

## **Review of "Improving radar-based rainfall nowcast by a nearest neighbour approach: Part I – Storm Characteristics" by Bora Shehu and Uwe Haberlandt**

Seppo Pulkkinen  
seppo.pulkkinen@fmi.fi  
January 6 2022

### **Summary**

I acknowledge that the authors have adequately addressed my previous concerns. I still have a list of mostly minor technical issues related to the presentation of the results. Once these have been addressed, I'm willing to recommend the manuscript for publication.

### **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?

### **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.
- Line 7: erratic ← unpredictable?
- 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?
- 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.
- 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.
- Line 68: When talking about stratiform rainfall, I would not use the word "storm".
- Line 76: Rather than being observed directly, the velocity vectors are estimated 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?
- 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?
- 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 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 neighbor in the group).

- 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?
- Line 193: "storm is just recognized"? Do you mean that 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.
  - 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.
  - 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?
- 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?
- Line 315: Table 2: Should this be Table 3?
- Equation (6): The weights  $P_r$  for the deterministic nowcast are not explicitly specified. Are they set to  $1/k$  in this case?
- Equation (7): This should be immediately after equation (6), since the weights  $P_r$  are mentioned there for the first time.
- 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?
- 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).
- 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.
- 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  $PI\_sd2$ ?
- 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 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?
- 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?
- 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?
- 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?

- Page 22: The figures are in wrong order. Also, are the figure numbers correct?
- 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?
- 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).
- The appendix: The figures should have caption texts instead of placing the explanations in the subsection titles.

## 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?
- 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.
- Figure 5 and the caption text: You are using both  $t_{+TL}$  and  $t_{t+LT}$ . Use only either one to avoid confusion.
- Figure 5: It is not clear why you are using  $I$  with and without hat. What does the  $I$  with hat mean?
- 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.
- 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?
- 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.

## Technical corrections

- Line 42: "weather forecast at several days ahead" ← "weather forecast to several days ahead"
- Line 44: "short than an hour" ← "shorter than an hour"
- Line 48: scales ← resolutions?
- Line 55: nowcast ← nowcasts
- Line 62: intermittent? Do you mean continuous?
- Line 70: nowcast ← nowcasts
- Line 101: govern ← dominate?
- Line 102: remove the word "rainfall"?

- Line 103: consulting ← utilizing?
- Line 148: at ← for
- Line 161: check the language
- Line 206: "tracking identification and algorithm" ← "tracking and identification algorithm"
- Line 209: clutters ← clutter
- Line 306: averages ← average
- Line 326: Please add subscripts  $i$  to  $R$  and  $P_r$ .
- Equation (7): In the numerator, you have  $\text{Rank}_i$  but in the denominator you have  $\text{Rank}_i$ . Should the  $i$  be written as subscript?
- Figure 5, caption text: "The nowcast is issued time  $t_0$ " ← "The nowcast is issued at time  $t_0$ "
- Line 342: "specifically per each variable" ← "separately for each variable"
- line 387: enambles ← enables
- Line 454: "important analysis" ← "importance analysis"
- Line 563: introduces ← introduced
- Line 566: are dependent ← depend
- Page 22: Figure 14 ← Should this be Figure 11?
- Line 643: towards ← over
- Page 25: Figure 154 ← Should this be Figure 14?
- Line 690: Remove "of the".
- Line 699: high ← long
- Line 701: "on two measurements of similarity" ← "based on two similarity metrics"
- Line 712: "combination of the" ← remove the word "the"
- Line 725: works ← improvement
- Line 735: lightening ← lightning
- Line 736: angels ← angles