Using Machine Learning to Improve Sub-Seasonal Climate Prediction

Richard Garmong[1],
Becky Bolinger[2], Russ Schumacher[2]
[1]University Of Georgia
[2]Colorado State University

Introduction

- Sub-Seasonal (SS) Climate prediction is important given the expected increase in climate variability with global warming.
- One vulnerable region to climate variability is the Upper Colorado River Basin (UCRB).
- The Climate Prediction Center's (CPC) outlook products are the most widely used in SS prediction.
- North American Multi-Model Ensemble (NMME) is a recent advancement in SS prediction.
- Current SS prediction lacks skill, especially in the UCRB.

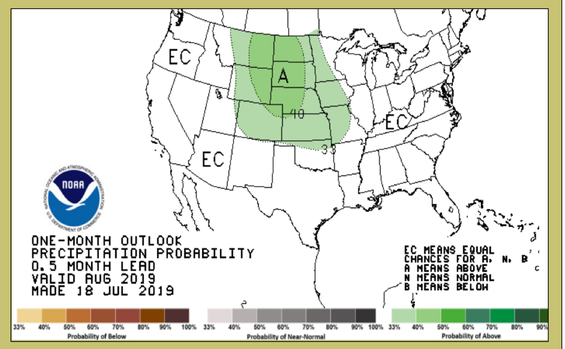


Figure 1. Current CPC next month precipitation forecast

Data and Methods

- The type of machine learning chosen was the Random Forest Classifier
- The dynamical model inputs were mapped to the Upper Colorado river basins
- A separate model was created for each sub-basin.
- Trained and tested upon a 1982-2010 dataset, cross verified on a 2017-2019 dataset
- Classification is based upon 3 categories of dry, wet and normal
- Forecasts were created for Lead .5 forecast based upon previous month's data

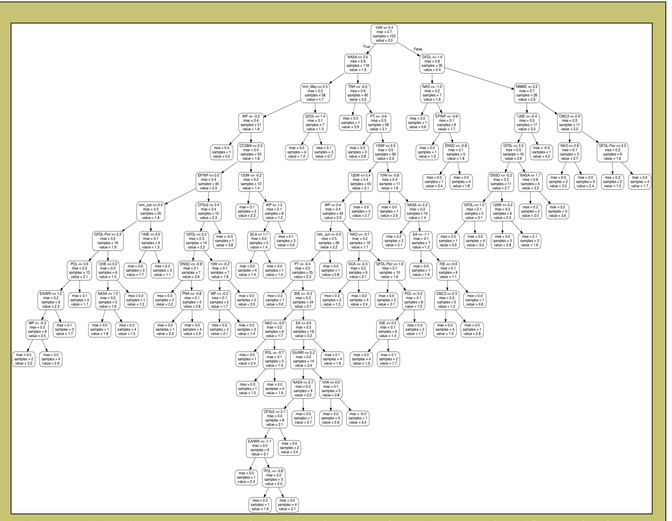


Figure 2. Decision tree used within the Random Forest Classifier, which uses hundreds of similar trees to create an ensemble that chooses the category

Random Forest Model inputs

- NMME suite (NMME, CFSv2, GFDL 2.1, etc.)
- Teleconnections (NAO, PNA, EA, etc.)
- Multivariate ENSO index
- Madden-Julian Oscillation
- Previous Month Precipitation
- Forecast Month

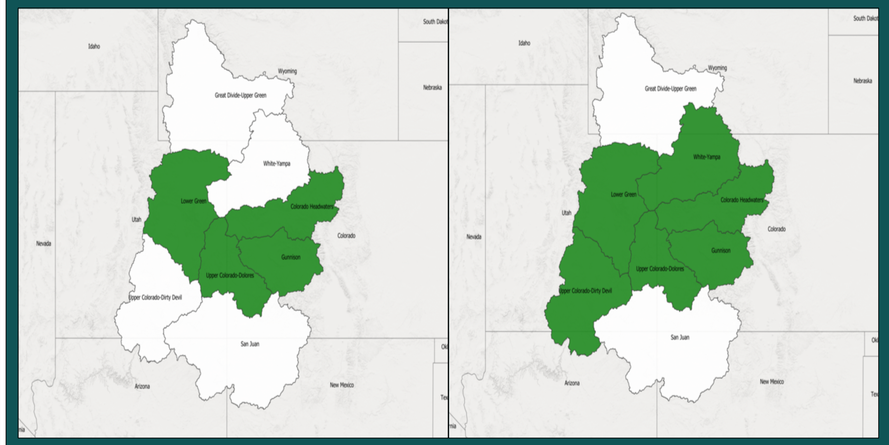


Figure 3. Comparison of model forecast (left) versus verification (right) for October 2018 . Green is for areas that are wetter than average, red is for areas that are dryer than average, and white indicates a normal amount of precipitation

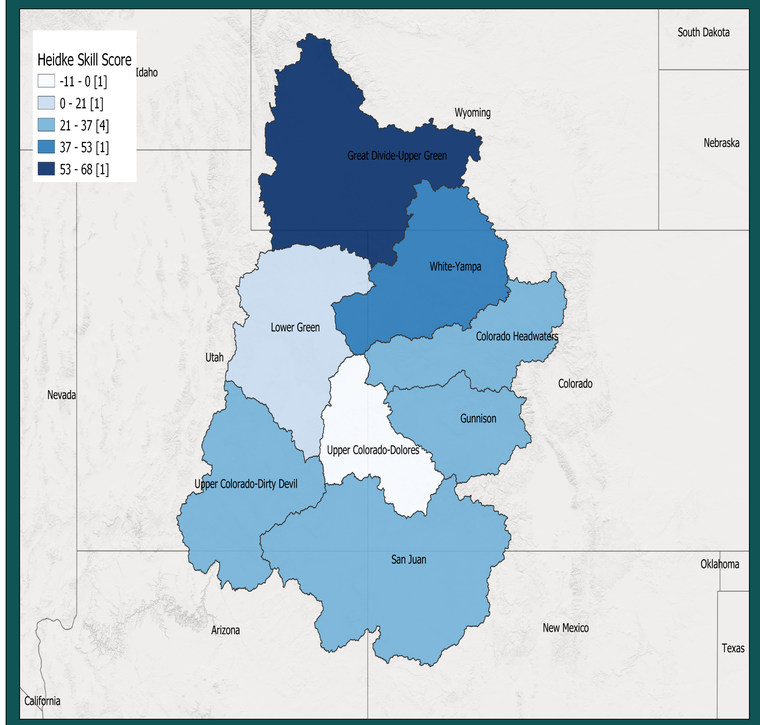
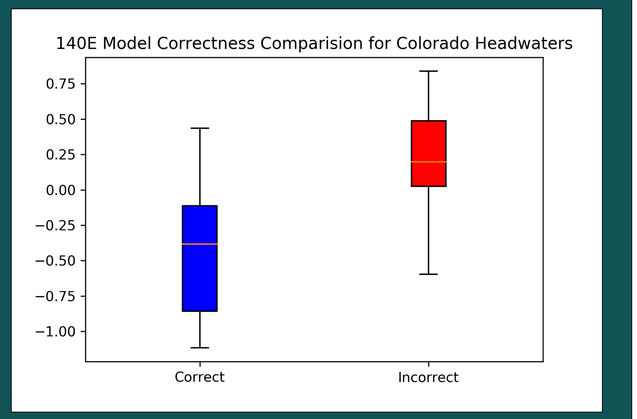


Figure 4. Heidke Skill Score across the test period. Areas that are white indicate that the model has no increased skill over random chance



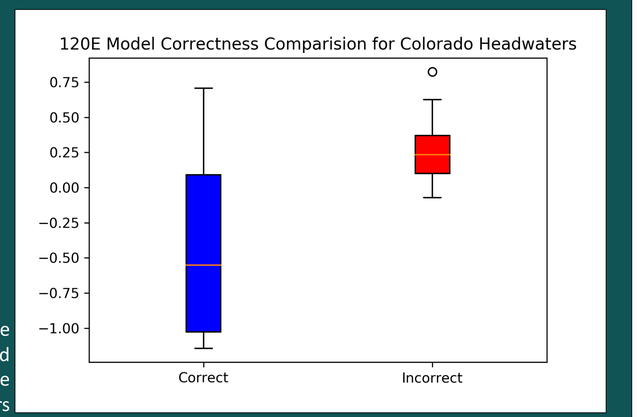


Figure 5 & 6. Difference in MJO phase associated with model skill in the Colorado Headwaters

Results and Discussion

- All models had skill above random error, with the exception of the Upper Colorado-Dolores Sub-basin
- Results suggest better skill than CPC forecasts
- Qualitatively, October 2018 is noteworthy due to the model identifying the relief from a period of extreme drought.
- The model failed to identify the dry event that occurred across every basin, the previous month.
- Seems to have a slight wet bias, potentially due to climate change?
- MJO phase is related to model skill

Acknowledgements This work has been supported by the National Science Foundation Research Experiences for Undergraduates Site in Climate Science at Colorado State University under the cooperative agreement No. AGS-1461270. Special thanks to Sarah Baker and Andy Wood.

Conclusion This study looked to determine if machine learning could improve prediction within the Upper Colorado River Basin. Machine learning has additional skill over the dynamical models alone and seems to have at least comparable skill to human forecasts. Model suggests that MJO phase is associated with increased forecast skill, however sample size is small. Future work includes using the Random Forest probability to provide a more "apples to apples" forecast with the CPC, as well as further expanding the basins explored.

Contact info: richard.garmong@uga.edu