Responses to Reviewer #1

Point #1

Review to Yang et ., 2021, Bias-correcting individual inputs prior to combined calibration leads to more skillful forecasts of reference crop evapotranspiration. HESSD.

In this study, the authors investigated a critical issue in the forecasting of short-term reference crop evapotranspiration (ETo) based on NWP outputs. It is getting popular that weather forecasts from NWP models are used to predict water loss through evapotranspiration. Such information is highly valuable for the effective management of water resources, particularly in arid/semi-arid regions. This investigation develops a new methodology that effectively corrects errors in ETo forecasts, and adds extra skills to statistical calibration. I believe this new post-processing strategy could benefit future NWP-based ETo forecasting. To improve this work, the authors should pay special attention to the following key issues:

Response: We appreciate the reviewer's insightful comments. We also believe the findings of this work could contribute to improving future NWP-based ETo forecasting. We address your constructive comments thoroughly and carefully and believe this work has been improved significantly. Please find more details in our point-by-point response.

Point #2

1, Presentation of the results could be improved. Currently, the authors use maps to show/compare results from different model experiments. These figures could demonstrate the spatial patterns of modeling results. However, it might be more useful if the authors could summarize regional results in a different way, such as using boxplots. I believe that will better show readers the overall statistical information across the whole country than simply plotting the results as maps.

Response: Thank you for the valuable suggestions. We create boxplots for all the maps shown in the main text. Since we already have eight figures in the main text and seven figures in the supplementary material, we think it is better not to add too many new figures. We combine these new boxplots with maps for Figures 2-7, which have extra zoom for adding new subplots. For Figures 1 and 7, which already include many subplots, we present the corresponding boxplots in the Supplementary Material. Please find the boxplots as follows:
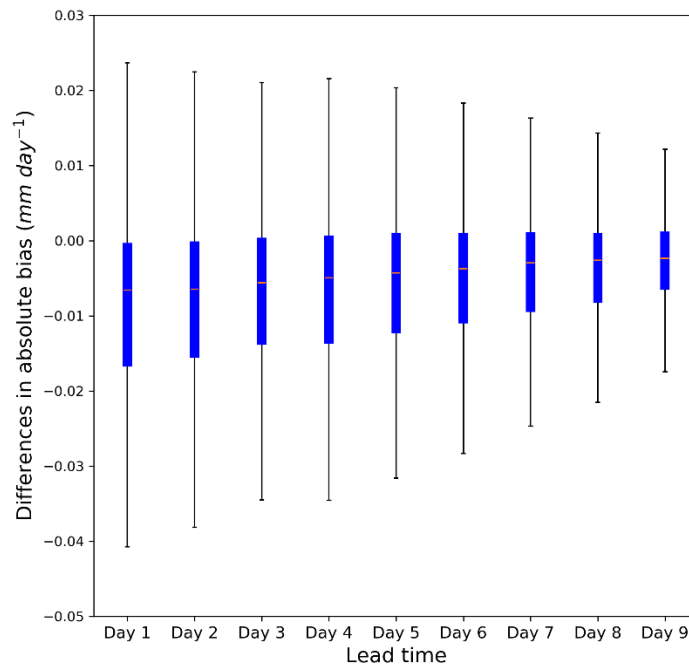
34
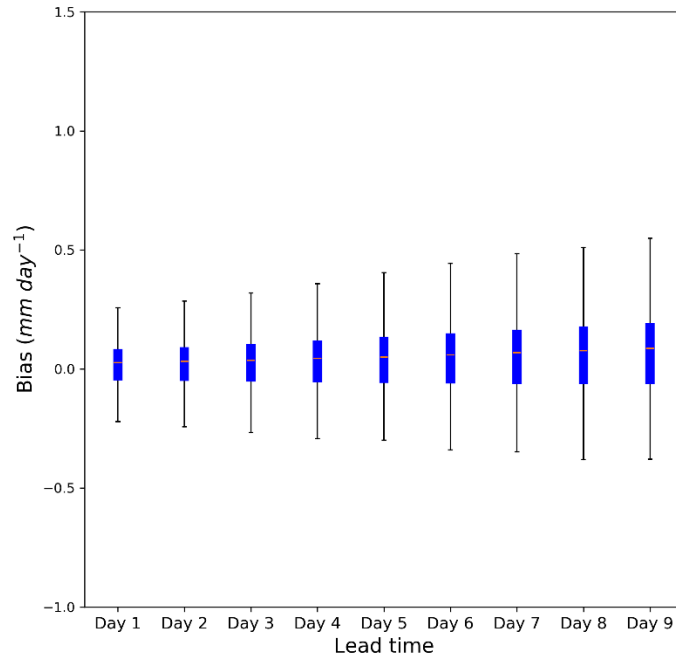
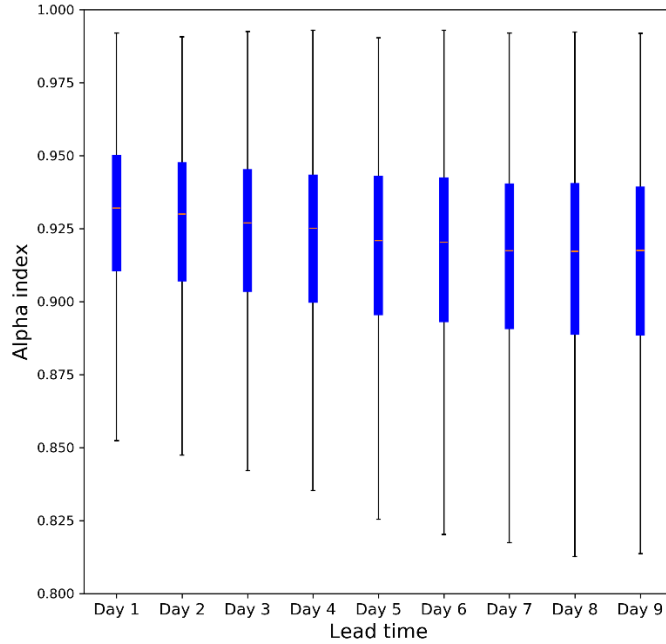Figure 2 Boxplot summarizing bias in calibrated ETo forecasts

36

37

Figure 3 Boxplot summarizing differences in absolute bias between calibrated ETo forecasts from Calibration 2 with Calibration 1
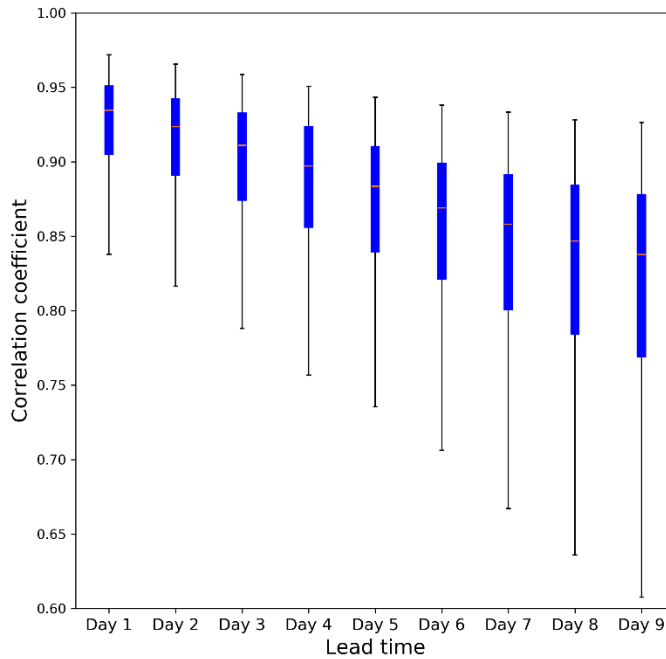
40

41

42

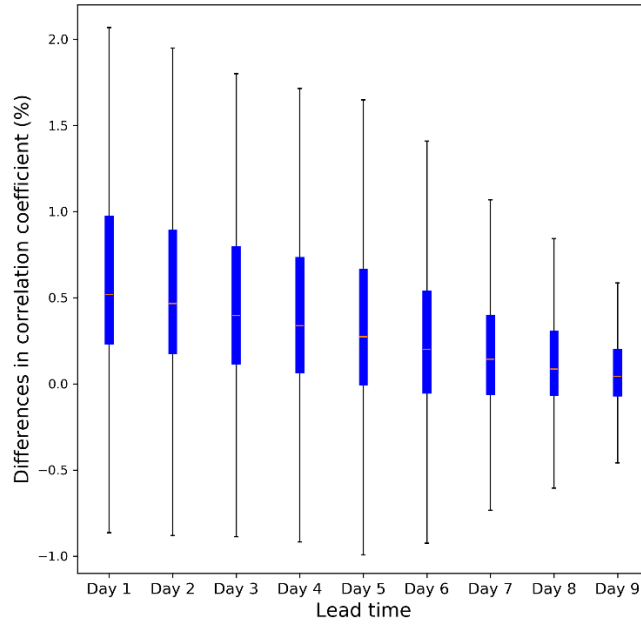**Figure 4 Boxplot summarizing the alpha index in the calibrated ETo forecasts**

44

Figure 5 Boxplot summarizing correlation coefficient between calibrated ETo forecasts from Calibration 2 and AWAP ETo data
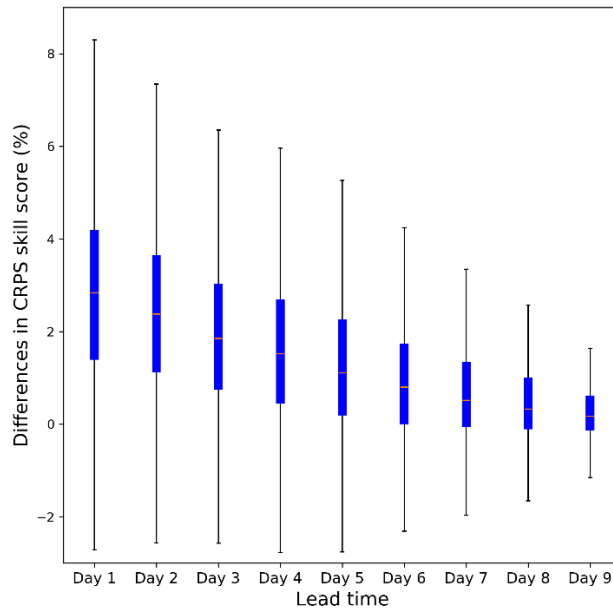
47

48

3

49

Figure 6 Boxplot summarizing differences in the correlation coefficient (calibrated forecasts vs. AWAP ETo) between Calibrations 2 and 1

52

53

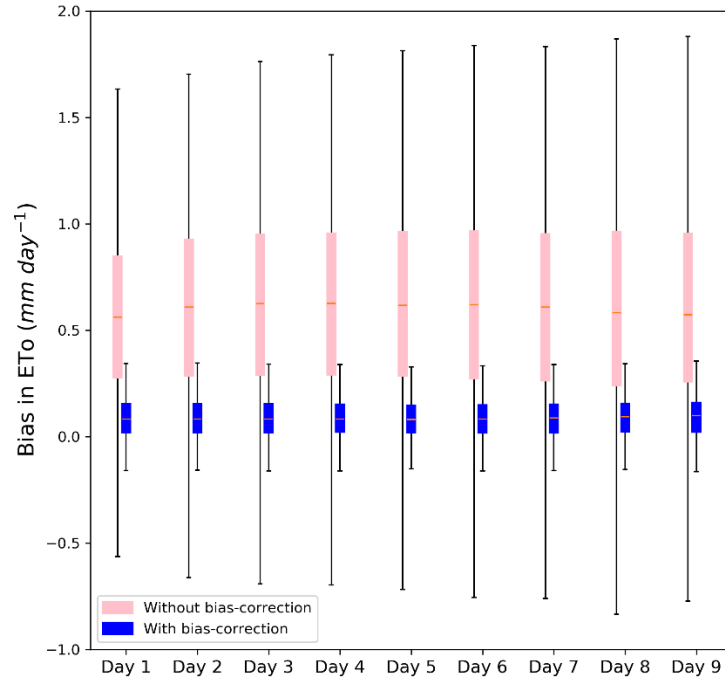**Figure 8 Boxplot summarizing differences in CRPS skill scores between the calibrated forecast from Calibration 2 with those from Calibration 1**
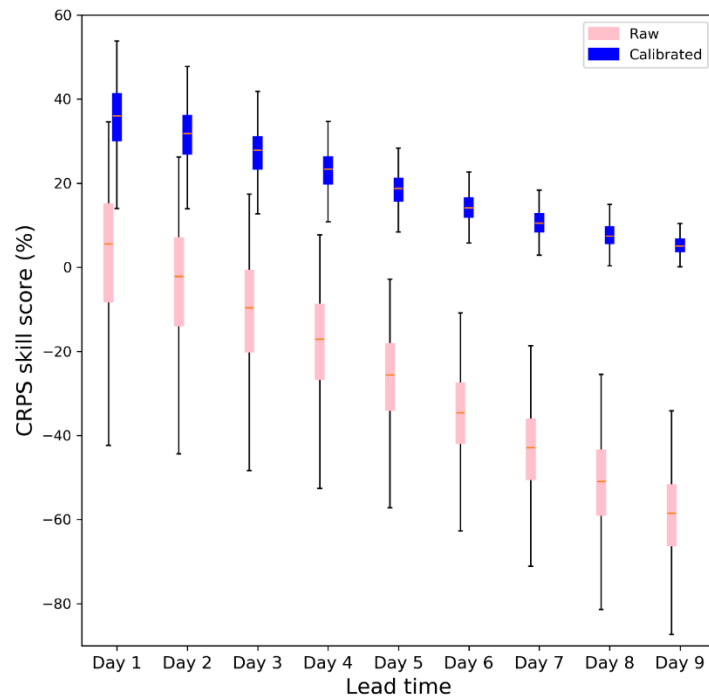
56

Figure S7. Boxplot of biases in raw ETo forecasts constructed without bias-corrected input variables (pink) and correct inputs (blue)

59

60

Figure S9. Boxplot of CRPS skill score in raw (pink) and calibrated ETo forecasts (blue) from Calibration 2

Point #3

2, Implications for ETo forecasting at the monthly or seasonal scales should be further discussed. ETo forecasting based on monthly or seasonal climate forecasts from GCMs is also widely performed. This study develops the new strategy for short-term forecasts. The applicability of this method to ETo forecasting based on GCM forecasts should be briefly discussed, to benefit a broader range of readers.

Response: We agree with the reviewer that ETo forecasting with longer forecast horizons (e.g., monthly and seasonal) based on GCM forecasts is increasingly performed, and it is necessary to evaluate whether the post-processing strategy developed in this investigation is applicable to the GCM-based seasonal ETo forecasting. As we have shown in this manuscript, the reduction of error propagation from the input variables to ETo is the key reason why the new strategy has better performance than the original strategy (no improvement to raw forecasts of input variables). We expect this will be the case for GCM-based seasonal forecasts. However, testing this idea will be beyond the scope of this current study. To highlight the necessity of adopting this strategy in seasonal ETo forecasting, we add the following paragraph to section 4.2 (Implications for forecasting of integrated variables and future work):

" In addition, seasonal ETo forecasting based on GCM climate forecast has been increasingly performed (Tian et al., 2014; Zhao et al., 2019b). In these investigations, raw ETo forecasts were also constructed with raw GCM forecasts. As a result, it is unavoidable that these investigations have suffered from error propagation from input variables to seasonal ETo forecasts. We expect that the calibration strategy (strategy ii) tested in this study will be applicable to seasonal ETo forecasting, considering its capability in reducing errors that could not be corrected through statistical calibration. Further investigations are needed to examine how the bias-correction of raw forecasts of input variables will affect the calibration of GCM-based seasonal ETo forecasts."

Point #4

Specific comments:

Line 20, rewrite this sentence. Not clear

Response: we replace the original sentence:

"This calibration strategy is expected to enhance future NWP-based ETo forecasting."

with

" We anticipate that future NWP-based ETo forecasting will benefit from adopting the calibration strategy developed in this study to produce more skillful ETo forecasts. "

Point #5

Line 74 Calibrate->calibrate

99   **Response: We correct the word accordingly.**

100

101   ## Point #6
102   Line 80 compiled as the inputs…..

103   **Response: We improve the sentence of:**

104   "Weather forecasts from the ACCESS-G2 model are compiled to generate ETo forecasts."

105   **with:**

106   "Weather forecasts from the ACCESS-G2 model are extracted as inputs for the calculation of ETo
107   forecasts."

108

109   ## Point #7
110   Line 95 10m -> 10 m.

111   **Response: We add a space between the number and the unit. We also check the entire**
112   **manuscript to correct this issue.**

113

114   ## Point #8
115   Line 107-108, need to clarify what the anomaly and climatological mean are referring to

116   **Response: To clarify how the anomaly and climatological mean are derived, we replace the**
117   **sentence:**

118   "Our recent investigation suggests that ETo forecast calibration based on anomaly and climatological
119   mean produces more skillful calibrated forecasts than calibrating ETo forecasts directly."

120   **with:**

121   "Our recent investigation suggests calibrating anomalies of raw forecasts, which are calculated as the
122   departure from the observed climatological mean, could produce more skillful calibrated forecasts than
123   calibrating ETo forecasts directly."

124

125   ## Point #9
126   Line 165 consider rewriting this sentence. Does not read well.

127   **Response: We replace the original sentence of**

128   "Once we obtain all the parameters for the BN distribution (equation 4), a conditional distribution is

129   established for $o(t)$ when a raw forecast ($f(t)$) is provided."

130   **with:**

131   "With the optimized parameters (means, standard deviations, and correlations) for the BN distribution

132   (equation 4), a conditional distribution for o(t) for a given raw forecast (f(t)) is derived "

133

134   ## Point #10

135   Line 172, what is specific month

136   **Response: we replace "specific" with "unselected" to make the wording more specific.**

137

138   ## Point #11

139   Figures in Results: shouldn't the figures be centralized?

140   **Response: The original format following a template from HESS. After we add boxplots to**

141   **these maps, the empty space for each figure is significantly reduced. We keep them aligned**

142   **to the left to be consistent with the provided template.**

143

144   ## Point #12

145   Line 360, not calibrate directly, should be without correcting forecasts of the inputs

146   **Response: Thank you for the suggestion. The key message we want to present here is that**

147   **statistical models may not be able to correct all errors in integrated variables (such as ETo).**

148   **However, when the input variables are corrected first, error propagation from inputs to**

149   **integrated variables, particularly for the errors which could not be corrected by calibration**

150   **models, will be reduced. To make it clear, we improved the original sentence of:**

151   **"**Our investigation suggests that improving the input variables may help correct errors that could not be

152   fixed when calibrating the integrated variables directly."

153   **with:**

154   **"** Our investigation suggests that improving the input variables could reduce error propagation from

155   inputs to integrated variables, and particularly reduce errors that could not be corrected by the calibration

156   model. "

157

158

Point #13

Line 365, consider rewriting this sentence

Response: Thank you for the suggestion. We replace the original sentence:

"As a result, using a more sophisticated calibration method to correct errors in input variables, is expected to further improve forecasts of these input variables, resulting in more significant improvements in the final calibrated ETo forecasts"

with:

" If a more sophisticated calibration method is employed to correct errors in input variables, error propagation from input variables to ETo forecasts will likely be further reduced. As a result, we anticipate that the calibrated ETo forecast will gain further improvements in forecast skills."

Point #14

Line 377-378, two' calibration models' consider to rewrite

Response: We improve the original sentence:

"Additional investigations using other calibration models will help clarify whether the improvements will hold for other calibration models."

With

" Additional evaluations using other calibration models will be needed to ascertain whether the improvements will be achieved when the calibration is conducted with a different model. "

Point #15

Line 385, in the calibrated forecasts

Response: We add the missing 'in' to this sentence.

Point #16

Line 386, consider making it shorter and clearer

Response: We improve the following sentence:

"Further investigation indicates that the contribution of improving input variables to the ETo forecasting tends to be independent of the calibration method applied to raw ETo forecasts."

With

189   "Further investigation indicates that the improvements tend to be independent of the calibration method
190   applied to raw ETo forecasts."