NHESS-2020-397
Authors' Responses to Reviewer 2 (RC2, anonymous)
Date: 10 Oct 2021

Title: Evaluation of Mei-yu Heavy-Rainfall Quantitative Precipitation Forecasts in Taiwan by A Cloud-Resolving Model for Three Seasons of 2012–2014
Authors: C.-C. Wang, P.-Y. Chuang, C.-S. Chang, K. Tsuboki, S.-Y. Huang, and G.-C. Leu

Reply:

The efforts and comments from this reviewer (Reviewer 2) are deeply appreciated, and the paper has been revised accordingly. In the revision (color-coded version), the changes made in response to Reviewer 2 and by ourselves (mostly minor changes in English) are marked in blue and orange, respectively. A point-by-point response to each of the comments from this reviewer are given below following their order. In each point, how and where the revision is made in the text is also specified.

1. General comments:

1) In the Abstract and Conclusions sections, again, the values of the TS are given with no comparison to the values of previous studies to clearly demonstrate the benefit of the CrESS simulation, e.g. L17, L472-L473. Although the TS values of previous studies are referenced in the introduction, the scores of TS should also be included in the Abstract and Conclusions to assert the real improvement of the CreSS simulations. Either write explicitly the values of TS from the previous experiment or provide the relative increase of TS in the CreSS simulations with respect to the previous papers.

Reply: Thank you for this suggestion. The TS values of previous studies are explicitly indicated (TSs < 0.1 at 100 mm and nearly zero at 250 mm and beyond) in both the Abstract and Section 6 (summary and conclusion) to indicate the improvement of the 2.5-km CReSS simulations, as suggested (L19, L461-462).

2) The concepts, score and skill of a model are confused in the manuscript. Scores (TS, POD, FAR) are a quantitative metric of the accuracy of a forecast. Skill of a model is the relative improvement of a QPF with respect to a reference value (most of the times a climatology). The metrics used for verification of the QPFs do not provide a measure of the skill of the model. All mentions to the "skill of the model" should be changed by mentions to the "accuracy of the model".

Reply: Thank you for this suggestion and we agree with it. In the revision, the word "skill" is replaced by "accuracy" (or "ability" or "performance" in some instances) to reference the categorical metrics such as the TS values (L53, L76, L86, L162, L191, L193, L198, L216, L229-231, L284, L334, L336, L409, L476), and retained only when the scores are compared to a reference to show improvement, as suggested.

3) In Fig. 12, an estimation of the number of items per weather type is introduced. While this graph clearly shows that large-scale systems (Meso-L, Through, Fronts) are associated with a larger number of items, this does not imply that "such conditions […] appear to allow for better model performance in QPFs". That particular aspect is not assessed. Either provide an assessment of the validation metrics (TS, POD, etc.) for each studied weather pattern (Fronts, Meso-L, Throughs etc.) or delete the figure 12 and lines L453 to L457, i.e., the complete sentence "The results (Fig. 12) … given a sufficient resolution".

Reply: Thank you for this suggestion. In the revision, the sentence is modified to say only what can be concluded from our analysis linked to Fig. 12, and is therefore revised to "… the synoptic and meso-α-scale conditions tend to be more favorable in larger events, which in general also correspond to higher TS values (Figs. 3 and 11) in combination with the orographic forcing in Taiwan", along the lines as suggested (L444-445).

2. Specific comments:

1) L109 – Normally at the resolutions of your simulations (2.5 km) shallow convection needs to be parameterized because the model cannot resolve it explicitly. What do you mean by "without any cumulus parameterization scheme"? Is there any active shallow convection scheme? Please provide an explanation in the manuscript.

Reply: In the revision, it is clarified that no cumulus (or shallow convection) parameterization is used in the CReSS model, as suggested (L109-110).

2) L200 & L279 – In atmospheric sciences, one can only say that a change is significant after performing a statistical test, for example the t-test. Either provide assessment of the significance of your changes or rephrase these sentences.

Reply: In the revision, the word "significantly" is changed to other words such as "clearly" or

"considerably", along the lines as suggested (L200, L276).

3) L211 – Please provide the physical/mathematical explanation suggested by Chien et al., (2002, 2006), Chien and Jou (2004) and Yang et al., (2004) that a value of TS >=0.15 is a threshold for an accurate forecast and include that information in the manuscript. Also please adapt the word skill for accuracy (see general comment).

Reply: The value of TS ≥ 0.15 was used in some previous studies based on experience (without physical explanation), and we adopt the same criterion here. Because this standard may be somewhat arbitrary, the sentence is revised to "… if we select TS ≥ 0.15 to indicate some level of accuracy…", to soften the tone, following the suggestion (L211).

4) L232 – How come that at 130 mm a TS=0.07 implies a good forecast, if the threshold for a good forecast is TS>=0.15? Is it because the TS threshold varies depending on the analysed precipitation intensity? Also please adapt the word skill for accuracy (see general comment).

Reply: In the revision, this sentence is modified to make no reference to the accuracy of the forecast, as suggested (L231-232).

3. Writting comments:

L32 – Should read: "Quantitative Precipitation Forecasting (QPF) …"
Reply: Revised as suggested (L31).

L35 – Should read: "… rather frequently, mainly during …"
Reply: Revised as suggested (L34).

L36 – Delete "when"
Reply: In the revision, the long sentence is broken into shorter sentences, and revised to "…Chang et al., 2013). The landslides and …" for better readability, similar to as suggested (L35).

L39 – Delete: "model"
Reply: Deleted as suggested (L39).

L57-L67 – Sentence is too long. Split into several sentences

Reply: In the revision, the long sentences are broken into shorter sentences, as suggested (L56-61).

L64: Should read: "… mean (WEPS) for thresholds between 50-200 mm …"

Reply: Revised as suggested (L64).

L72 – Should read: "… more comparable to research studies"

Reply: Revised as suggested (L72).

L74 – Should read: "… forecasts showed …"

Reply: Revised as suggested (L74).

L76 – Should read: "… is remarkably higher for typhoon …"

Reply: Revised as suggested (L76-77).

L89 – Delete complete sentence: "As none of these above questions … our objectives"

Reply: The sentence is deleted as suggested (L89).

L103 – Change "without nesting" for "without intermediate nesting"

Reply: Changed as suggested (L103).

L206 – Delete: "and the second question in our objectives is answered"

Reply: Deleted as suggested (L206).

L208 – Delete: "for example"

Reply: Deleted as suggested (L208).

L220 – Change: "at times, or do not drop at all" for "for"

Reply: Changed as suggested (L220).

L262-L264 – Delete complete sentence: "Thus, as exemplified … but small (unimportant) events."

Reply: Deleted as suggested (L261).

Figure 4 – Delete "rounded to two decimal places"

Reply: Deleted as suggested (L267).

L279-L280 – Delete: “Thus, the first objective of this study is fulfilled”

Reply: Deleted as suggested (L276).

L285 – Delete: “shown above”

Reply: Deleted as suggested (L282).

L286 – Delete: “for further examination and discussion”

Reply: Deleted as suggested (L283).

L287 -Delete: “with the corresponding scores” and “in detail”

Reply: Both deleted as suggested (L284).

L288-L289 -Delete: “thereby to shed light on the source of skill seen in Fig. 4”

Reply: Deleted as suggested (L285).

L292 -Delete: “and the model’s performance in predicting this event is thus highly relevant”

Reply: Deleted as suggested (L288).

L295 -Delete: “Reminiscent to the season average (cf Fig.3)”

Reply: Deleted as suggested (L290).

L313 -Delete: “in the same column”

Reply: Deleted as suggested (L308).

L314 – Change “lengthy” for “long-lasting”

Reply: Changed as suggested (L309).

L319 – Change “… somewhat too early …” for “ … somewhat earlier …” and “with apparent” for “showing”

Reply: Both instances changed as suggested (L314).

L322 – Delete “however”

Reply: Deleted as suggested (L316).

L324 – Change “Compare” for “Compared”

Reply: Corrected as suggested (L319).

L328 – Change “observation” for “observations”

Reply: Changed as suggested (L322).

L328 – Delete “must”

Reply: Deleted as suggested (L323).

L329 – Delete “in this event”.

Reply: Deleted as suggested (L323).

L334 – L336 – Delete complete sentence: “With such information … can also be verified here”.

Reply: The sentence is deleted, as suggested (L328).

L337 – Change “commencement” for “beginning”

Reply: Changed as suggested (L328).

L346 – Start new paragraph at “In Fig. 8”.

Reply: Revised as suggested (L338-344).

Figure 7 - Description is unintelligible. Please rewrite. Use points and split into shorter sentences.

Reply: Rewritten following the suggestion (L346-350).

L361-362– Change “…, while the dependency on event magnitude also exists, the QPFs… ” for “…,the dependency on event magnitude exists, but the QPFs… ”

Reply: Changed as suggested (L351).

L362 – L363 – Delete “as mentioned. All of good forecast quality (cf Figs 6ab, columns 4-7)”

Reply: Deleted as suggested (L352).

L370 – Change “lone” for “only”

Reply: Changed as suggested (L360).

L413-L414 – Delete complete sentence “Thus, through … to a certain degree.”

Reply: Deleted as suggested (L403).

L471 – Delete: “In this section, … remarks are given”.

Reply: Deleted as suggested (L459).

L476 – L477 – Rephrase the sentence as “The ability to represent the extreme and top events (group A+) in terms of the TS are much higher when a proper classification based on observed rain area size (i.e., event magnitude) is used.”

Reply: Revised as suggested (L464-465).

L478 – Delete “at the same thresholds”

Reply: Deleted as suggested (L466).

L482 – Change “Through further analysis of example case, the improvement…” for “For a selected case study, the improvement…”

Reply: Changed as suggested (L470).

L484 – Delete “stationary”

Reply: Deleted as suggested (L471-472).

L485-L486 – Change “the concentrated rainfall” for “the accuracy of QPFs for concentrated rainfall”.

Reply: Changed as suggested (L473-474).

L486 – Change “… squall lines) is highly random and more difficult to predict…” for “… squall lines) could not be demonstrated probably due to the difficulty to predict…”

Reply: Changed as suggested (L474-475).

L 487-L488 – Delete “ and the potential of heavy rainfall indicated (unless at short ranges). For such rainfall model QPFs are still informative and useful, but the categorical scores may not be.”

Reply: Deleted as suggested (L476).

L489 – Change “may possess” for “showed a”.

Reply: Changed as suggested (L476).

L490 – Change “skill” for “accuracy” (see general comment).

Reply: Changed as suggested (L476).

L 491- Change “ordinary ones at least for regions like” for “coarser resolution models over the orographic region of Taiwan”

Reply: Changed as suggested (L477).