NHESS-2023-192

Authors' Responses to Reviewer 2 (RC2, anonymous)

Date: 30 May 2024

Title: Investigation of an extreme rainfall event during 8–12 December 2018 over central Viet Nam – Part 2: An evaluation of predictability using a time-lagged cloud-resolving ensemble system

Authors: C. C. Wang et al.

Firstly, **we thank the reviewer for the valuable comments that have significantly improved the clarity and highlighted important points of the paper**

**COMMENTS**

Substantial Comments (Comments are not listed in order of importance):

**Comment 01.** I find the evaluation of predictability for this event incomplete. Grouping ensemble members together is a very superficial evaluation of predictability that does not go into the depth needed to examine the actual cause of the lack of predictability. The ensemble sensitivity analysis is a logical next step, but it is simply presented in the manuscript. The results are not interpreted or physically linked back to the event. They are simply presented.

**Reply:** Thank you for your comments. We are reinstructing our paper to make our current study more logical and clearer. Specifically, we moved Fig. 4 to section 3.2, adding more new results, as well as put more emphasis on the analysis to (1) better clarify the predictability of D18 event, and to (2) present the new findings in our study more effectively and more clearly.

a. The result of predictability increases as lead time decreases is not a new result to the body of literature on meteorological prediction.

**Reply**: You are right. However, our results once again reconfirm that predictability decreases as lead time increases in general, especially for a region with complex terrain and weather systems like Vietnam. Furthermore, in our study, it is shown that the lead time is not the only factor to influence the predictability (as one may suspect). Specifically, there is a good forecast at a longer lead time, made by the

member ran at 1800 UTC 4 Dec. This indicates the potential for heavy rainfall to occur with more time for preparation. We will put more emphasis on this member in the revision, and conclude that it is still possible to have good forecasts at a lead time up to 5 days.

b. An analysis of the differences, physically, between each ensemble run that might be the cause of the lack of predictability should be undertaken. This should be more than just low-level RH and surface winds, as it is well known that large scale features are important to controlling these factors (see the results from Part 1). This will then lend context to the ESA and identify how these sensitivities feedback into the prediction.

**Reply:** Thank you for comments. Based on the results from Part 1, the low-level wind convergence led to moisture convergence and these conditions resulted in the D18 event. Furthermore, the southward movement of the low-level wind convergence also dictated the movement of heavy rainband during the event. Therefore, we would like to put more focus on these two aspects in the revision.

To clarify your concern, we are plotting more results from NCEP FNL analysis data. We would like to compare the evolution of synoptic-scale patterns (features) in the sole good member at a longer lead time (at 1800 UTC 4 Dec) and bad member with that in NCEP FNL analyses to better link the performance of CReSS runs back to the physics, and also better point out a new interesting result in this study. that is still possible to have good forecasts at lead time up to 5 days.

Besides, we also plot the sensitivity at t-36 to better identify the timing with higher predictability. Because we found a new interesting result (new finding) in section 3.2 that the synoptic pattern already developed into (or toward) what would cause the rainfall later at timing more than 24h earlier. We believe that these could explain the predictability of D18 event.

**Comment 02**. No hypotheses are presented in this work. This leads to the manuscript being unorganized, and the results unclear in the context of the broader literature. Having model simulations are not alone publishable. It is thus important to outline scientific based hypotheses in which the experiments in the manuscript are designed to evaluate, which will then make it clearer how the work adds to the body of literature.

**Reply:** As we replied to your comment 01 that we are reorganizing the manuscript to better explain the predictability of D18 event as well as the performance of the CReSS model. We also add more results in the revision, as replied to comment 01,

to have deeper look at interesting results (new findings), including (1) it is still possible to have good forecasts at longer lead time (up to 5 days) for record-break events like D18 event. (2) the development of synoptic patterns into (or toward) what could lead to an extreme rainfall later at timing earlier in the past. By these changes, we believe that our present study will have valuable contributions to the body of literature.

**Comment 03**. Ensemble spread is not purely error or a representation of accuracy. The goal of a well calibrated ensemble is to represent the forecast probability density function. Thus, if there is high uncertainty, we want the ensemble to have a large amount of spread. If there is small uncertainty in the system, we want the ensemble to have little spread. The usage of spread as an error metric needs to be done within this context.

a. It is also not clear to me where the spread analysis is undertaken within the paper.

**Reply:** We apologize for the confusion. In the time-lagged approach, only one run is executed at each initial time, so for each time it is not possible to derive the probability information. However, when successive runs are grouped together, one can see that the spread (e.g., standard deviation) still evolves with time just like a multi-member ensemble, as shown in Fig. 8. As the lead time shortened, the high SD region from the last several members became more focused along the coast, indicating high rainfall amounts (also in Fig. 9) there but with uncertainty in exact locations. We will direct our discussion more toward this direction, and better clarify the evolution of the spread in the revision.

b. Some discussion and framing of the work here from a context of intrinsic versus practical predictability is needed. Additionally, the scale dependence of predictability. I suggest Melhauser and Zhang (2012), Nielsen and Schumacher (2016), Weyn and Durran (2018), and citations within as starting points. There is also some useful suggestions from an ensemble analysis within these papers.

**Reply:** Thank you for your comments and your information. We will add more discussion (as mentioned above) and cite these references in the revision.

**Comment 04**. The results presented in this paper would have made an interesting section in Part 1 paper but because of these issues outline above it is not in the current state publishable on their own.

**Reply:** Thank you for the comment. As it is, the Part I paper has already been published, and we certainly would like to published the high-resolution time-lagged

QPF result on the D18 event as Part II, if at all possible. We will devote efforts to clarify the predictability, factors affecting the predictability, and the usefulness of the time-lagged ensemble in this event. We hope that our reply (including those stated above) and future revision will shed light on the above issues, be satisfactory, and make contributions to the existing literature on heavy-rainfall QPFs.

Additional Comments (Comments are not listed in order of importance):

1. Lines 21: typo “predicts”

**Reply:** Thank you for your comment. We corrected it.

2. Lines 53-55: The phrasing of this sentence is awkward. I recommend removing “until now” and adding something like “to improve predictability” to the end of the sentence.

**Reply:** Thank you for your suggestion. We updated it, as suggested.

3. Lines 85-87: Citations are needed to support this statement.

**Reply:** Thank you for your suggestion. We added the information, as suggested.

4. Lines 182-184: What version of the GFS?

**Reply:** we used the GFS version ds084.1. We also added the information in the revision for clarification.

5. Section 2.1.3: Is there any citation or information about the observational error associated with the gauges in this network?

**Reply:** Thank you for your question. The automated rain-gauge network in Vietnam is operated, maintained, and managed by Vietnam National Centre for Hydro-Meteorological Network (NCN). Therefore, the network meets the World Meteorological Organization standard. The observed dataset in this study is provided by NCN through the Mid-Central Regional Hydro Meteorological Center.

6. Section 2.1.4: A citation needs to be added for IMERG. Additionally, please let us know what version you used.

**Reply:** Thank you for your comment. We used the IMERG Final Run V07 data production. We also clarified this and added a citation, as suggested.

7. Line 247: Table 2 does not appear anywhere in the manuscript.

**Reply:** Thank you for your comments. We had removed that information. We created the table 2 in the first version of the manuscript. However, we had removed it from the submitted version of manuscript. But we were missed to remove the relevant information at line 247.

8. Lines 485-488, Lines 508-509: It is also possible the convective inflow to the storms is not at the surface but is elevated. Again, a much more detailed examination into the variables that control the ingredients for extreme rainfall is needed in the ensemble runs and ESA.

**Reply:** You are right. It is necessary to add more detailed examinations. Therefore, we will put more emphasis on the analysis along with adding new results to figure out the quantitative contribution of selected variables to the D18 event. We will let you see it in the revision.

9. Section 3.2: What does “per SD” mean?

**Reply:** We apologize for the confusion. It means per standard deviation. We will add more information along with quantitative values in the revision for clarification.