Review of “Improving object-oriented radar based nowcast by a nearest neighbor approach” by Bora Shehu and Uwe Haberlandt

Seppo Pulkkinen
seppo.pulkkinen@fmi.fi
July 4 2021

Summary

The authors propose a methodology for predicting the evolution of storm cells by comparing them to past storms. This is done by using a nearest neighbor approach. Overall, I find this a highly relevant and well-written paper with scientifically significant results. The proposed approach has a high novelty value. I did not find any fundamental flaws in the methodology. The paper is also well within the scope of HESS. However, I have several questions and comments about the data selection, how the verification is done and some requests for clarification. The literature review is somewhat lacking to put the work into a broader context. There is also room for improvement in how the results are presented and the language needs some polishing. However, all these are minor details that can be improved with a small amount of additional work.

I'm looking forward to see the revised paper. I can recommend it for publication once the following concerns have been addressed.

General comments

- To make the title better reflect the content, you could add the word "rainfall" because the paper is about rainfall nowcasting.
- Before going directly to the matter, it could be worthwhile to add one general paragraph about nowcasting. Like why nowcasting is done, its societal need and what kind of hazards can be prevented with reliable rainfall nowcasts.
- In the beginning of the introduction, the authors should make a more clear distinction between the two nowcasting approaches (field- and object-based) and add more description about what purposes they are used for. For instance, mention that field-based methods are well-suited for predicting large-scale stratiform precipitation systems but cell-based methods are best-suited for predicting the motion of intense convective cells.
- To put their work into a broader context, the authors could mention in the introduction that the proposed approach is conceptually similar to the so-called analogue-based nowcasting. The idea of this approach is to look for similar events from a large sample of archived radar data. See, for instance:

  L. Panziera, U. Germann, M. Gabella and P. Mandapaka: NORA–Nowcasting of Orographic Rainfall by means of Analogues, Quarterly Journal of the Royal Meteorological Society, 137(661), 2106-2123, 2011.

  There are a number of others, so I recommend the authors to do a literature review. However, all the previous studies I know attempt to find analogs from full radar images, not from individual cells or their features. This is a novel aspect, which should be clearly pointed out in the manuscript.
- I have concerns about the choice of the predictors. The proposed methodology opens the possibility to use a large number of different predictors (and targets).

However, the set chosen in the study is in my opinion quite limited and the capability of the model is thus not fully utilized. In addition, they are more or less correlated with each other, and also with the target variables, which the authors admit. I think that using the following additional predictors could reveal the full potential of the model:

- convective available potential energy (CAPE)
- convective inhibition (CIN)
- signatures from radar-measured Doppler and polarimetric parameters, as well as vertical profile information obtained by using all elevation angles
- lightning flash density
- geographical features like terrain altitude or proximity of water bodies

These are probably beyond the scope of this study, but I encourage the authors to include them in a follow-up paper. In addition, the authors could replace the generic description of additional predictors in the last paragraph of Section 5 by specifically mentioning some of the above.

Note that a relationship between CAPE and CIN and the life cycle of convective cells is suggested in:

C. Moseley, O. Henneberg and J. O. Harter: A Statistical Model for Isolated Convective Precipitation Events, Journal of Advances in Modeling Earth Systems, 11(1), 360-375, 2019.

- A fundamental reason why similar storm cells behave similarly is that their life cycles follow characteristic patterns. For instance, the areal extent and intensity of storm cells are related to each other and the storm lifetime. In particular, the future behavior of cells depends on what stage they are in their life cycle. I think this aspect needs to be discussed more in the paper with literature references to put the research in a broader context. To this end, the authors could study the following papers:

  H. Kyznarova and P. Novak: CELLTRACK-Convective cell tracking algorithm and its use for deriving life cycle characteristics, Atmospheric Research, 93(1-3), 317-327, 2009

  C. Moseley, P. Berg and J. O. Haerter: Probing the precipitation life cycle by iterative rain cell tracking, JGR: Atmospheres, 118(24), 361-370, 2013

  There is also a large amount of meteorological literature, where the life cycles of convective storms are discussed.
- A major limitation of the k-neighbors approach is that it cannot generalize beyond the training data. How would the proposed method perform for extreme events that have a very limited number or no training samples? Please add more discussion or analysis about this.
- In many places, the authors are describing results that are not shown anywhere, so the reader cannot verify the validity of the claims. An example of this can be seen at lines 362-374. Could you include some of the not shown results that are discussed in the text in an appendix or in supplementary material?

Specific comments

- Line 101: What does "step 3" refer to? Storm extrapolation in Figure 1?
- Figure 2:
  - What do the numbers represent in the x- and y-tick labels? It think it would be more informative to show the distance from the radar in kilometers.
  - Does DEM mean altitude obtained from a digital elevation model?
- Lines 131-137: What is the justification for these threshold choices?
- Lines 139-140: What is the "spatial rainfall intensity of a storm". Is it some kind of average or maximum value taken inside the storm object?
- Line 142: I'm curious how the ellipsoid is fitted. Please provide a more detailed explanation (though no need to include this in the paper).
- Line 145: Please give a more detailed description about how the storm velocities are estimated?
- Line 145 and Figure 3: The merges are mentioned in the text but not shown in the figure.
- Figure 3:
  - The very high velocities of 5-minute storms look suspicious to me. How can you even estimate the velocity of a storm if its duration is only 5 minutes (i.e. one time step)?
  - Is there a reason for specifying the duration intervals in inclusive way? I would use separate intervals (i.e. 0-1h, 1-3h, 3-6h, 6-12h).
  - I'm very surprised to see how the 5-minute storms have such a large area. I would expect all storms having area over 500 km^2 to have lifetime longer than 5 minutes. Can you explain this?
- Line 152: Please give numbers describing "high intensity" and "low areal coverage".
- Lines 153-154: What is the evidence for making this conclusion? At least this cannot be seen from Figure 3.
- Line 210 onwards: It is not obvious to me how the partial information correlation is better able to capture non-linear behavior than the Pearson correlation coefficient. Can you add more discussion about this?
- Equation (3): How is PI defined?
- Equation (4): The notation is confusing. What does X(-j) mean? Maybe you should use subscripts for j and -j instead.
- Table 1: I would not use the word "predicted" for the target variables. I thought that they are obtained form observations, not predictions. For the velocities, I would use the word "estimated" since they are not directly observed but estimated by using some method.
- Table 1: You could attempt to eliminate the dependency of the standard deviation on the mean value by using the coefficient of variation instead (i.e. standard deviation divided by the mean).
- Lines 255-262: The actual procedure for generating the ensemble is not well-described. Are the ensemble members somehow randomly assigned based on their probabilities?
- Line 261: What is the justification for choosing the value 0.5? Is the model sensitive to this value?

- Equations (6) and (7): To me it appears that the same symbol R is used for two different purposes: response and rank. Could you use different symbols?
- Equation (8): Should the summation terms be taken their absolute values? To me summation over the differences does not make much sense if it's used as the objective function.
- Line 304: Please define precisely the concept of Lagrangian persistence in this context. It can be defined in many different ways depending on the type of the nowcast (i.e. grid- or object-based). Here it means that all storm attributes (not only the shape) are taken from the most recent values and they remain constant for all lead times. Right?
- Figure 6:
  - In Table 1, the target variables A, I, V_x and V_y have the subscript denoting lead time. However, these are is omitted in Figure 6. Thus, it is not clear to me what lead times do the correlations shown in Figure 6 represent. The text is just saying that the values are averaged from three different lead times.
  - The correlations depend on the lead time. Would it make sense to show the correlations separately for each of the chosen lead times instead of averaging over different lead times?
- Lines 365-366: This is difficult to follow. It is confusing that the authors mention both mean and median but are not showing the latter anywhere. In addition, you should clearly state in the caption of Figure 7 that it shows the mean.
- Lines 391 and 397: The authors are using a confusing term "event-based" that is not defined previously. Does this mean storm-based?
- Figure 7: The interpretation of the Total Lifetime figure on the right was not immediately clear to me. In particular, the connection of the black and red lines labeled as VS1 and VS2 to the boxes shown in the figure could be more clear.
- Figure 9:
  - What is the "Timestep of nowcast"? This should be clearly explained in the caption text. Now I found it from line 403 only after reading through the main text.
  - As in Figure 7, it was not immediately obvious to me how to interpret the right pane.
- Figure 10:
  - Again, clarify the meaning of the "Timestep of nowcast". Please explain it in the figure caption.
  - To me it is striking that in the worst case the 4-NN nowcast can perform more than 100% worse than the Lagrangian persistence. This is lacking discussion in the text that focuses mainly on the improvement from the Lagrangian persistence.
  - My advice here is to explore, or at least mention the possibility of blending the Lagrangian persistence and the 4-NN nowcast by using weights that depend on the lead time. This would combine the strengths of both approaches.
- Figure 12: It could be more informative to compute instead the fraction of verifying observations within (or outside) the ensemble and average this statistic over the events. The "% of timesteps" statistic gives no information about this fraction.
- Sections 4.2 and 4.3: The authors use the terms lead time and timestep interchangeably. When reading the text, their correspondence is not immediately clear to the reader (until the reader goes to Section 2 to recall that the time step is 5 minutes). Could you use only one of them?

- Section 4.4: I have some doubts whether the "best ensemble member" or "% of ensemble members better than Lagrangian persistence" verification approach is meaningful. These are not standard verification metrics. In practice, one does not know a priori which ensemble members to choose to obtain the best forecast skill. There are more elaborate ways of showing the advantages of ensemble-based predictions over deterministic ones. For instance, you can compute the continuous ranked probability score (CRPS), which is a generalization of mean absolute error (MAE) for deterministic nowcasts. The CRPS of the ensemble nowcast should be lower than the MAE, which indicates the added value of the ensemble nowcast.
- Line 511 onwards: I guess that the authors mean individual ensemble members, not whole ensembles?
- Lines 538-539: Can you give some numbers to describe what are the fine spatial and temporal scales?
- Line 546: Where does the number 5200 come from? It is mentioned for the first time in the conclusions. Perhaps it should be mentioned in Section 2 as well.

Technical corrections

- I'm not sure if it's proper to use the word "object-oriented". It refers to programming terminology. Could you use object-based instead?
- Figure 2: Should the legend read "Lower Saxony border"?
- Line 211: important ← importance?
- Lines 226-227: "the α_j the predictors weight" ← "α_j denote the predictors weight"?
- Line 256: 30-ensembles ← 30-member ensembles?
- Line 447: persistence ← persistent?
- Line 544: behaviours ← behaviour
- Line 578: An increment in the sample size ← Increase in the sample size?
- Figure 6: Should this be titled as a figure or a table?