of source contributions to Arctic black carbon. *Atmospheric Chemistry and Physics*, 20**,** 1641-1656.

Salminen-Paatero, S., Vira, J., and Paatero, J.: Measurements and modeling of airborne plutonium in Subarctic Finland between 1965 and 2011, Atmos. Chem. Phys., 20, 5759–5769, https://doi.org/10.5194/acp-20-5759-2020, 2020.

Evangeliou, N., Balkanski, Y., Cozic, A., and Møller, A. P.: Simulations of the transport and deposition of 137Cs over Europe after the Chernobyl Nuclear Power Plant accident: influence of varying emission-altitude and model horizontal and vertical resolution, Atmos. Chem. Phys., 13, 7183–7198, https://doi.org/10.5194/acp-13-7183-2013, 2013.

Stohl, A., Seibert, P., Wotawa, G., Arnold, D., Burkhart, J. F., Eckhardt, S., Tapia, C., Vargas, A., and Yasunari, T. J.: Xenon-133 and caesium-137 releases into the atmosphere from the Fukushima Dai-ichi nuclear power plant: determination of the source term, atmospheric dispersion, and deposition, Atmos. Chem. Phys., 12, 2313–2343, https://doi.org/10.5194/acp-12-2313-2012, 2012.

Our responses to all comments made by referee #2 are as follows:

- A general comment is that the authors use short (24h) releases (for 365 days) and follow these for 96h. If I understood correctly, in evaluating the impact of these releases the authors look for i) the (seasonal/annual) median of the maximum concentration over each period (24h release with 96h tracking) and ii) the median of the maximum (total over 96h) deposition over each period.

✓ **Indeed, this is exactly what we have implemented. In the revised submission, according to the last comment of the first referee, instead of the maximum value of I-131 concentrations in each 96-hour simulation period ($^{131}I^{conc\_seas\_max}$), the 96-hour integration of I-131 concentrations ($^{131}I^{intg\_conc\_seas}$) is used to calculate inhalation doses (WHO, 2012). The seasonal median of recalculated inhalation doses are presented in figures 5 (for Adult), subplot 7-A (full-year analysis for adults), S7 (for infants), and S8 (for children).**

✓ **No changes were necessary in the analysis related to ^{137}Cs deposition while the color scale has been unified to better visualize inter-seasonal and inter-model variations of radionuclide deposition and concentrations (please see figures 5, 6, 7, S7, and S8). For clarity, instead of $^{137}Cs^{depos\_seas\_max}$, the new abbreviation $^{137}Cs^{tot\_depos\_seas}$ is applied to the seasonal median of total ^{137}Cs deposition.**

- However, usually the release from a nuclear accident has a longer (than 24h) duration. The authors should clearly discuss in the paper how these results may be used to understand what happens in a real case (multiple days release). For example, can this be considered a sort of median daily worst-case scenario? Could it be converted linearly in a season/yearly median worst case by considering it over a longer release period in any season or over the year?

✓ **We conducted consecutive daily simulations with release periods of 24 hours to investigate the effect of variability of atmospheric conditions with hourly/daily temporal resolution on the median distribution of airborne and deposited radionuclides. Our study is not intended to provide actual simulations of radionuclide concentration and deposition levels after a specific nuclear accident in the study area. Rather, our study aims to uniquely capture the range of diurnal and seasonal variations in the transport processes (subsection 3.1), and concentrations and deposition (subsection 3.2) of radionuclides. With our methodology, we can investigate the times of day and year when there is a certain**

probability (risk) of radioactive materials transport and deposition in the study area, and in particular the possible population exposure. We examined the level of uncertainty in the above research questions using different meteorological inputs and different model parameterisations and codes for the entire study.

- ✓ **To emphasize these points, the following paragraphs are added to the revised manuscript.**

  Added to the lines 76 to 78:

  To the best of our knowledge, this is the first time that such a study is conducted for potential radionuclide releases in the study region, and we expect that our results can contribute to the formulation of preparedness plans.

  Added to the lines 184 to 187:

  We simulated a fictitious release of radioactivity at a level comparable to the Fukushima nuclear accident. However, our study does not replicate previous accidents or simulates a specific real-world case; rather we designed this analysis to determine what time of the day and year is associated with particular probabilities of transport of dense radionuclide plumes from any hypothetical release in the study area.

  Added to the lines 319 to 322:

  We note that one should not expect our results to reproduce the simulations of the Fukushima NPP accident. In this study, we primarily aim to determine the relative risk by highlighting variations in the spatio-temporal distribution of radionuclides in the study area due to differences in the diurnal and seasonal atmospheric processes and modeling conditions.

Other comments

- 1) Line 201. The authors write "The relatively lower spatial resolution of CFSv2 caused a smooth distribution of its simulated air parcel ages that is close to the average of other distributions". This does not seem correct to me. For example, in fall, spring and summer (2A-all intensities) the value of CFSv2 before 25h is generally higher than all other ensemble members (therefore cannot look like an average). Moreover, "by eye" the smoothness does not seem different to me (2A-all). I suggest avoiding this statement as it is not necessary for the discussion.
- ✓ **Removed from the result and conclusion sections.**

- 2) Line 210, “age distribution produced by FNL-WRF was found to be more similar to the one produced by ERA5-WRF than by FNL”. This is very difficult to see from figure 2A-all in my opinion, it is somewhat clear in figure 2A-high. Did you use a metric? or is this a “by eye” evaluation? 3) Similarly, to (7) and (8) above, “Although the base model used for the production of FNL, the Global Forecast System (GFS), is also the atmospheric component of CFSv2, FNL age distributions look closer to those from ERA5- and FNL-WRF”. From the figure 2A-all it is very difficult for me to see these claimed similarity/difference. Perhaps you need to add a distance metric that may objectively evaluate what distributions are closer to each other.

    - ✓ **In the revised manuscript we employ a metric (the maximum normalized difference) to determine the similarity between distributions of smooth density estimates of air parcel ages. The following information was added to the revised manuscript.**

      Added to the lines 151 to 156:

      For the one-by-one comparison of age distributions of air parcels from ensemble members, we have used the maximum normalized difference as defined in Eq. 8. Assuming that a and b are two distributions of smooth density estimates of air parcel ages, their normalized difference is defined as the maximum value of the absolute differences of these two distributions divided by the maximum value of these two distributions. Higher values of this metric indicate greater differences in distributions (Jin and Kozhevnikov, 2011).

$$\textbf{maximum normalized difference} = \textbf{max}\big(\textbf{abs}(\textbf{a}-\textbf{b})\big) / \textbf{max}\,(\textbf{max}(\textbf{a}), \textbf{max}(\textbf{b})) \quad \textbf{(8)}$$

Added to the lines 218 to 225:

The maximum normalized distance of age distributions (Fig. 3) shows larger similarity between FNL-WRF and ERA5-WRF than between the former and FNL in all seasons other than the fall (0.3, 0.19, 0.25, and 0.38 vs. 0.3, 0.32, 0.32, and 0.41 in fall, spring, summer, winter). This may to be due to the use of meteorological inputs with the same spatio-temporal resolution and a common simulation code and, consequently, similar modeling schemes for the two former members. Although the base model used for the production of FNL, the Global Forecast System (GFS), is also the atmospheric component of CFSv2, the distribution differences were found

to be lower between the FNL age distribution and those from ERA5- and FNL-WRF for most seasons.

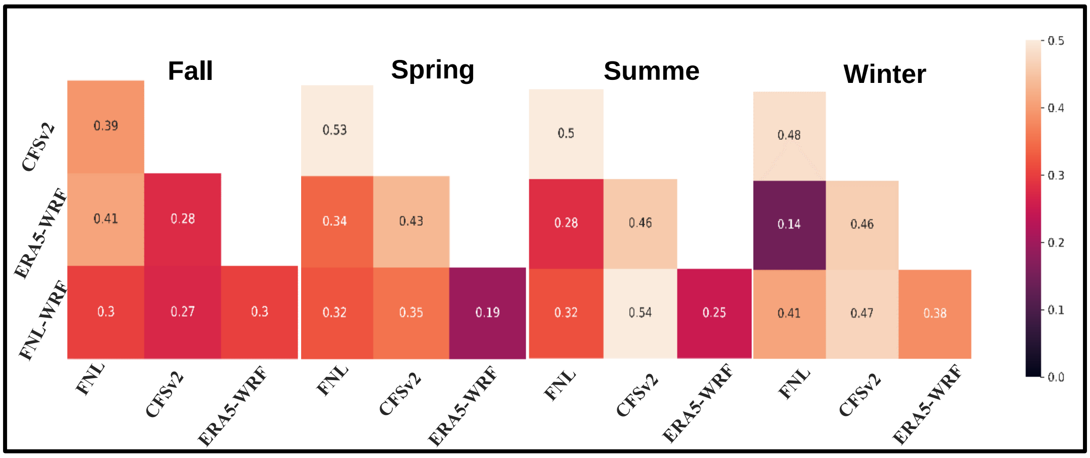

Figure 3 the seasonal maximum normalized difference of smooth density estimates of air parcel ages.

- 4) A general comment is that the discussion (lines 209-215) related to figure 2A-all (see point 7-9 above) comparing the age distributions over the whole 96h age interval seems not objective and perhaps not needed. I think that Plot 2A-all is useful for finding/pointing to specific differences that are obvious for a specific age intervals, e.g. the large peak in FNL and ERFA5-WRF in Winter at about 10hours, or e.g. what pointed out by authors in "air parcel ages are distributed in a wider range in all seasons in FNL (note the location of the first and last peaks", and afterward find the reason for the difference/similarity with a further analysis. On the other end the attempt to compare the full extension (all ages) and evaluate the overall similarity among (two or more) lines crossing repeatedly seems to me very difficult by eye (if not impossible). This comparision would need a specific metric objectively evaluating the overall distance between the lines. Concluding, in my opinion the authors should remove the discussions of 2A-all comparing curves over the whole extension or alternatively add a metric to evaluate the overall similarities/differences among the age distributions.

  ✓ **We followed your second suggestion. We expect that the comparison metric used is sufficiently objective to preserve the discussion related to the full extension of air parcel ages.**

- 5) Please define exactly the density plotted in Figure (2.B-up) and their normalization. Obviously, particles released later in the day have a shorter travel time, i.e. particle released at 24 hours can only travel for 96h-24h=72h. What is the integral under the curves in 0-6h, 6-12h, 12-18h, 18-23h?

  - ✓ **This figure (2.B-up in the first submission and 3.B-up in the revised manuscript) is similar to 3-A (2-A in the first submission), but shows the age distribution of particles stratified in four 6-hour parts of the day. For example, the upper left panel shows the age of the particles that released until 6 am on the first day of the simulations.**
  - ✓ **Results show that particle ages peak between 20 and 70 hours after the release (Figure 2) and that the age of the particles leading to moderate and high radionuclide intensities (Figures 2 and S3) is less than 20 hours. Hence, we found it unnecessary to normalize particle ages. Only a very small portion of the particles has transport time in excess of 80 hours. To clarify this, the following lines were added to the text.**

    Added to the lines 247 to 249:

    This panel is similar to 2-A, but shows the age distributions plotted separately for particles released in 6-hour periods of the first day of the simulations. For example, the upper left figure shows the age distribution of the particles that is released within the first 6 hours of simulations.

    Added to the lines 253 to 258:

    Because particles that are released at the end of the day have less time to travel by the end of simulation period, a sharper fall is observed in the right tail of the age distributions during the second half of the day. However, the lifetime of most simulated particles, especially of those that caused moderate (bottom row in Fig. S3-B and S4-B) and high intensities (bottom row in Fig. 2-B and S1-B), is found to be between 20 and 70 hours after release. Consequently, the difference in release time of Lagrangian particles is not significantly affected by their age spectrum.

- 6)Line 256/ figure 3. The authors should add the formula used to define the deposition as plotted in figure 3. The current explanation (by words) lacks clarity, and the exact mathematical formulation should be added.

- ✓ **Equation 9 is added to clarify the way we calculated the normalized deposition values.**

  Added to the lines 272 to 278:

  To analyze the relationship between the age composition of air parcels and the amount of ^{137}Cs deposition, the deposition values cumulatively summed through time steps (j) and age spectra (i) are normalized to the total amount of ^{137}Cs deposition simulated at each grid cell (k) at the end of each simulation run (l).

$$^{137}Cs_{klta_{norm\_depso}} = \begin{cases} \dfrac{\sum_{i=1}^{a} {}^{137}Cs_{klij}}{\sum_{j=1}^{96}\sum_{i=1}^{96} {}^{137}Cs_{klij}} & if\ j = 1 \\ \dfrac{\sum_{j=1}^{t-1}\sum_{i=1}^{96} {}^{137}Cs_{klij} + \sum_{i=1}^{a} {}^{137}Cs_{klij}}{\sum_{j=1}^{96}\sum_{i=1}^{96} {}^{137}Cs_{klij}} & if\ j > 1 \end{cases} \quad (9)$$

  Figure 4 shows the normalized deposition amounts ($^{137}Cs_{klta_{norm\_depso}}$) in winter when both dry and wet deposition occur in the study area.

- Also include the definition of upper and lower bounds of the green shaded interval.

  - ✓ **This is clarified in the text 289 to 290:**

    Error bars show the 25th and 75th percentiles (the lower and upper quartiles) and the range of normalized deposition (the upper and lower extremes) within the study area.

- Also, may you explain why the median in figure 3 occasionally decreases? Given the 30 years half-life of 137Cs, I would expect that in any grid cell the deposition increases toward its maximum at 96 hours. Therefore, the median should always increase.

  - ✓ **This is clarified in the text 281 to 283:**

    Given that the figure shows the amount of deposition across the whole study area, the decrease in the levels of accumulated deposition at the end of the simulation period pertains to the areas that are far from the source.

- 7)Line 257, the authors write "To perform analysis related to radionuclide concentrations, the average of the simulations in the lowest four layers of the model between 5 to 100 meters has been used". Please add the mathematical definition. Are these layers evenly spaced? If not the average over the four layers should be defined accordingly to the different vertical extensions of the layers (please specify).

- ✓ The information is added to subsections 2.1 and 3.2.

  Added to the lines 158 to 160:

  the thickness-weighted averages of simulations in the lowest four model levels between 5 and 100 m agl (with layer thicknesses of 5 m, 5 m, 40 m, and 50 m) are used for the spatial analysis (subsection 3.2)

  Added to the lines 293 to 294:

  To perform the analysis related to radionuclide concentrations, the thickness-weighted average of the simulations in the lowest four layers of the model between 5 to 100 meters has been used.

- 8)Figure 10, S10, S11. The quality of these plots is poor.

  - ✓ **All abovementioned Figures are recreated with higher quality.**

- 8.1) I think that the full year should not be overlapped with the seasons. 8.2) There is a lot of empty space on the right of the diagonal that can be used for plotting the full year separately.

  - ✓ **Changed. The full-year analysis is now shown in the upper triangular of matrix plot.**

- 8.3) On the diagonal, the colored areas should be replaced with lines so that all the seasons can be clearly distinguished. What is the title of the vertical axis?

  - ✓ **Changed.**

    Added to the lines 470 to 471:

    ***The density plot (unitless) in the main diagonal of the evaluation matrix shows the relative distribution of simulations*.**

Minor comments

- Eq. 2, dW_i should be dW_ j.

  - ✓ **Corrected.**

- Line 100-101. “Wiener process with mean zero and variance dt”, the “dt” is missing.

- ✓ **Corrected.**

- Line 108. I think that Cassiani et al (2013) should be (2015) as the reference.

  - ✓ **Corrected.**

4) In table 1, add a further column indicating the deposition scheme used in FLEXPART (10.4 vs 9.02).

- ✓ **Added.**

5) Line 202-203 the phrase "could not be so great …." is unclear. Please rephrase it.

- ✓ **Corrected, please see the lines 209 to 212:**
  The significant similarity of age distributions of ^{131}I and ^{137}Cs indicates that differences in the transport characteristics of these radionuclides, such as the wet and dry deposition rate and radioactive decay, are not large enough to cause the abundance of cases where ^{131}I and ^{137}Cs particles are not present in a common grid.

6)Line 203 what do you mean with "base concentration" ?

- ✓ **We meant low concentrations. Rephrased to *The close dispersion of ^{131}I and ^{137}Cs concentrations, especially at low intensities …. (lines 212-213)***

7) Line 239, may you clarify what do you mean with "to the further parts of the study area…. ". In relation with the peak at high particles age in the spring.

- ✓ **Corrected. please see the lines 244 to 246:**
  Therefore, it can be concluded that regional atmospheric circulations led to the more distant transport of radionuclides to northern parts of the study area in this season than in other seasons.

8)Figure 3, add axis titles on both the axes.

- ✓ **Done.**

9)Figure 4 and 5, add the grid spacing in the axis and explain the units of the contour lines.

- ✓ **Grid information and units are added.**

10) Line 354, "less than above thresholds" seems awkward language to me.

- ✓ Corrected, **please see the lines 376 to 378:**
  The populated areas (with a density of more than 15 people per arc-second) exhibit lower expected inhalation doses and $^{137}Cs^{tot\_depos\_seas}$ less than 200 μSv and 100 kBqm−2.

- Figure 6A and 6B, in my opinion it would be better to use a unique (for all models) color scale here.

  - ✓ **Done.**

- I think that Figure S11 should be included in the main manuscript since it is discussed in many details.

  - ✓ **Done. S11 is shown as Fig. 12 in the revised manuscript.**

13) Line 474, what do you mean with "iteratively" here? The models were simply run for the 365 days in the year.

- ✓ **We meant that FLEXPART with a same setting is executed iteratively for each single day of 2019.**

References:

JIN, D. Z. & KOZHEVNIKOV, A. A. 2011. A compact statistical model of the song syntax in Bengalese finch. *PLoS computational biology,* 7**,** e1001108.

WHO 2012. *Preliminary dose estimation from the nuclear accident after the 2011 Great East Japan Earthquake and Tsunami*, World Health Organization.