Dear Reviewers,

We would like to thank you once again for your feedback that has improved the manuscript. Following your comments, the following changes were done in the updated version of the manuscript:

➔ Comments from Reviewer RC1

1. General Comments:

- The authors provide useful information in line 200 and further, but can I ask to also include the given response (see above) in the text? Perhaps something is already there and I've missed it, but if not, I think it is useful information to include.

  Lines 215-217: "The duration (or Total Lifetime) of the storm is then the lifetime of the radar pixels group as dictated by the threshold used to recognize them and the tracking algorithm that decides if the same storm is observed at continuous time steps."

- This response may be an interesting discussion point to add to the paper.

  Lines 450-454:" The Total Lifetime is an easier target to be analysed, which means the values are not zero and its distribution is not as heavy tailed as the distribution of the other variables. The other variables, depending on the lead time, have more zeros included and have an asymptotic density function. It seems that, whenever zeros are not present, like in the case of storms lasting longer than 3 hours, the PIC is able to represent quite well the important predictors."

2. Specific Comments:

- I agree, but feel free to even mention that field-based nowcasting approaches generally have trouble capturing and forecasting this well, which would stress even more the reason to focus on object-based nowcasting methods.

  Lines 70-73:" Even though the field-based approached has gained popularity recently (Ayzel et al., 2020; Imhoff et al., 2020) they still have trouble nowcasting convective storms. Thus, the focus in this study is on object-based nowcasts as they are more convenient for convective storms that typically cause urban pluvial floods."

- The authors have, indeed, increased the font size of the figures. For some figures, the font size can still be considered somewhat small. Nevertheless, this is also a process of typesetting later on. A suggestion that I would like to make, is to sometimes consider placing the figures (e.g., but not limited to, figures 7 and 8) on multiple rows. That would allow for using larger figures and that would solve the problem, too.

  I think I haven't increased the label size of Figure 7 and 8, but I can change them to the same size like Figure 9. I am currently working on that. I would like to keep the same settings (in columns) as in the other following Figures, so it is easy to associate for the reader.

3. Technical corrections:

   All noted and changed!

➔ Comments from Reviewer RC2

1. General Comments:

- There is no description about how merging and splitting of storm cells are handled. The way this is done can have a significant impact on the results.
  - Are all merged/splitted cells included in the considered storm tracks (so that the storm tracks form a tree and a storm object at a given time may consist of multiple cells)? •
  - Or do you only consider storms that do not merge or split during their lifetime?

I suppose you are referring here to the HyRaTrac nowcast algorithm which serve as a base for the tracking and building the storm database. Merged/Splitted Cells are included in the storm tracking, please refer to one of your specific comments to see which line in the text was changed to clarify this process better. Regarding the k-NN, splitting of the cells is (because of HyRaTrac) included indirectly in the forecast. The k-NN is able to forecast splits if past neighbours have splatted during their lifetime. The Merging is also possible with the k-NN but needs an extra step (implemented in part 2 for rainfall intensities): two storms are predicted at the same region, and thus they are superimposed and considered as a merged storm.

2. Specific Comments:

- Manuscript title: since you now have "Part I" in the title, you should have more explicit discussion in the conclusions what would "Part II", or even "Part III" include.

Lines 753-769: "Improving the nowcasting of storm characteristics is the first step in improving rainfall nowcasting at fine temporal and spatial scales. On a second step, the knowledge about the storm characteristics (as nowcasted by the 30NNs) should be implemented on the spatial structure of the storms to estimate rainfall intensities at fine scales (1km^2 and 5min). There are two options to deal with the spatial distribution of the rainfall intensities inside the storm region (which is so far not treated in this study): 1. Increase/Reduce the area by the given nowcasted area (as target variable) for each lead time, scale the average intensity with the nowcasted intensity, and move the position of the storm in the future with the nowcasted velocity in x and y direction. 2. Take the spatial information of the selected neighbours, perform an optimisation in space (such that present storm and the neighbour's storms locations match) and assign this spatial information to the present storm for each lead time. The former is an extension of the target-based 30NNs, while the later an extension of the storm-based 30NNs. So far, the comparison between these two versions, showed that the target-based approach is better suited mainly to nowcast the velocity components, thus a merging of the two could also be reasonable: the storm-based approach is used for nowcasting Area-Intensity-Total Lifetime (features that are co-dependent based on the life cycle characteristics of convective storms), and the target-based approach for the nowcasting of the velocity components. Future works (Part II – Local Intensities) will include the integration of the developed 30NNs application in the object-oriented radar based nowcast to extend the rainfall predictability limit at fine spatial and temporal scales (1km^2 and 5min). The main focus of the Part II is to investigate if the methodology applied here can introduce improvements as well at the local scale, i.e. validation with the measurements from the rain gauge observations."

- Line 7: erratic ← unpredictable?

Noted and changed!

- Lines 44 and 57: I'm not sure if these claims are completely true for state of the art rapid-update limited-area NWP models that could be applicable to urban-scale nowcasting. Are there any more recent references about this topic?

  Lines 41-44: "The Numerical Weather Prediction Models (NWP) are typically used in hydrology for weather forecast to several days ahead, nevertheless they are not suitable for urban modelling as they still cannot produce reliable and accurate intensities for spatial scales smaller than 10km^2 and temporal time steps shorter than an hour (Kato et al., 2017; Surcel et al., 2015)."
- Line 46: You could add that the inability to capture the spatial structure of rainfall is due to the sparsity of the existing rain gauge networks.

  Lined 44-46: "Ground rainfall measurements (rain-gauges) are considered the true observation of rainfall but they are as well not adequate for QPFs because, due to the sparsity of the existing rain-gauge networks, they cannot capture the spatial structure of rainfall."
- Line 67: I'm not sure if it's necessary to say that you are using a region of size W. In my opinion, this is a technical detail that does not belong to the introduction.

  Noted and changed!
- Line 68: When talking about stratiform rainfall, I would not use the word "storm".

  Noted and changed!
- Line 76: Rather than being observed directly, the velocity vectors are estimated from consecutive storm objects.

  Lines 76-77: "… and velocities are assigned from consecutive storm objects…"
- Line 170: Please explain what is the accumulation time of the gauges. Or is the quantity measured by gauges a 5-minute averaged intensity?

  Lines 171-172: "…while the rain-gauges measure the rainfall intensities at 1min temporal resolution but are aggregated to 5min time steps…"
- Line 172-173: Taking a gauge-interpolated field instead of a radar image is highly questionable. It does not contain any information about the small-scale features, so why not completely exclude missing time stamps from the dataset?

  They are excluded from the application of k-NN, because their total lifetime (or duration) is only 5min. All storms recognized as single time steps are removed from the k-NN application.
- Lines 178-179: This is not clearly written. This seems to describe the condition for the end of an event, but it's not clear to me how the start of an event is defined.

  Lines 181-182: "… The start and the end of the rainfall event is determined when areal mean radar intensity is higher/lower than 0.05mm for more than 4 hours…"
- Lines 186-188: This is lacking essential information. Should you also mention that in such a group of pixels (grid cells), the pixels also need to be spatially connected (e.g. they have at least one neighbour in the group)

  Lines 189-190: "… A storm is initialized if a group of spatially connected radar grid cells (> 64) has a reflectivity higher than Z=20dBz…"
- Lines 185-194: Estimation of the storm displacements from cross-correlation between two images is described here, but should you also describe the matching of storm objects between different time steps in more detail? And how are merges and splits handled?

  Lines 195-203: "…Once storms at different time steps are recognized, they are matched as evolution of a single storm, if the centre of intensity of storm at t=0 falls within the boundary box of the storm at t-5 min. The tracking of individual storms in consecutive images is done by the cross-correlation optimization between the last 2 images (t=0 and t-5 min), and local displacement vectors for each storm are calculated. In case a storm is just recognized (the storm does not yet have previous history), then global displacement vectors based on cross-correlation of the entire radar image are assigned to them. It is usually the case, that two storms merge together at a certain time, or a single storm splits between several daughter storms. The splitting and merging of the storms is considered here if two criteria are met: a) the minimum distance between the storms that have splatted or merged is smaller than the perimeter of the merged or that-is-splitting storm, and b) the position of the centre of intensity of former/latter storms is within the boundaries of the latter/former storm."

- Line 193: "storm is just recognized"? Do you mean that the storm does not yet have previous history?

  Lines 198-199: "…(the storm does not yet have previous history)…"

- Line 198 and the following text: The term "state" is not precisely defined when you first introduce it, which makes its meaning unclear to the reader.

  Lines 205-207: "…and for each time step of the storm evolution the spatial information is saved and various features are calculated. Here the features computed from the spatial information of the rainfall inside the storm boundaries at a given time step (in 5min) of the storms' life, is referred to as the "state" of the storm…"

- At lines 197 and 198, you define the state as the "spatial structure of the rainfall inside the storm boundaries". It is not clear what this means. Please elaborate.

  Here we meant the spatial information, but the sentence was updated (see above) to avoid the confusion.

- I would call all the features of the object together as the state of the storm. Later (e.g. line 313) you in fact use the term in this way because you are comparing the states of the storms against each other. So, could you define the state in this way when you first introduce the term at line 197?

  Noted and changed! Please see the two points before.

- Line 214: Again, I don't think that it makes much sense to use gauge-only fields as inputs. Could you just exclude time stamps with missing radar data from your dataset?

  They are already excluded from the application of k-NN, because they live only 5min. All storms recognized for single time steps are removed from the k-NN application.

- Line 315: Table 2: Should this be Table 3?

  Noted and changed!

- Equation (6): The weights Pr for the deterministic nowcast are not explicitly specified. Are they set to 1/k in this case?

  Line 338: "… where k is the number of neighbours obtained from optimization, Ri and Pri (from Equation 7) are respectively the response and weight of the ith neighbour…"

- Equation (7): This should be immediately after equation (6), since the weights Pr are mentioned there for the first time.

  Please see the point above.

- Equation (8): Should the MAE terms be the absolute values of the differences, and not the differences of absolute values, as it is currently written?

  I'm sorry for the confusion, but yes you are right and Equation 8 has been updated accordingly.

- Lines 385-386: It is stated that Y' is the finite first moment? Is this correct? Isn't Y' a random variable? And please define the symbol E (expectation).

  Line 339-400: "… where F is a probabilistic forecast, y the observed value, Y and Y´ independent random variables with CDF of F and finite first moment E (Gneiting and Katzfuss, 2014)."

- Section 4: I'm not able to follow how you compute the MAE when verifying the nowcasts. As in Section 3.2.1 (equation 8), you also need to explicitly define the MAE used for the verification in Section 3.2.2.

Lines 389-392; "i) absolute error per lead time and target variable computed for each event and for a specific selected nowcast time

$$MAE_{target} = \sum_{i=1}^{N} (|Pred_{i,+LT} - Obs_{i,+LT}|)/N ,$$

where the P_{red} is the predicted response, Obs the observed response for the i^{th} storm, +LT the lead time and N the number of storms considered inside an event."

- Table 3: I'm not fully able to follow the notation. Why are you using the symbols I in Table 1 but in Table 3 you use PI? And what are PI_sd1 and PI_sd2?

  I'm sorry for the confusion. The Table 1 and 3 notations (also in the Appendix) are now changed to be consistent with one another.

- Lines 398-402: It is confusing that in Table 3 you show the correlation coefficients but then directly move into the predictor weights (the tables in the appendix) without explicitly explaining how the weights are obtained from the correlation coefficients. Are the former directly obtained from the latter? Please make this more clear.

  Lines 291-292: "...Here, the Pearson correlation absolute values are used directly as predictors weights in the k-NN application."

- Lines 416-417: I'm not able to follow this. What rows/columns of Table 3 or the tables in the appendix are you looking at when deciding what predictors are the most important?

  Lines 437-440: "Hence based on the Pearson correlation values from Table 3 the following most important predictors were selected: Area –A (as maximum correlation value from first row), Intensity –PIsd1 (as maximum correlation value from second row), - Velocity X – Vx30 (as maximum correlation value from third row), Velocity Y –Vy30 (as maximum correlation value from fourth row),Total Lifetime – A (as maximum correlation value from fifth row)."

- Figure 8 and line 482: It is not clear to me how you minimize the ME. It is a quantity that may have arbitrarily large negative values. Are you taking absolute value somewhere in the minimization process?

  Lines 369-374: "...The objective function is the minimization of the mean absolute error (Equation 8) and of the absolute mean error (Equation 9) between predicted and observed target variables at lead times from +5min to +180 min:

  $$MAE_{target} = \sum_{i=1}^{N} (|Pred_{i,+LT} - Obs_{i,+LT}|)/N , \quad (1)$$

  $$ME_{target} = |\sum_{i=1}^{N} (Pred_{i,+LT} - Obs_{i,+LT})/N | , \quad (2)$$

  where the Pred is the predicted response, Obs the observed response for the ith storm, +LT the lead time and N the number of storms considered inside an event"

- Figure 9: I'm not able to follow how you compute the MAE for a nowcast longer than 30 minutes for a storm, whose lifetime is less than 30 minutes. What are you comparing the nowcast against?

  Lines 413-416: "Lastly, it is important to notice, that the performance criteria can be calculated even for nowcast times longer than the storm lifetime, if the nowcast fails to capture the dissipation of the storms. In this case, Area, Intensity, Velocity in X and Y Direction are compared against zero, while the Total Lifetime against the total observed lifetime of the storms."

- Line 534: I'm unable to find Figure 10 in the manuscript. Where does this refer to? Should Figure 11 on page 20 be Figure 10?

  Yes! The Figure numbering has been correctly updated.

- Page 22: The figures are in wrong order. Also, are the figure numbers correct?

  Yes! The Figure numbering has been correctly updated.

- Lines 647-661: This discussion is beyond the scope of the current section (verification of the ensemble nowcasts). Could it be moved to Section 5?

  Noted and changed!

- Line 680: It is stated that the predictability limit of the Lagrangian persistence is one hour. Please make clear what type of Lagrangian persistence are you talking about because in the introduction you give different predictability limits for different nowcast types (grid- vs. object-based).

  Lines 691-692: "…and more importantly the extrapolation of the storms in the future based on the Lagrangian persistence, are limiting the forecast horizons of such object-oriented radar based nowcasts to 30-45 min for convective storms and to 1 hour for stratiform events (Shehu & Haberlandt, 2021)."

- The appendix: The figures should have caption texts instead of placing the explanations in the subsection titles.

  Appendix 8-1 Strength of relationship between the selected predictors and the target variables averaged for three lead times and storm duration groups computed from Pearson correlation. The green shade indicates the strength of the relationship: with 0 for no relationship at all, and 1 for highest dependency. The averaged computed values for each target variable (last row) are used as input for Table 3. The correlation weights are absolute values of the correlation values between the predictors at specific lead times and target variables.

  Appendix 8-2 Strength of relationship between the selected predictors and the target variables averaged for three lead times and storm duration groups computed from PIC method. The green shade indicates the strength of the relationship: with 0 for no relationship at all, and 1 for highest dependency. The averaged computed values for each target variable (last row) are used as input for Table 3. For intensity, velocity in x and y direction, since the PIC recognized only one predictor as important, the average values is fixed as 1 for the respective predictor.

  Appendix 8-3 The standard deviation of the Pearson Correlation Weights between predictors and target variables obtained from a cross-sampling of the events (leave one event out). The boxplot for each target variable describes the spread of the standard deviation over all selected predictors.

3. Figures:

- Figure 1: The notation in the figure caption is inconsistent. In t-delta_t and t+LT, you are using notation with and without subscripts. Please use only one notation. And should the subscript 0 also be included to t in the middle and the right pane?

  Noted and changed!

- Figure 3: Should this be split into separate figures? Only Figure 3a is referred in the same section with the figure. Figures b and c are defined only much later.

  I understand your concern, but I would like to keep it like this because it saves place, and keeps all the information at one place. I hope this won't bother the reader too much.

- Figure 5 and the caption text: You are using both t_+TL and t_t+LT. Use only either one to avoid confusion.

  Noted and changed!

- Figure 5: It is not clear why you are using I with and without hat. What does the I with hat mean?

  I have changed it to simply I.

- Figure 5: Here you are using psi for the orientation angle, but in Table 1 you use phi. Please use either one to avoid confusion.

  I have updated the Figure and used only phi.

- Figure 9: Please explain the meaning of "nowcast time" more clearly. Does it mean the current lifetime of the storm when the nowcast is issued?

  The following explanation was added to the caption of the figure: "… Nowcast time dictates when the nowcast is issued relative to storm initiation."

- Figures 11 and 12: To me it looks like that the line styles don't match the descriptions in the caption text. It is stated in the text that the probabilistic nowcasts consistently outperform the deterministic ones in terms of CRPS. This would be the case if the probabilistic nowcasts were plotted with solid lines and the deterministic nowcasts with dashed lines, which is the opposite as stated in the caption text. Also note that the labels in the legends inside the plots contradict with the caption texts.

  Yes indeed, the caption was wrong and it has been updated accordingly.

4. <u>Technical corrections:</u>

   All Noted and changed!