

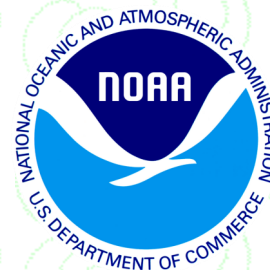
46th NOAA Climate Diagnostics and Prediction Workshop
Special Issue



Virtual Online

Climate Prediction S&T Digest

<https://doi.org/10.25923/rj6c-rk11>



July 2022

NWS Science & Technology Infusion Climate Bulletin Supplement

Inside this issue:

1. Subseasonal to Seasonal Precipitation Prediction
2. Extremes and Extreme Events
3. Applications of Modern Technologies
4. Development and Use of Climate Data Records

NOAA's National Weather Service

Office of Science and Technology
Integration

1325 East West Highway
Silver Spring, MD 20910

Climate Prediction Center
5830 University Research Court
College Park, MD 20740

Article Citation:

Author(s), 2022: Article title. Extended Summary, *Climate Prediction S&T Digest*, 46th NOAA Climate Diagnostics and Prediction Workshop, Virtual Online, DOC/NOAA, page range.
DOI: 10.25923/rj6c-rk11

Although the skill of current operational climate prediction is limited and the research on the topic presents many challenges, there are promises of improvement on the horizon. To accelerate advancement in climate services, an effective mechanism of S&T infusion from research to operation for application is much needed. This bulletin has been established to clarify science-related problems and relevant issues identified in operation, inviting our partners in the research community to work together on improvement of national climate prediction services.

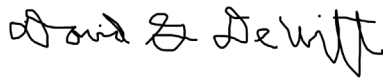
Science and Technology Infusion Climate Bulletin
<https://vlab.noaa.gov/web/osti-r2o/climate/sti-climate-bulletin>

National Weather Service
National Oceanic and Atmospheric Administration
U.S. Department of Commerce

PREFACE

It is with great pleasure that the Climate Prediction Center (CPC) and the Office of Science and Technology Integration (STI) offer you this synthesis of the 46th Climate Diagnostics and Prediction Workshop (CDPW). The CDPW remains a must attend workshop for the climate monitoring and prediction community. As is clearly evident in this digest, considerable progress is being made both in our ability to monitor and predict climate. The purpose of this digest is to ensure that climate research advances are shared with the broader community and also transitioned into operations. This is especially important as NOAA works to enhance climate services both across the agency and with external partners. We hope you find this digest to be useful and stimulating. And please drop me a note if you have suggestions to improve the digest.

I would like to thank Dr. Jiayu Zhou of the Office of Science and Technology Integration, for developing the digest concept and seeing it through to completion. This partnership between STI and CPC is an essential element of NOAA climate services.

A handwritten signature in black ink, reading "David G. DeWitt". The signature is fluid and cursive, with the first name "David" and last name "DeWitt" clearly legible.

David G. DeWitt
Director, Climate Prediction Center
National Centers for Environmental Prediction
NOAA's National Weather Service

CONTENTS

OVERVIEW	1
1 SUBSEASONAL-TO-SEASONAL PREDICTION	3
Do models generate realistic simulations of the North Atlantic SST? <i>Timothy DelSole, and Michael K. Tippett</i>	4
Preconditions for extreme wet winters over the Contiguous United States <i>Andrew Hoell, Martin Hoerling, Jon Eischeid, and Joseph Barsugli</i>	6
MJO impacts on winter weather event frequency <i>Stephen Foskey, and Naoko Sakaeda</i>	8
Prediction challenges associated with errors in linear trends of tropical Pacific sea surface temperature <i>Michelle L. L'Heureux, Michael K. Tippett, and Wanqiu Wang</i>	12
Skillful long-lead prediction of summertime heavy rainfall in the US Midwest from ocean salinity <i>Laifang Li, Raymond W. Schmitt, and Caroline C. Ummenhofer</i>	16
Evaluating the potential of a blocking predictor in a hybridized dynamical-statistical model for improved week 3-4 temperature and precipitation outlooks <i>Cory F. Baggett, Laura Ciasto, Daniel Harnos, Emerson LaJoie, Dan Collins, Muthuvel Chelliah, Gregory Jennrich, Evan Oswald, Kyle MacRitchie, Jon Gottschalck, and Michael Halpert</i>	21
Dynamical weighting of the week 3-4 models based on forecasts of opportunity <i>Cory Baggett, Emerson LaJoie, Daniel Collins, Johnna Infanti, and Peitao Peng</i>	27
Evaluation of subseasonal Arctic sea ice hindcasts in an NCEP's UFS-based system <i>Yanyun Liu, Wanqiu Wang, Weiyu Yang, Jieshun Zhu, Arun Kumar, and David DeWitt</i>	32
Regional and global climate drivers of marine heat waves and related atmosphere-ocean anomalies in the Eastern North Pacific <i>Mitch Porter, Katie Kohlman, Seth Madden, and Tom Murphree</i>	36
Mechanism of the centennial subpolar North Atlantic cooling trend in the FGOALS-g2 historical simulation <i>Yifei Fan, Jianhua Lu, and Laifang Li</i>	39
Initialized and uninitialized ENSO predictability in year 2+ <i>Nathan Lenssen, Lisa Goddard, Simon Mason, and Yochanan Kushnir</i>	41
Seasonal tropical-extratropical teleconnections originating from tropical rainfall modes beyond canonical ENSO for northern winters <i>Peitao Peng, Wanqiu Wang and Arun Kumar</i>	46

2 EXTREMES AND EXTREME EVENTS	53
Global seasonal forecasts of marine heatwaves	
<i>Michael G. Jacox, Michael A. Alexander, Dillon Amaya, Emily Becker, Steven J. Bograd, Stephanie Brodie, Elliott L. Hazen, Mercedes Pozo Buil, and Desiree Tommasi</i>	54
Was the February 2021 cold air outbreak over the central U.S. a subseasonal forecast of opportunity?	
<i>John R. Albers, Matthew Newman, Andrew Hoell, Sam Lillo, Yan Wang, and Melissa Breeden</i>	56
A diagnosis of 2018/19 and 2019/20 winter CPC outlooks and model forecasts	58
<i>Arun Kumar, Zeng-Zhen Hu, Bhaskar Jha, and Mingyue Chen</i>	
The 2020-21 Southwestern U.S. drought: A preliminary study on its causes and prediction	
<i>Hailan Wang, Arun Kumar, Brad Pugh, Mingyue Chen, and Bala Narapusetty</i>	62
Impact of the MJO on the forecast skill of week-2 severe weather over the United States	68
<i>Hui Wang, Arun Kumar, Wanqiu Wang, David DeWitt, and Matthew Rosencrans</i>	
Fewer troughs, not more ridges, have led to a drying trend in the Western United States	
<i>Wei Zhang, Vittal Hari, Simon S-Y Wang, Matthew D. LaPlante, Gregg Garfin, Grace Affram, and Rohini Kumar</i>	73
Predictability of summer monsoon extreme rainfall events over Taiwan using NCEP GEFSv12 reforecast	76
<i>M. M. Nageswararao, Yuejian Zhu, and Vijay Tallapragada</i>	
Sea-surface temperatures and vertical wind shear as precursors to tropical cyclone activity in the Caribbean and an expanding Main Developing Region	83
<i>Keneshia Hibbert, Thomas M. Smith, Equisha Glenn, and Jorge E. Gonzalez</i>	
3 APPLICATIONS OF MODERN TECHNOLOGIES	89
Deep learning for subseasonal precipitation and temperature errors	
<i>Maria J. Molina, Jadwiga Richter, Judith Berner, Anne A. Glanville, Katie Dagon, Abby Jaye, Aixue Hu, and Gerald Meehl</i>	90
Ensemble predictability of week 3/4 precipitation and temperature over the United States via cluster analysis of the large-scale circulation	94
<i>Greg Jennrich, David Straus, Muthuvel Chelliah, and Cory Baggett</i>	
Meta-heuristic ant colony optimization technique to forecast the amount of summer monsoon rainfall: Skill comparison with Markov chain model	99
<i>Sayantika Mukherjee</i>	
Enhancing subseasonal temperature prediction by bridging a statistical model with dynamical Arctic Oscillation forecasting	100
<i>Changhyun Yoo, Minju Kim, and Jung Choi</i>	

4 DEVELOPMENT AND USE OF CLIMATE DATA RECORDS	101
U.S. climatological standard normals: A utilitarian workhorse	102
<i>Mike Palecki</i>	
Communicating uncertainty in SST analysis	
<i>Caihong Wen, Arun Kumar, Wanqiu Wang, Michelle L'Heureux, Pingping Xie, Zeng-Zhen Hu, and Bert Katz</i>	106
Updating the CPC T2M observational verification dataset and impact on the seasonal T2M GPRA	114
<i>Mike Halpert, David DeWitt, and Jon Gottschalck</i>	
Understanding US Drought in Past 120+ Years	116
<i>Yun Fan</i>	
Developing an experimental week 3-4 storm track outlook over North Pacific, North America, and North Atlantic	123
<i>Yutong Pan, Wanqiu Wang, Edmund K.M. Chang, Hui Wang, and David DeWitt</i>	
CPC's climate assessment database version 2	127
<i>Melissa Ou, Tom Collow, Donald Garrett</i>	
Moisture based agroclimate indices across the Canadian Prairies under a changing climate	133
<i>Aston Chipanshi, Dongzhi Qi, Yinsuo Zhang, and Mark Berry</i>	
Extended range verification using economic value	139
<i>Eva Regnier, and Joel Feldmeier</i>	

OVERVIEW

Due to the COVID-19 pandemic NOAA's 46th Climate Diagnostics and Prediction Workshop was held online on 26–28 October 2021. The workshop was hosted by the Climate Prediction Center (CPC) of the National Centers for Environmental Prediction (NCEP) and the Climate Services Branch (CSB) of the National Weather Service (NWS) Headquarters.

The workshop focused on four major themes, with an emphasis on seasonal and subseasonal prediction, monitoring, attribution, diagnostics, and service delivery related to:

1. The development and use of climate data records, including the impacts of updating base period climate normals on defining climatological base periods, verification, and trends. Topics also included the need for additional observational databases and improvements to existing databases, such as snowfall;
2. Prediction, monitoring, and verification of extremes and extreme events related to the surface, ocean, atmosphere, land, and ice;
3. Applications of modern technologies including geographic information system (GIS), machine learning, and software development at Subseasonal-to-Seasonal (S2S) time scales;
4. Improving S2S precipitation prediction capabilities, sources of predictability, and user engagement practices for application in water resources.

The workshop featured oral and poster presentations, lightning talks, invited speakers, and group discussions. This Digest volume is a collection of extended summaries of the presentations contributed by participants.

The workshop is continuing to grow and is expected to provide a stimulus for further improvements in climate monitoring, diagnostics, prediction, applications and services.



1. Subseasonal-to- Seasonal Prediction



Do Models Generate Realistic Simulations of the North Atlantic SST?

Timothy DelSole¹ and Michael K. Tippett²

¹Department of Atmospheric, Oceanic, and Earth Sciences, George Mason University, Fairfax, VA

²Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY

ABSTRACT

After a new climate model is constructed, a natural question is whether it generates realistic simulations. Here, "realistic" does not mean that the detailed patterns on a particular day are correct, but rather that the statistics over many years are realistic. Past approaches to answering this question often neglect correlations in space and time. This paper proposes a method for answering this question that accounts for correlations in space and time. The basic idea is to fit each multivariate time series to a Vector Autoregressive (VAR) Model, and then test the hypothesis that the parameters of the two models are equal. Because a VAR model uniquely specifies the autocorrelation function and the power spectra, equality of VAR models implies equality of autocorrelation functions and equality of power spectra. In the special case of a first-order VAR model, the model is a Linear Inverse Model (LIM) and the test constitutes a difference-in-LIM test. The likelihood ratio test for this problem and the associated sampling distributions are derived. This derivation leads to a deviance statistic that measures the difference between VAR processes and can be used to rank models based on their "closeness" to the VAR process inferred from observations. This test is applied to decide if climate models generate realistic internal variability of annual mean North Atlantic Sea Surface Temperature (NASST). Given the disputed origin of multidecadal variability in NASST (*e.g.*, it could be forced by anthropogenic aerosols or it might arise naturally from internal variability), the time series are filtered in two different ways appropriate to the two driving mechanisms. In either case, only a few climate models out of three dozen are found to generate internal variability consistent with observations. In fact, it is shown that climate models differ not only from observations, but they also differ from each other, unless they come from the same modeling center. In addition to these discrepancies in internal variability, other

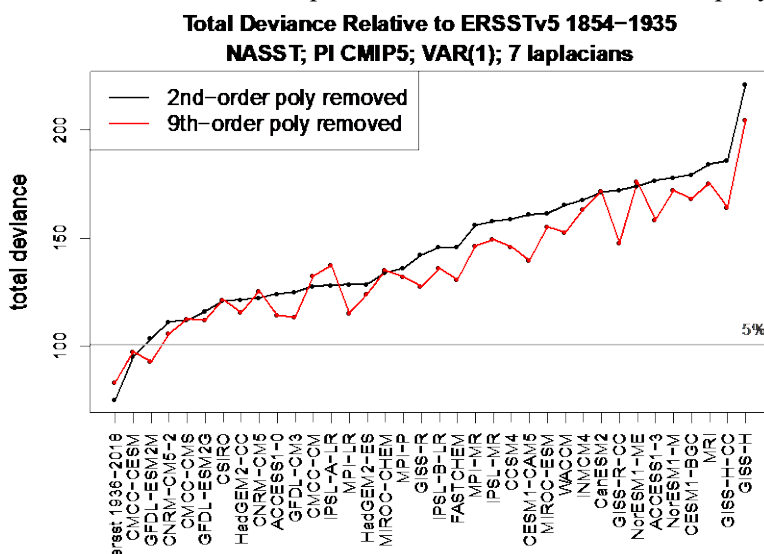


Fig. 1 Deviance between ERSSTv5 1854-1935 and 82-year segments from 36 CMIP5 pre-industrial control simulations. The black and red curves show, respectively, results after removing a second- and ninth-order polynomial in time over 1854-2018 before evaluating the deviance. Also shown is the deviance between ERSSTv5 1854-1935 and ERSSTv5 1937-2018 (first item on x-axis). The latter deviance falls below the 5% threshold and hence indicates no significant difference in internal variability between two halves of ERSST, regardless of polynomial fit. This result is consistent with the hypothesis that ERSST is a stationary VAR process after removing either a second- or ninth-order polynomial. Only one CMIP5 model is consistent with ERSST when a 2nd-order polynomial is removed, and only two CMIP5 models are consistent with ERSST when a 9th-order polynomial is removed. We conclude that the vast majority of CMIP5 models generate unrealistic internal variability.

studies show that models exhibit significant discrepancies with observations in terms of the response to external forcing. Taken together, these discrepancies imply that, at the present time, climate models do not provide a satisfactory explanation of observed variability in the North Atlantic.

This study has been published in *Advances in Statistical Climatology, Meteorology and Oceanography* in 2020. The follow-up researches can be found on Timothy DelSole's publication website at <http://cola.gmu.edu/delsole/publications.html>.

Reference

DelSole, T., and M. K. Tippett, 2020: Comparing climate time series – Part 1: Univariate test, *Adv. Stat. Clim. Meteorol. Oceanogr.*, **6**, 159–175, doi:10.5194/asmo-6-159-2020.

Preconditions for Extreme Wet Winters over the Contiguous United States

Andrew Hoell,¹ Martin Hoerling,¹ Jon Eischeid,^{1,2} and Joseph Barsugli^{1,2}

¹NOAA Physical Sciences Laboratory, Boulder, CO

²Cooperative Institute for Research in the Environmental Sciences,
 University of Colorado Boulder, Boulder, CO

ABSTRACT

We identify physical factors leading to extreme wet winters over the contiguous U.S. and examine whether preconditions operated during winter 2019 (December 2018 to February 2019) when record precipitation occurred (Fig. 1) that led to billion-dollar flood disasters along the Missouri and Mississippi Rivers. Models and observations are used to determine the effect of slow-varying forcing that may lead to practical forecast skill for extreme wet winters. Atmospheric models indicate that sea surface temperatures during strong eastern Pacific El Niño events like 1983 and 1998 can drive extreme wet winters over the contiguous U.S. These strong El Niños shift the distribution of contiguous U.S. precipitation to wetter conditions with a mean wetting of 1.5-2.0 standard deviations of the interannual variability. The shift to wetter conditions leads to a fivefold increase in the probability of wet winters of the magnitude observed in 2019. On longer timescales, observations indicate contiguous U.S. winter precipitation has increased over the last century. Analysis of historical coupled model simulations indicates anthropogenically-forced shifts to wetter conditions over the last century of 0.2-0.4 standard deviations of the interannual variability. While increasing the risk of extreme wet winters like 2019, this effect is a limited source of predictability during any particular winter. Concerning 2019 specifically, preconditioning factors of the risk for extreme contiguous U.S. winter wetness were weak or absent and offered little practical early warning (Fig. 2). The ongoing central Pacific El Niño that winter did not significantly alter the risk of the wetness, and thus the extreme 2019 conditions are judged not to have been a seasonal forecast of opportunity.

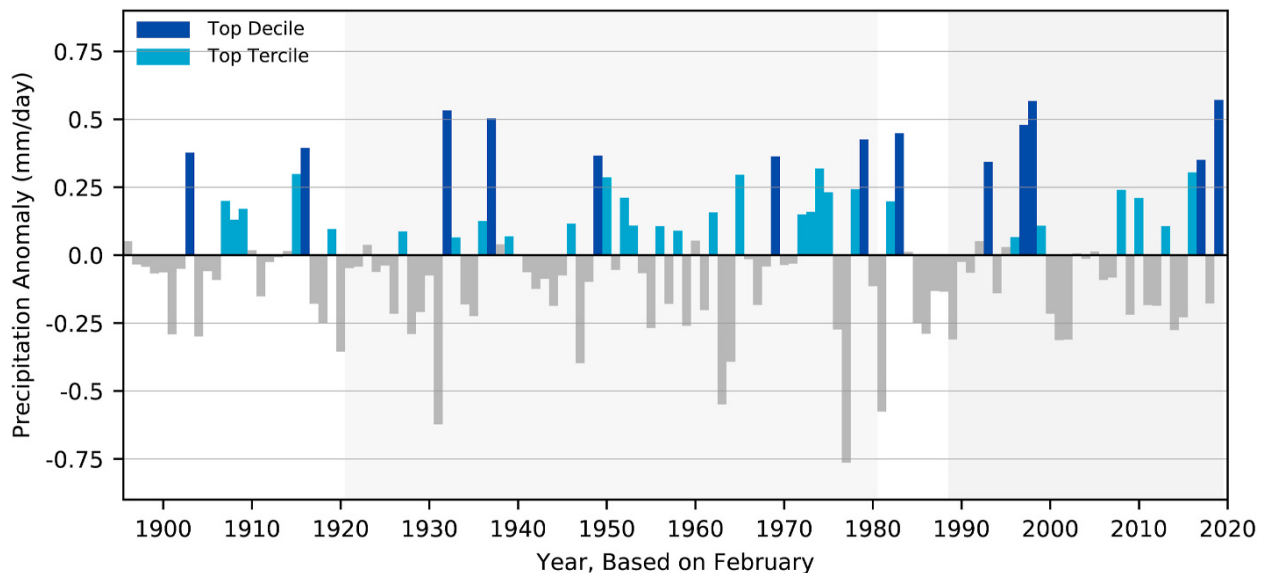


Fig. 1 December–February 1896–2019 contiguous U.S. average observed precipitation anomaly in mm/day from nClimGrid/CLIMGRID (Vose *et al.*, 2014). Anomalies are based on 1981–2010 and percentile ranks are based on 1896–2019. 1989–2018 and 1921–1980 are shaded to highlight the focus eras of anthropogenic preconditioning in discussion.

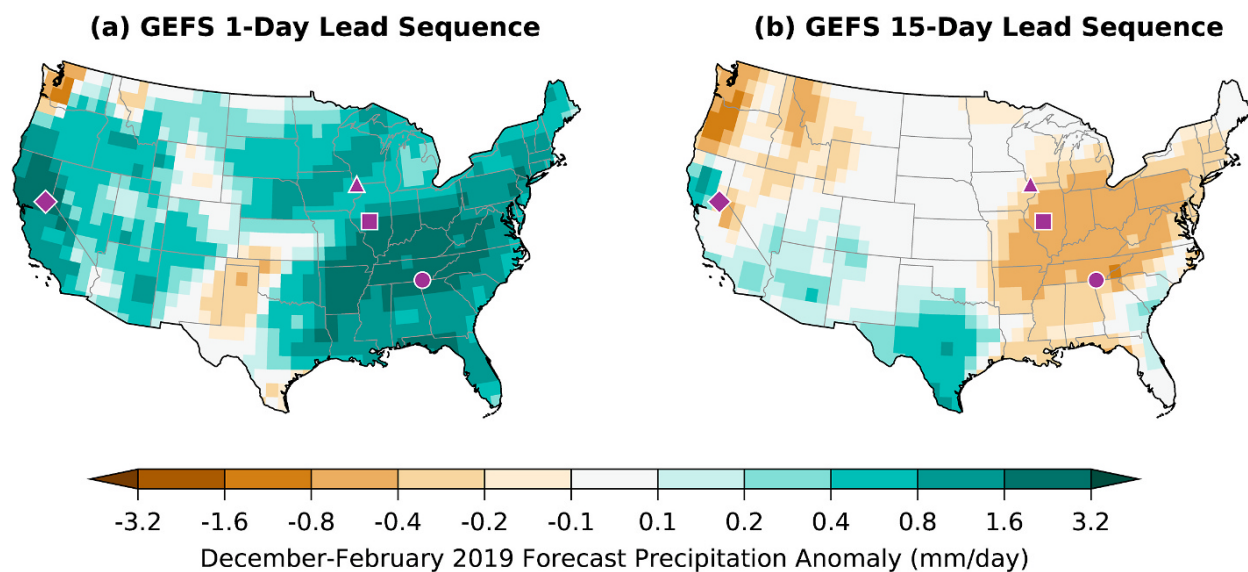


Fig. 2 December–February 2019 precipitation anomaly in mm/day constructed from sequences of GEFS ensemble average (a) 1-day lead forecasts and (b) 15-day lead forecasts initialized daily. Chattanooga, Tennessee (circle), Springfield, Illinois (square), Dubuque, Iowa (triangle), Pacific House, California (diamond) are indicated.

* The map of accumulated daily precipitation anomalies based on sequences of 1-day leads closely resemble the observations as expected, affirming the well-known high skill of synoptically driven cold-season precipitation at such short leads. By contrast, the 15-day leads provide no skill for the wetness. The severe degradation in quality of the lead-dependent weather predictions by the second week affirms that boundary forcing unlikely played an important role in creating America's wettest winter.

This work has been published in *Weather and Climate Extremes* in 2021.

References

- Hoell, A., M. Hoerling, J. Eischeid, and J. Barsugli, 2021: Preconditions for extreme wet winters over the contiguous United States. *Weather Clim. Extrem.*, **33**, 1–14, doi:10.1016/j.wace.2021.100333.
- Vose, R. S., and Coauthors, 2014: Improved historical temperature and precipitation time series for U.S. climate divisions. *J. Appl. Meteorol. Climatol.*, **53**, 1232–1251.

MJO Impacts on Winter Weather Event Frequency

Stephen Foskey and Naoko Sakaeda
University of Oklahoma, Norman, OK

1. Introduction

Winter weather can have a major impact on society. Improving forecasting of winter weather would be beneficial to individuals and society by providing them with more time to prepare. The Madden-Julian Oscillation (MJO) has a major influence on weather and climate patterns around the globe. The MJO has been shown to influence the predictability of 500 hPa height patterns out to two or three weeks (Tseng *et al.* 2017). The MJO can also impact temperature patterns across the globe (*e.g.*, Zhou *et al.* 2012), as well as stratospheric conditions (Green and Furtado 2019). If the MJO can be used to help predict winter weather, that would extend the forecast lead time for winter weather beyond the typical 7-10 days for synoptic scale systems (Zhang *et al.* 2019). Some research has been done on how the MJO influences snowfall over eastern New England (Klotzbach *et al.* 2016), but there has been less research on the MJO and winter weather over the rest of the country. The goal of this research is to determine how the phase of the MJO impacts the frequency of winter weather events over the contiguous United States. We hypothesize that the MJO impacts winter weather over the United States through the changes in the flow pattern influenced by the MJO and its effect on thermodynamic and dynamic properties. The upper-level divergence associated with the MJO has been shown by Sardeshmukh and Hoskins (1988) to lead to the generation of Rossby waves, which can become amplified and lead to blocking. This blocking causes temperature anomalies in areas downstream from the ridging and a consistent jet pattern that is supportive of winter storms.

2. Data and methods

We used winter storm data from the NCEI Storm Events Database. This database is a collection of extreme weather events from 1996 to the present. For this study, events were chosen which met winter storm warning criteria for local forecast offices and were divided into three categories: winter storm, ice storm, or heavy snow. Heavy snow events have only snow as a precipitation type, ice storms have only freezing rain, and winter storms have some combination of snow, sleet, and freezing rain. The most frequent months for winter weather were December through March. These months were selected for further analysis due to their increased frequency of winter weather. The strength and phase of the MJO are primarily identified using the Outgoing Longwave Radiation (OLR) MJO Index (OMI) (Kiladis *et al.* 2014). The results based the OMI were compared to the Real-time MJO Monitoring Index (RMM) to check consistency.

Using the NCEI Storm Events Database, the frequency of severe winter weather events in the different storm categories were calculated by weather forecast office (WFO). The calculation of storm frequency by WFO provides consistency in storm criteria because each WFO has its unique criteria for classifying weather events that are reported. The frequency is defined as the number of reports per day normalized by the number of zones within each WFO. To determine the impact of the MJO on the frequency of severe winter weather events, we found the ratio of the frequency of winter weather events in a given phase of the MJO index to its climatological frequency. Frequency ratios greater than one indicate events are more frequent than climatology, while those less than one indicate events that are less frequent than climatology.

3. Analysis

The climatological frequency of extreme winter weather events shows that heavy snow is most frequent in the Northeast and Northwest U.S., winter storms are most frequent in the Northeast and North Central U.S., and ice storms are most frequent in the Central U.S. The maximum frequency of heavy snow and winter storms is around 3-4 reports per year, while it is less than 0.5 report per year for ice storms.

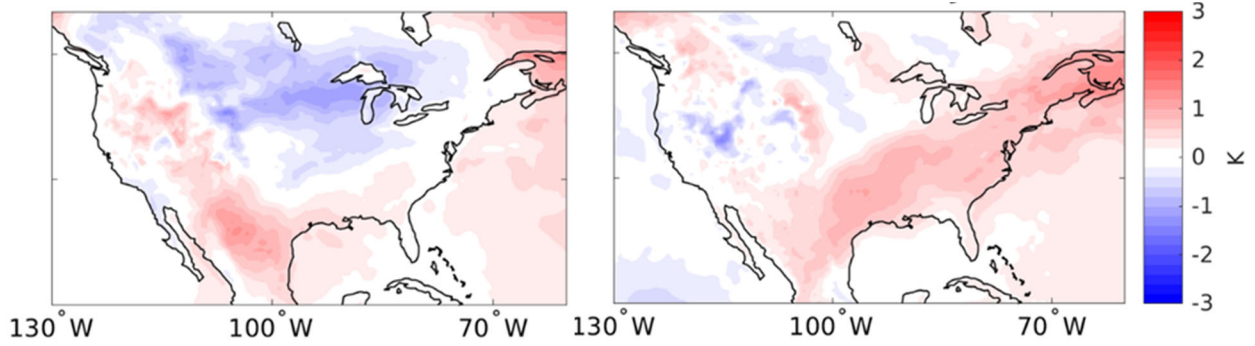


Fig. 1 DJFM surface temperature anomalies for OMI phase 2 (left) and phase 4 (right) of the MJO.

The MJO phases analyzed here were phase 2 and phase 4, due to their relatively large frequency of winter weather events. During phase 2 of OMI or RMM, which is associated with enhanced MJO convection over the Indian Ocean, surface temperature tends to be colder over the United States (Fig. 1, left) (Zhou *et al.* 2012). Phase 4 (which has enhanced MJO convection over the Maritime Continent) is associated with warmer temperatures over the southeastern half of the United States, and near-normal temperatures over the rest of the country (Fig. 1, right). Phase 2 featured a trough over the northern U.S. and southwesterly flow across the middle of the country (Fig. 2, top). Phase 4 featured a trough in the northwest corner of the country and ridging to the east (Fig. 2, bottom). There was a strong jet in the western half of the country, with weak flow to the southeast.

There are significant variations in the frequency ratio of winter weather events by MJO phase. These variations did not always align with temperature anomalies. For heavy snow events, phase 2 had large frequency ratios in the Southern U.S., while phase 4 had large frequency ratios in the Northern Plains (Fig. 3a, b). This is in spite of the Northern Plains having near-normal to above-normal temperatures. However, the areas of greatest frequency ratios are near the areas of sharpest temperature gradient. For winter storms, the pattern is different. In phase 2, the greatest frequency ratios are in the Northeast, while in phase 4, they shift to the western U.S. (Fig. 3c, d). The area of large frequency ratios is very pronounced and includes areas with both warmer and colder than normal temperature anomalies. Ice storm reports are quite frequent across the eastern half of the country in phase 2, but less so in phase 4 (Fig. 3e, f). There are fewer areas of significance for ice storms because of the smaller sample size. The ice storm frequency ratios are more closely correlated to temperature. When comparing the OMI and RMM, the results are similar, however they can vary slightly over particular regions.

4. Conclusion

The frequency of winter weather across the United States is affected by the MJO phase. However, the impacts vary based on storm type and region. The original hypothesis that the impacts vary primarily on temperature will need to be revisited. That pattern appears to be true for ice storms, but not other storm types to the same extent. It appears likely that wind patterns including areas of convergence and divergence also play

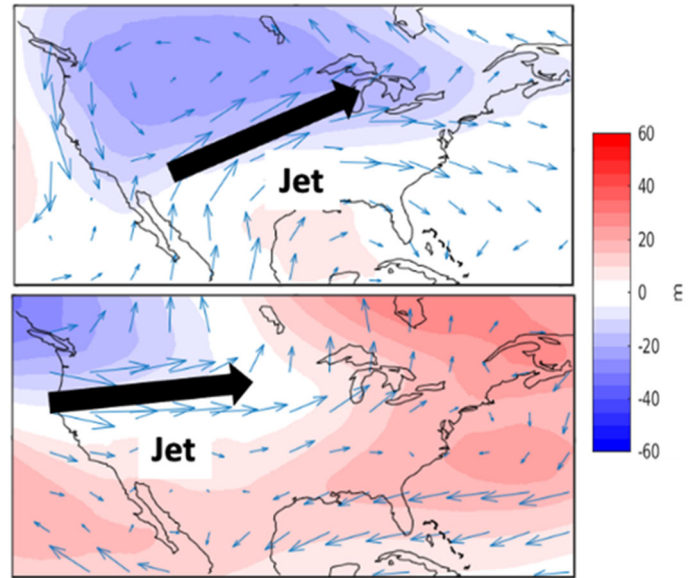


Fig. 2 Geopotential height anomalies at 500 hPa and wind vector anomalies at 250 hPa for phase 2 (top) and phase 4 (bottom) of the MJO.

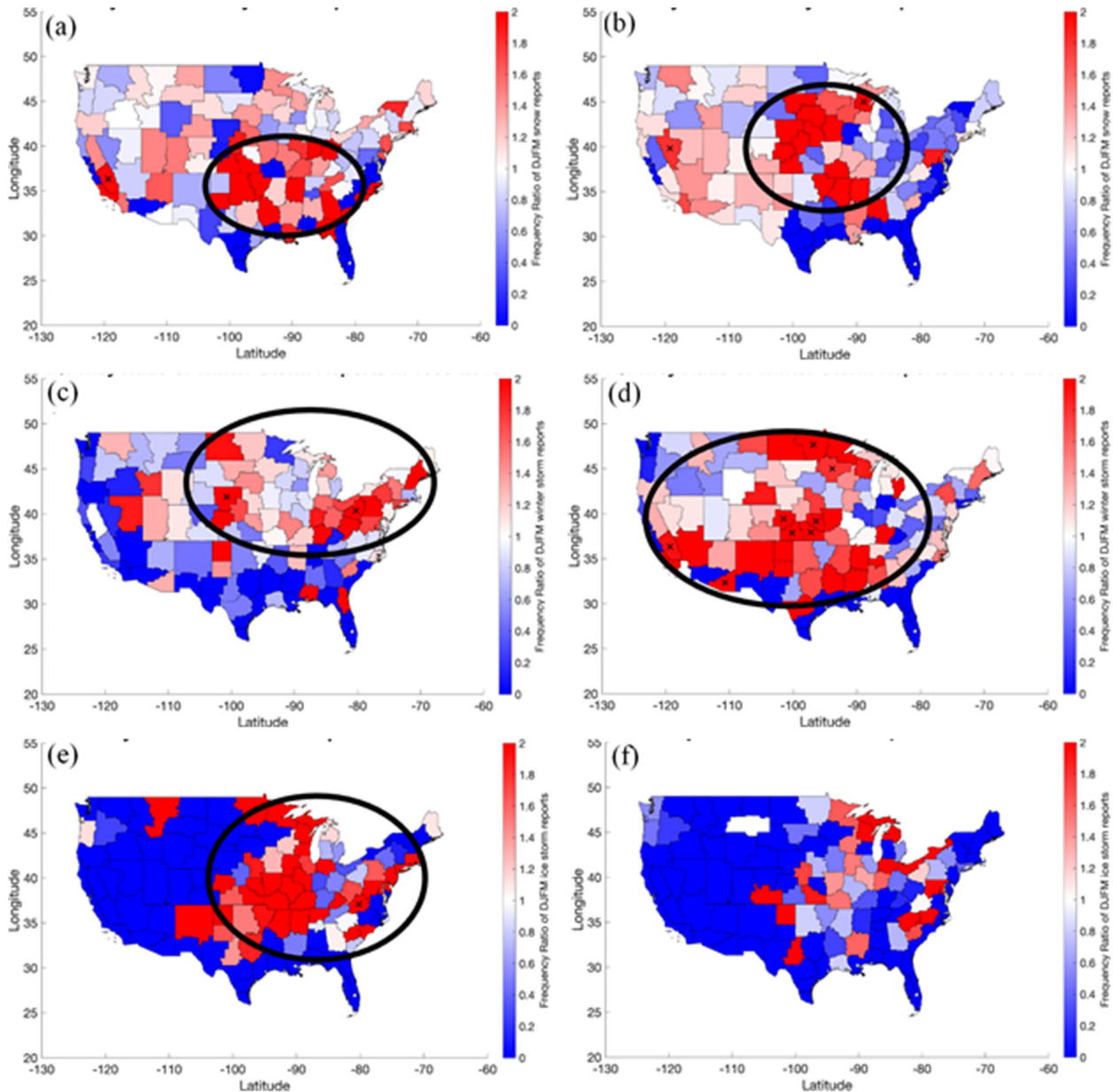


Fig. 3 Frequency ratios from 1996-2018 for winter event types. Subplots are as follows: a) heavy snow reports for phase 2 and b) phase 4 of the MJO, c) winter storm reports for phase 2 and d) phase 4 of the MJO, e) ice storm reports for phase 2 and f) phase 4 of the MJO. Crosses indicate significance at the 90% level using bootstrapping.

a role. Further research since the time of the conference has focused on analyzing these topics and looking at other properties such as moisture.

References

- Green, M. R. and J. C. Furtado, 2019: Evaluating the joint influence of the Madden-Julian Oscillation and the stratospheric polar vortex on weather patterns in the Northern Hemisphere. *J. Geophys. Res. Atmos.*, **124**, 11693-11709, <https://doi.org/10.1029/2019JD030771>.
- Kiladis, G. N., J. Dias, K. H. Straub, M. C. Wheeler, S. N. Tulich, K. Kikuchi, K. M. Weickmann, and M. J. Ventrice, 2014: A comparison of OLR and circulation-based indices for tracking the MJO. *Mon. Wea. Rev.*, **142**, 1697-1715, <https://doi.org/10.1175/MWR-D-13-00301.1>.

-
- Klotzbach, P. J., E. C. J. Oliver, R. D. Leeper, and C. J. Schreck III, 2016: The relationship between the Madden-Julian Oscillation (MJO) and southeastern New England snowfall. *Mon. Wea. Rev.*, **144**, 1355-1362, <https://doi.org/10.1175/MWR-D-15-0434.1>.
- Sardeshmukh, P. D. and B. J. Hoskins, 1988: The generation of global rotational flow by steady idealized tropical divergence. *J. Atmos. Sci.*, **45**, 1228-1251, [https://doi.org/10.1175/1520-0469\(1988\)045<1228:TGOGRF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1988)045<1228:TGOGRF>2.0.CO;2).
- Tseng, K.-C., E. A. Barnes, and E. D. Maloney, 2017: Prediction of the midlatitude response to strong Madden-Julian Oscillation events on S2S time scales. *Geophys. Res. Lett.*, **45**, 463-470, <https://doi.org/10.1002/2017GL075734>.
- Zhang, F., Y. Q. Sun, L. Magnusson, R. Buizza, S. Lin, J. Chen, and K. Emanuel, 2019: What is the predictability limit of midlatitude weather? *J. Atmos. Sci.*, **76**, 1077-1091, <https://doi.org/10.1175/JAS-D-18-0269.1>.
- Zhou, S., M. L'Heureux, S. Weaver, and A. Kumar, 2012: A composite study of the MJO influence on the surface air temperature and precipitation over the continental United States. *Clim. Dyn.*, **38**, 1459-1471, <https://doi.org/10.1007/s00382-011-1001-9>.

Prediction Challenges Associated with Errors in Linear Trends of Tropical Pacific Sea Surface Temperature

Michelle L. L'Heureux,¹ Michael K. Tippett,² and Wanqiu Wang¹

¹Climate Prediction Center, NOAA/NWS/NCEP, College Park, MD

²Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY

The research associated with the presentation is now published at Frontiers in Climate (L'Heureux et al. 2022), which is freely available and open-access at the website:

<https://doi.org/10.3389/fclim.2022.837483>

What follows is a brief summary, including a subset of figures that were not featured in the paper.

SUMMARY

Across the tropical Pacific Ocean, linear trends in sea surface temperature (SST) were analyzed over all calendar months from 1982-2020 in the observations (Fig. 1) and in predictions from the North American Multi-model Ensemble (NMME; Fig. 2). This period was chosen because reforecasts from the NMME are only available since 1982. To summarize the model data, the lead-0.5, lead-4.5, and lead-8.5 monthly forecasts are shown in the subpanels of Fig. 2.

Figure 1 shows that, over 1982-2020, the largest positive linear trends were observed in the western equatorial Pacific Ocean, while trends are very small and insignificant across the east-central and eastern equatorial Pacific Ocean. In the NMME predictions (Fig. 2), the spatial distribution of these linear trends are generally captured at the shortest lead times, with the exception of the NCEP-CFSv2 and CCSM4 models, which have immediate positive trend errors at the shortest lead time. This is due to errors arising from their initial condition, which are from the CFSR, which has a discontinuity in ~1999 so that forecasts are warmer after 1999 than before (Xue *et al.*, 2011; Kumar *et al.* 2012). In operations, this is addressed by removing two different monthly mean climatology periods (*e.g.* 1982-1998 and 1999-2010) prior to computing an anomaly. However, here, we are examining total SSTs and therefore do not correct for this error.

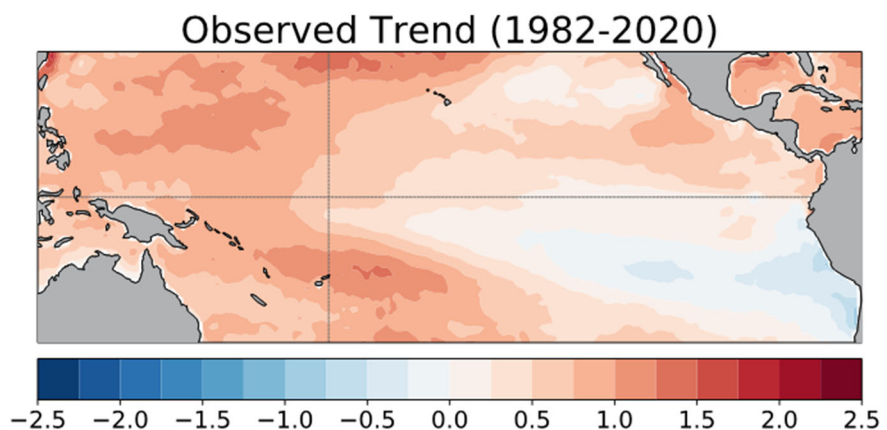


Fig. 1 Linear trend of monthly sea surface temperature from January 1982 to December 2020. Units are in degrees Celsius change over January 1982 to December 2020. The thin gray lines indicate the equator and the International Date Line. Data is based on OISSTv2.

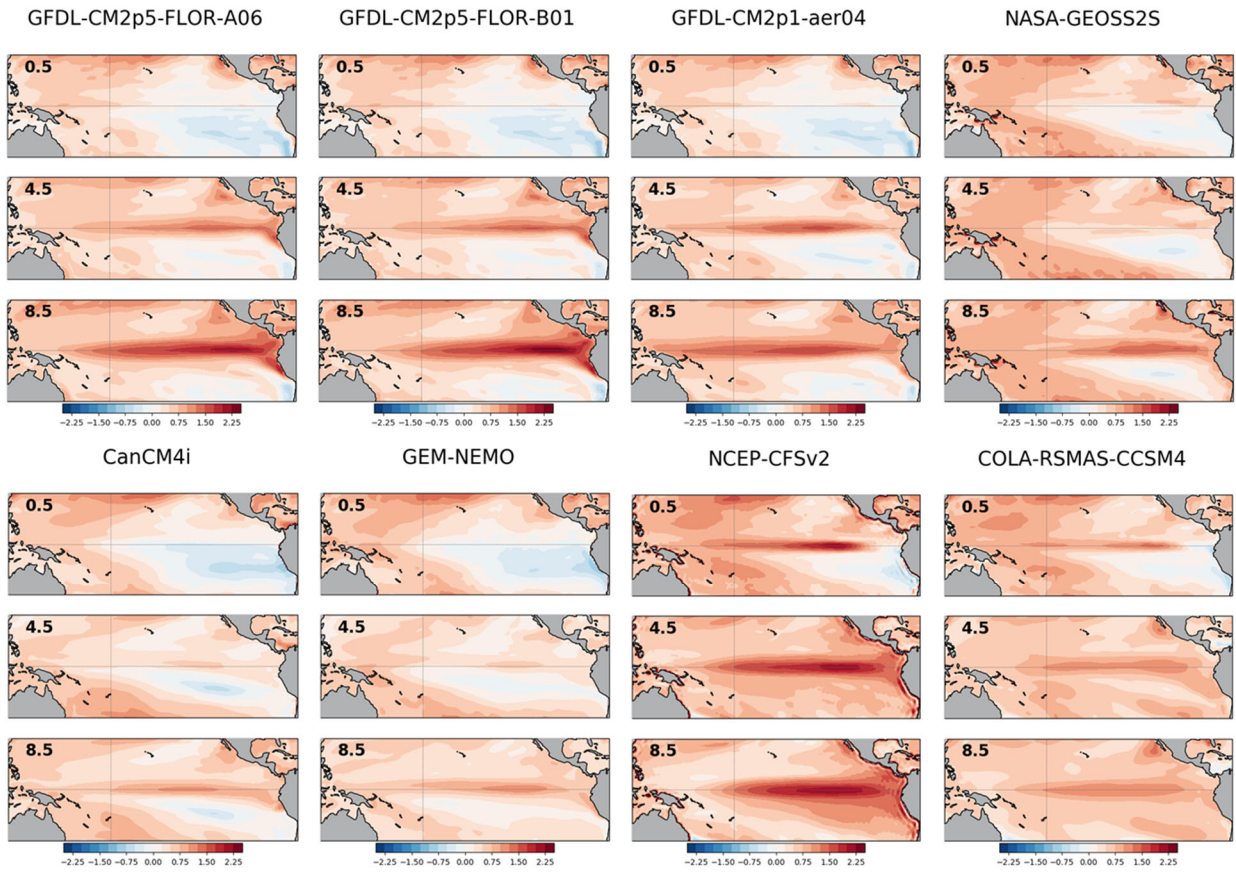


Fig. 2 For each model in the NMME, linear trend (forecasted coefficient minus observed coefficient) in monthly sea surface temperatures from January 1982 to December 2020. The bold number in each panel shows the forecast lead time (0.5-, 4.5-, and 8.5-month leads). Units are in the degrees Celsius change over January 1982 to December 2020. The thin gray lines indicate the equator and the International Date Line.

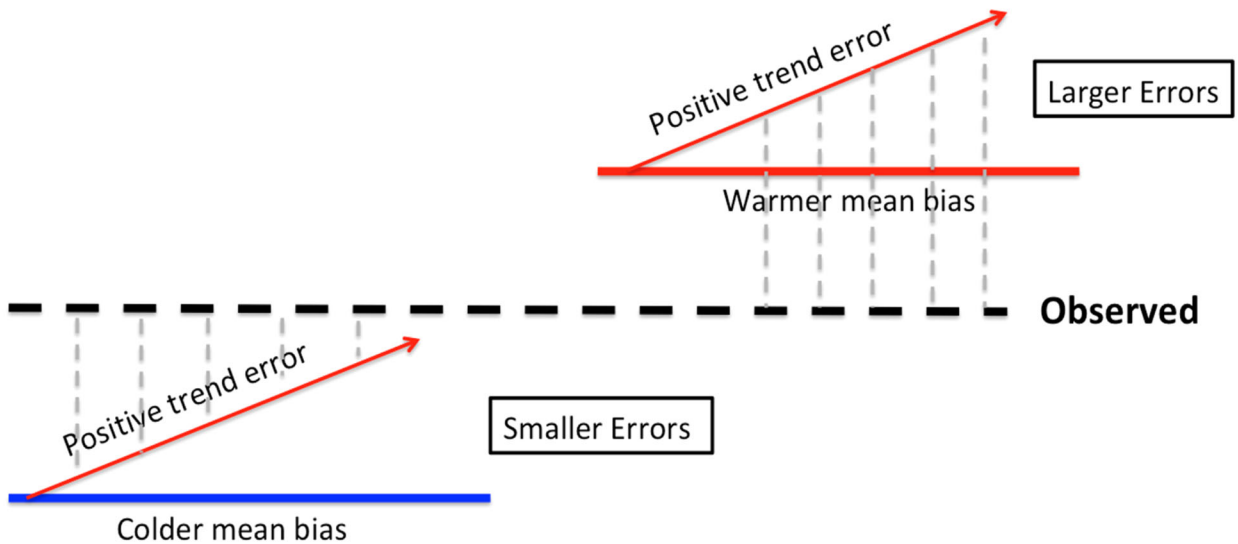


Fig. 3 Schematic of positive SST trends superimposed on models with a colder or warmer mean SST bias. Vertical grey lines show the increase in error with time for warm mean biased models and decrease in error for cold mean biased models.

At longer monthly lead times, which become apparent after the ~4.5-month lead time, there are positive trends in the eastern and central Pacific Ocean across all models, which is in stark contrast to the observed linear trends presented in Fig. 1. These positive trends grow with lead-time, so that all models have substantial positive trends in the central and eastern Pacific Ocean by lead-8.5.

These linear trend errors in the NMME are also associated with linear trends in squared errors (forecast minus observations), so that some models predictions are improving over time (1982-2020), while other models worsen (*not shown*). It turns out that the direction of the error is determined by the sign of the model's mean bias, which is shown as a schematic in Fig. 3. Some models have a cold mean bias (blue line) and so positive trend errors result in forecasts that are closer to the observations over time (black dashed line), and subsequently, smaller errors over time (dashed grey lines). In contrast, other models have a warm mean bias (red line) and positive trend errors result in predictions that are farther from the observations over time.

In addition to the SST trend errors over the 1982-2020 period, there are precipitation trend errors over the same period (*not shown*). Accompanying the too warm SST predictions, the precipitation predictions have trended in the direction of being too wet. Overly warm and wet predictions are accompanied by an increase in El Niño False Alarms, or cases when El Niño was predicted and did not occur.

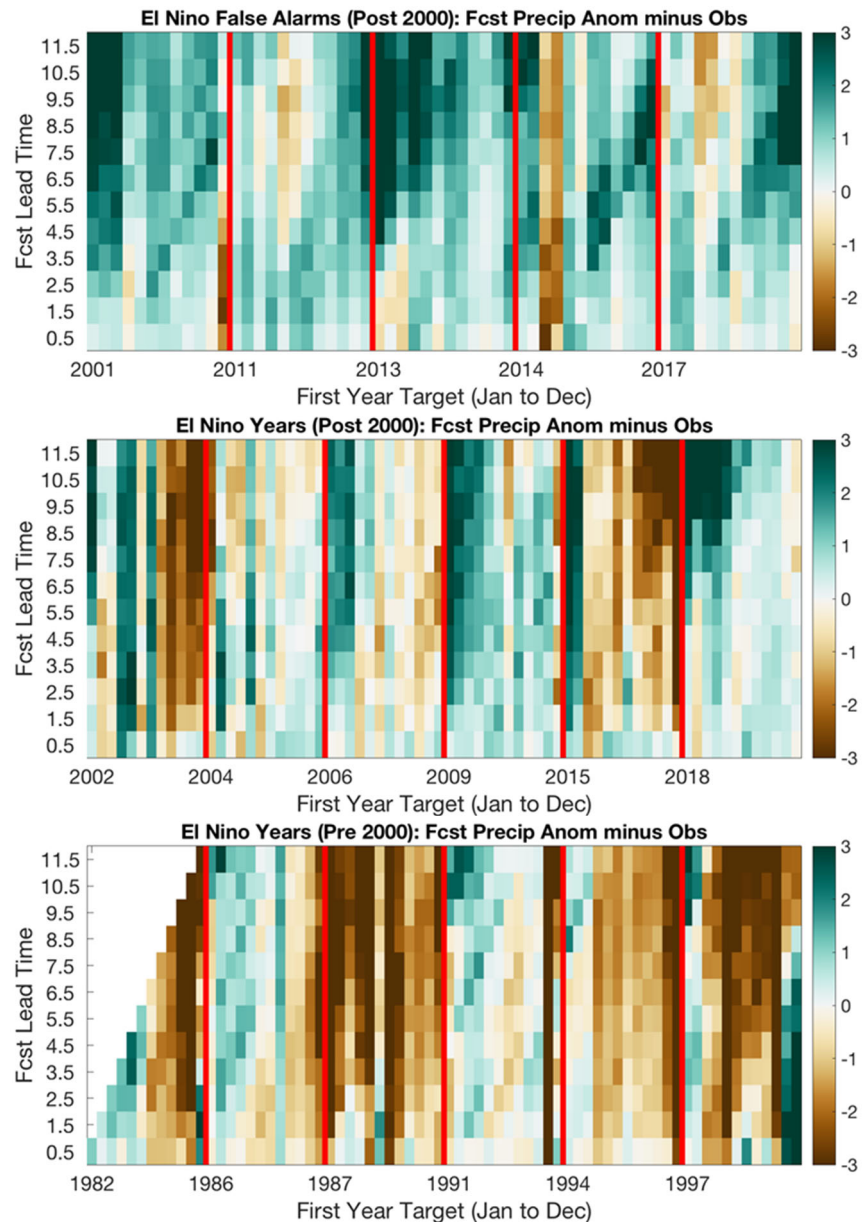


Fig. 4 The error (forecast minus observations) of central Pacific precipitation anomalies for the NMME model average across monthly lead-time (y-axis). The x-axis displays El Niño years (middle and bottom panels) and El Niño false alarm years (top panel). Along the x-axis, there are twelve squares between the vertical red lines, which represent January-December target times for the first (onset) calendar year of the El Niño or El Niño false alarm case. The middle panel shows error for El Niño cases after 2000 and the bottom panel shows El Niño cases prior to 2000. The top panel includes El Niño false alarm cases after 2000.

Averaging precipitation anomalies in the equatorial central Pacific Ocean, we examined the precipitation anomaly errors during the first calendar year of El Niño events between 1982-2020 (Fig. 4 — bottom two panels) and also five El Niño false alarm events (Fig. 4 — top panel). It is clear that predictions of precipitation anomalies during El Niño events after 2000 (middle panel) are biased wetter than predictions before 2000

(brown shading in bottom panel). However, the largest positive precipitation anomaly errors (forecasts were too wet) occur during El Niño false alarms (top panel). Interestingly, four of the five El Niño false alarms have occurred since 2010 (Tippett *et al.*, 2020), which implies the forecast trends toward a warmer and wetter tropical Pacific may be increasing the false alarm frequency (see additional analysis in *Frontiers in Climate*).

References

- Kumar, A., M. Chen, L. Zhang, W. Wang, Y. Xue, C. Wen, and Coauthors, 2012: An analysis of the nonstationarity in the bias of sea surface temperature forecasts for the NCEP Climate Forecast System (CFS) version 2. *Mon. Wea. Rev.*, **140**, 3003–3016, doi: 10.1175/MWR-D-11-00335.1.
- L'Heureux, M. L., M. K. Tippett, and W. Wang, 2022: Prediction challenges from errors in tropical Pacific sea surface temperature trends. *Front. Clim.*, **4**, doi:10.3389/fclim.2022.837483.
- Tippett, M. K., M. L. L'Heureux, E. J. Becker, and A. Kumar, 2020: Excessive momentum and false alarms in late-spring ENSO forecasts. *Geophys. Res. Lett.*, **47**, e2020GL087008. doi: 10.1029/2020GL087008.
- Xue, Y., B. Huang, Z.-Z. Hu, A. Kumar, C. Wen, D. Behringer, and Coauthors, 2011: An assessment of oceanic variability in the NCEP climate forecast system reanalysis. *Clim. Dyn.*, **37**, 2511–2539. doi: 10.1007/s00382-010-0954-4

Skillful Long-lead Prediction of Summertime Heavy Rainfall in the US Midwest from Ocean Salinity

Laifang Li,^{1,3,4} Raymond W. Schmitt,² and Caroline Ummenhofer²

¹Department of Meteorology and Atmospheric Science, Penn. State University, University Park, PA

²Physical Oceanography Department, Woods Hole Oceanographic Institution, Woods Hole, MA

³Institute of Computational and Data Science, Penn. State University, University Park, PA

⁴Earth and Environmental Science Institute, the Penn. State University, University Park, PA

1. Introduction

The hazardous impacts of heavy rainfall on agriculture and the economy in the US Midwest cannot be overstated (Ray *et al.* 2015). These heavy rainfall events are generated in atmospheric circulation patterns distinct from moderate or light rainfall, which challenges heavy rainfall prediction (*e.g.*, Doblas-Reyes *et al.* 2013; White *et al.* 2017). Previous attempts to predict Midwest heavy rain based on pre-season sea surface temperature anomaly (SSTA) (*e.g.*, McCabe *et al.* 2004; Schubert *et al.* 2016) hardly generated skill that met the forecast accuracy for socioeconomical needs (Li *et al.* 2018; Carter *et al.* 2021). We are motivated to search for more skillful predictors for summertime heavy rain in the Midwest.

These predictors can be sought from the oceanic water cycle, as the ocean is the ultimate moisture source for precipitation on land (Gimeno *et al.* 2010), and the oceanic influence amplifies as rainfall intensity increases (*e.g.*, Jana *et al.* 2018; Liu *et al.* 2021). The oceanic moisture export is also a forcing mechanism on sea surface salinity (SSS), an ocean state variable highly sensitive to moisture flux over the ocean (Schanze *et al.* 2010; Gordon 2016; Vinogradova *et al.* 2019). The close connection between SSS, oceanic moisture, and terrestrial precipitation suggests that the upstream SSS anomaly (SSSA) could be a skillful predictor of precipitation on land. Encouraging skill of SSSA-based seasonal rainfall forecasts was shown for the African Sahel (Li *et al.* 2016a), the US (Li *et al.* 2016b; Liu *et al.* 2018), China (Zeng *et al.* 2019), and Australia (Rathore *et al.* 2021). Here, we present an additional line of evidence that springtime SSSA in the tropical Pacific and subtropical North Atlantic are skillful predictors of summertime heavy rain in the US Midwest.

2. Data and methods

2.1 Bayesian inference on summer-season heavy rainfall events over the US Midwest

A Bayesian 3-cluster Gaussian mixture model (GMM; Li and Li 2013) is implemented to objectively categorize Midwest daily precipitation into light, moderate, and heavy rainfall:

$$y_i | \pi, \mu, \phi \sim \sum_{h=1}^3 \pi_h \text{Gauss}(y_i | \mu_h, \phi_h^{-1}), \quad (1)$$

where $y_i = \ln Pr_i$ is the natural log of precipitation in the i^{th} day, π is cluster weight (*i.e.*, rainfall frequency), μ is cluster mean (*i.e.*, rainfall intensity), and ϕ is the precision parameter. $h \in \{1, 2, 3\}$ is the cluster index, respectively representing light, moderate, and heavy rainfall cluster. The distribution parameters (π, μ, ϕ) are sampled using Markov Chain Monte Carlo algorithm (Gelfand 2000). The code for the Bayesian model is available at: <https://doi.org/10.5281/zenodo.5389218>.

2.2 Logistic regression

To assess the predictive skills of preseason SSTA and SSSA in forecasting US Midwest heavy rainfall frequency (π_3 in the GMM), we construct a logistic regression model:

$$\pi_3 = \frac{1}{1 + e^{-\theta^T x + \varepsilon}}. \quad (2)$$

In Eq. (2), x is the predictors, and θ is the regression coefficients. $\varepsilon \sim N(0, \sigma^2)$ is the residual of the regression which is normally distributed with zero mean and variance of σ^2 .

In constructing the regression model, we randomly draw half the data as training samples and the remaining half are independent samples for validation. The regression coefficients are first derived from the training sample and then applied to the validation samples to estimate the prediction errors. This training-validation process is repeated 1000 times. The parameter set that minimizes the prediction errors of the validation samples is selected to create the final logistic regression model that is applied to the entire dataset.

3. Results

3.1 Heavy precipitation explains the majority of the US Midwest summer precipitation variance

According to the GMM (Eq. 1), the US Midwest receives an accumulation of 261.7 mm of precipitation, with $38.5 (\pm 7.2)$ mm from light rainfall, $121.9 (\pm 23.7)$ mm from moderate rainfall, and $101.3 (\pm 43.3)$ mm from the heavy rainfall cluster during the summer season (Fig. 1a). At an interannual time scale, the variation of summer precipitation is almost exclusively explained by heavy rain. The R^2 between the heavy rainfall cluster and season-total precipitation is 0.81 (Fig. 1a). In contrast, the variance explained by moderate (light) rainfall is lower at $R^2 = 0.29$ (negative at $R^2 = -0.21$) (Fig. 1a). Our analysis suggests that heavy rainfall, despite weather event, exerts climatic impacts on water resources in the summer. Further, heavy rainfall frequency is overwhelmingly more important than heavy rainfall intensity to interannual variation of seasonal precipitation, *i.e.*, the correlation between seasonal-mean precipitation and heavy rainfall intensity is 0.18 ($p = 0.13$) but reaches 0.90 ($p \leq 0.00$) with heavy rainfall frequency (Fig. 1b-c).

3.2 Springtime salinity provides predictive values to US Midwest heavy rainfall

We explore oceanic precursors of US Midwest heavy rainfall by regressing March-April-May (MAM) SSTA and SSSA upon summertime heavy rainfall frequency (π_3 in Eq. 1). More frequent heavy rainfall would be expected following an anomalously warm (cold) tropical central Pacific (Gulf of Mexico (GOM)) (Fig. 2a). Corresponding to the SSTA, the center of tropical convection is shifted near the dateline, resulting in a dipole SSSA pattern with anomalously low salinity along the equator and high salinity to the south (Fig. 2b; Guimbar *et al.* 2017). The low SSSA along the equator reaches -0.1 PSU (Fig. 2b). With the typical SSS (33PSU) in this region, and mixed layer depth of 15m, the observed SSSA could correspond to an 1.8 mm day^{-1} increase in precipitation ($P' \sim -\frac{\rho_0 h_m \text{SSSA}}{\bar{S} \Delta t}$) in the spring, forcing atmospheric teleconnection and thus the response in extratropical precipitation (Horel and Wallace 1981; Liu and Alexander 2007). In the North Atlantic, the preseason SSSA is most significant off the Eastern coast of the US (Fig. 2b). The SSSA pattern is consistent with our previous studies linking a saltier subtropical North Atlantic to increased summer precipitation over the Midwest (Li *et al.* 2016b, 2018).

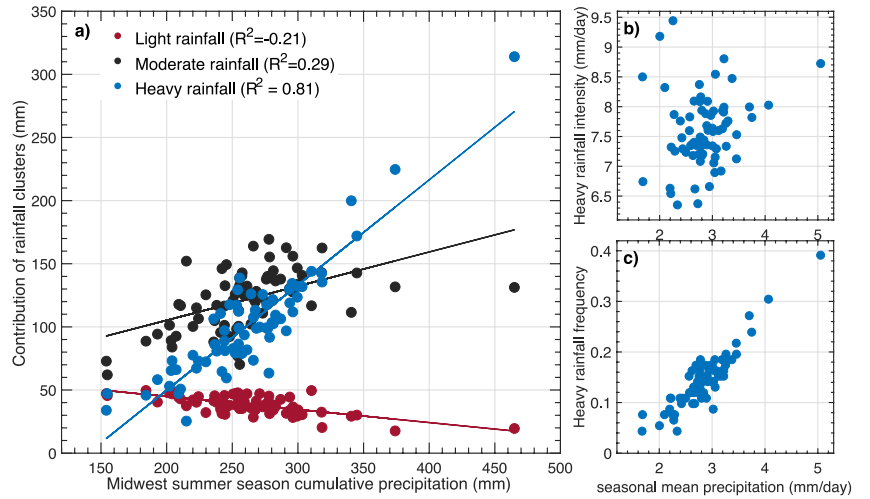


Fig. 1 a) Contribution of the light (red dots), moderate (black dots), and heavy (blue dots) rainfall to summer season cumulative precipitation in the Midwest. The lines are the linear regression of cluster total precipitation against cumulative precipitation in summer. b) and c) are, respectively, the scatter plots of summer-season precipitation versus heavy rainfall intensity and frequency. (from Li *et al.* 2022, in review with GRL)

Based on MAM SSTA and SSSA in Fig. 2a-b, we define four predictors: Niño 3.4 SSTA, GOM SSTA, Tropical western Pacific SSSA dipole index, and North Atlantic SSSA. These four predictors are supplied to the logistic regression model (Eq. 2) to predict summertime heavy rainfall frequency (Fig. 2c-e). The R^2 value between the observations and the prediction is 0.46 (Fig. 2c). In particular, the model predicts a doubling of heavy rainfall in the wettest summer of 1993 and reduced heavy rainfall in the dry summer of 2012 (Fig. 2c), suggesting the capability of the logistic regression model to forecast extreme events one season ahead.

Overall, SSSA-based predictors are almost twice as skillful ($R^2 = 0.27$; Fig. 2e) than SSTA-based predictors ($R^2 = 0.14$; Fig. 2d). It is noteworthy that the Pacific SSSA dipole significantly correlates with Niño 3.4 but is more skillful in predicting Midwest heavy rainfall. The superior skill of the North Pacific SSSA dipole is attributable to tropical convection as the dominant forcing on the SSSA (Sean-Martins and Stammer, 2015). Without a damping mechanism, the SSSA reflects cumulative tropical convection and latent heating released into the atmosphere, a driver of atmospheric teleconnection (Gill *et al.* 1980; Scaife *et al.* 2017). Via tropical-extratropical teleconnection, the SSSA dipole turns out to be a more skillful predictor of precipitation on land.

4. Conclusion and discussion

According to the Bayesian GMM applied to summertime daily precipitation in the US Midwest, heavy rainfall events are primarily responsible for the year-to-year variation of cumulative seasonal precipitation, a meteorological variable that significantly impacts yields in this important crop production region. Here, we explore pre-season oceanic state variables for predicting summer-season heavy rainfall in the US Midwest. Based on the logistic regression, skillful prediction of US Midwest heavy rain is achieved one season ahead with the knowledge of SSTA over the Niño 3.4 region and GOM and SSSA over the tropical western Pacific and subtropical North Atlantic. The combination of the four predictors leads to prediction with $R^2 = 0.46$, in which the SSSA-based predictors provide superior skill compared to SSTA-based predictors (Fig. 2).

The potential mechanisms underlying the SSSA's predictive skill could involve both atmospheric teleconnection and soil moisture feedback. Specifically, the Pacific salinity dipole is indicative of tropical convection and the resultant tropical-extratropical teleconnection. Through the long memory of extratropical SSTA, and the coupling between SSTA, pressure, and precipitation, a quasi-barotropic wave train that resembles the typical circulation during US Midwest heavy rainfall events is triggered and maintained. In addition, the subtropical North Atlantic SSSA is skillful in predicting summer-season heavy rain mainly due to a positive soil moisture feedback mechanism as identified previously (Li *et al.* 2016b, 2018). With the cumulative effects of soil moisture feedback and tropical-extratropical teleconnection, the combination of Pacific SSSA dipole and North Atlantic SSSA provides indispensable value to predict US Midwest heavy rainfall, complementing the current statistical forecasting based solely on pre-season SSTA.

Furthermore, the newly identified pre-season SSSA patterns could become more valuable as the climate warms: ENSO teleconnections are sensitive to a changing background state in response to natural and anthropogenic forcing (McPhaden *et al.* 2011; Yeh *et al.* 2019). In addition, the oceanic water cycle is predicted to intensify in the future (Durack *et al.* 2012; Levang *et al.* 2015), and the oceanic moisture export to more

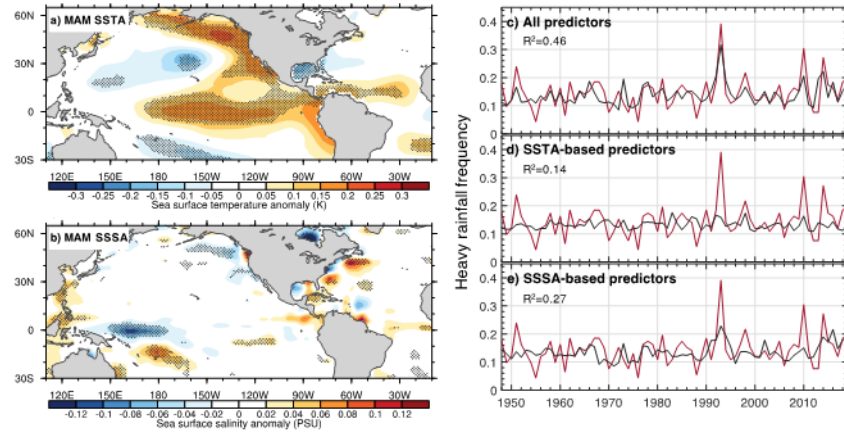


Fig. 2 MAM SSTA (a) and SSSA (b) regressed upon Midwest heavy rainfall frequency. Grid cells with regression coefficients significant at $\alpha = 0.05$ level are stippled; c-e) are US Midwest heavy rainfall frequency predicted by the logistic regression model: c) all predictors; d) SSTA-based predictors; e) SSSA-based predictors. The red curves are the observations. (from Li *et al.* 2022, in review with GRL)

significantly influence the water balance on land (Findell *et al.* 2019). SSS is a standard ocean state variable that has been routinely measured back to the late 19th century. The long observational records provide an enriched data source to train and validate prediction models, a critical step towards improved operational forecasts of high-impact hydroclimatic events. Thus, sustained measurement and monitoring of SSS will greatly benefit the prediction of heavy rainfall and early warning of flooding events for the US Midwest.

References

- Carter, E., D. A. Herrera, and S. Steinschneider, 2021: Feature engineering for subseasonal to seasonal warm-season precipitation forecasts in the Midwestern US: Towards a unifying hypothesis of anomalous warm-season hydroclimatic circulation. *J. Climate*, **34**, 8291–8318.
- Doblas-Reyes, F. J., J. Garcia-Serrano, F. Lienert, A. P. Biescas, and L. R. L. Rodrigues, 2013: Seasonal climate predictability and forecasting: status and prospects. *Wires Clim. Change*, **4**, 245–268.
- Durack, P. J., S. E. Wijffels, and R. J. Matear, 2012: Ocean salinities reveal strong global water cycle intensification during 1950 to 2000. *Science*, **336**, 455–458.
- Findell, K. L., P. W. Keys, R. J. van der Ent, B. R. Lintner, A. Berg, and J. P. Krasting, 2019: Rising temperatures increase importance of oceanic evaporation as a source for continental precipitation. *J. Climate*, **32**, 7713–7726.
- Gelfand, A. E., 2000: Gibbs sampling. *J. Am. Stat. Assoc.* **95**, 1300–1304.
- Gill, A. E., 1980: Some simple solutions for the heat induced tropical circulation. *Quart. J. Roy. Meteor. Soc.*, **106**, 447–462.
- Gimeno, L., A. Drumond, R. Nieto, R. M. Trigo, and A. Stohl, 2010: On the origin of continental precipitation. *Geophys. Res. Lett.*, **37**, 1–7.
- Gordon, A. L., 2016: The marine hydrological cycle: The ocean's floods and droughts. *Geophys. Res. Lett.*, **43**, 7649–7652.
- Guimbard, S., N. Reul, B. Chapron, M. Umberto, and C. Maes, 2017: Seasonal and interannual variability of the Eastern Tropical Pacific Fresh Pool. *J. Geophys. Res.*, **122**, 1749–1771.
- Horel, J. D., and J. M. Wallace, 1981: Planetary-scale atmospheric phenomena associated with the Southern Oscillation. *Mon. Wea. Rev.*, **109**, 813–829.
- Jana, S., B. Rajagopalan, M. A. Alexander, and A. J. Ray, 2018: Understanding the dominant sources and tracks of moisture for summer rainfall in the southwest United States. *J. Geophys. Res.: Atmo.*, **123**, 4850–4870.
- Levang, S. J., and R. W. Schmitt, 2015: Centennial changes of the global water cycle in CMIP5 models. *J. Climate*, **28**, 6489–6502.
- Li, L., and W. Li, 2013: Southeastern United States summer rainfall framework and its implication for seasonal prediction. *Environ. Res. Lett.*, **8**, 044017.
- , R. W. Schmitt, C. C. Ummenhofer, and K. B. Karnauskas, 2016a: North Atlantic salinity as a predictor of Sahel rainfall. *Sci. Adv.*, **2**, e1501588.
- , —, —, and —, 2016b: Implications of North Atlantic sea surface salinity for summer precipitation over the US Midwest: Mechanisms and predictive values. *J. Climate*, **29**, 3143–3159.
- , —, and —, 2018: The role of the subtropical North Atlantic water cycle in recent US extreme precipitation events. *Clim. Dyn.* **50**, 1291–1305.
- Liu, T., R. W. Schmitt, and L. Li, 2018: Global search for autumn-lead sea surface salinity predictors of winter precipitation in southwestern United States. *Geophys. Res. Lett.*, **45**, 8445–8454.
- Liu, Y., and Coauthors, 2021: Moisture sources variation for summer rainfall in different intensity classes over Huaihe River Valley. *Clim. Dyn.*, **57**, 1121–1133.
- Liu, Z., and M. Alexander, 2007: Atmospheric bridge, oceanic tunnel, and global climatic teleconnections. *Rev. Geophys.*, **45**, RG2005.

- McCabe, G. J., M. A. Palecki, and J. L. Betancourt, 2004: Pacific and Atlantic Ocean influence on multidecadal drought frequency in the United States. *Proc. Natl. Acad. Sci.*, **101**, 4136-4141.
- McPhaden, M. J., T. Lee, and D. McClurg, 2011: El Niño and its relationship to changing background conditions in the tropical Pacific Ocean. *Geophys. Res. Lett.*, **38**, L15709.
- Rathore, S., N. L. Bindoff, C. C. Ummenhofer, H. E. Phillips, M. Feng, and M. Mishra, 2021: Improving Australian rainfall prediction using sea surface salinity. *J. Climate*, **34**, 2473–2490.
- Ray, D. K., J. S. Gerber, G. K. MacDonald, and P. C. West, 2015: Climate variation explains a third of global crop yield variability. *Nat. Commun.*, **6**, 5989.
- Scaife, A. A., and Coauthors, 2017: Tropical rainfall, Rossby waves and regional winter climate predictions. *Quart. J. Roy. Meteorol. Soc.*, **143**, 1-11.
- Schanze, J. J., R. W. Schmitt, and L. L. Yu, 2010: The global oceanic freshwater cycle: A state-of-the-art quantification. *J. Mar. Res.*, **68**, 569-595.
- Schubert, S. D., and Coauthors, 2016: Global meteorological drought: A synthesis of current understanding with a focus on SST drivers of precipitation deficits. *J. Climate*, **29**, 3989-4019.
- Sena-Martins, M., and D. Stammer, 2015: Pacific Ocean surface freshwater variability underneath the double ITCZ as seen by satellite sea surface salinity retrievals. *J. Geophys. Res.: Oceans*, **120**, 5870–5885.
- Vinogradova, N., and Coauthors, 2019: Satellite salinity observing system: Recent discoveries and the way forward. *Front. Mar. Sci.*, **22**, doi: <https://doi.org/10.3389/fmars.2019.00243>.
- White, C. J., and Coauthors, 2017: Potential applications of subseasonal-to-seasonal (S2S) predictions. *Meteorol. Appl.*, **24**, 315–325.
- Yeh, S.-W., and Coauthors, 2019: ENSO atmospheric teleconnections and their response to greenhouse gas forcing. *Rev. Geophys.*, **56**, 185–206.
- Zeng, L., R. W. Schmitt, L. Li, Q. Wang, and D. Wang, 2019: Forecast of summer precipitation in the Yangtze River Valley based on South China Sea springtime sea surface salinity. *Clim. Dyn.*, **53**, 5495–5509.

Evaluating the Potential of a Blocking Predictor in a Hybridized Dynamical-Statistical Model for Improved Week 3-4 Temperature and Precipitation Outlooks

Cory F. Baggett,^{1,2,3} Laura Ciasto,¹ Daniel Harnos,¹ Emerson LaJoie,¹ Dan Collins,¹
Muthuvel Chelliah,¹ Gregory Jennrich,^{1,2,3} Evan Oswald,^{1,2,3} Kyle MacRitchie,¹
Jon Gottschalck,¹ and Michael Halpert¹

¹*Climate Prediction Center, NOAA/NWS/NCEP, College Park, MD*

²*Earth Resources Technology, Inc., Laurel, MD*

³*Innovim, LLC, Greenbelt, MD*

1. Introduction

Atmospheric blocking has long been recognized as an important phenomenon that manifests at large, quasi-stationary anticyclones that block the storm track and reverse the typical climatological westerlies in the mid-latitudes (Rex 1950; Miller and Zhou 2019 and references therein). The impacts of blocking on both upstream and downstream temperature and precipitation anomalies are well known and are profound due to the prolonged, several week timescale over which blocking can persist, often extending into Week 3-4. For example, persistent drought and heat in California has been linked to blocking over the northeastern Pacific Ocean (*e.g.*, Wise 2016), while these same blocks can lead to atmospheric river landfalls in Alaska (Baggett *et al.* 2016). Persistent extreme cold conditions over the eastern United States have been linked to blocking over the western Atlantic (*e.g.*, Wang *et al.* 2010), and blocks are capable of steering tropical systems such as Hurricane Sandy into the East Coast in 2012 (Mattingly *et al.* 2015). Also, blocking is known to have a strong relationship with the stratosphere and sudden stratospheric warmings (Martius *et al.* 2009; Butler *et al.* 2017). Any potential forewarning of the occurrence of a sudden stratospheric warming and its subsequent impacts would be beneficial. Thus, a greater understanding of blocking and its impacts has the ability to improve subseasonal outlooks, including CPC's Week 3-4 temperature and precipitation outlooks.

2. Methods

In this work, we investigate the extent to which Week 3-4 prediction skill can be improved by adding blocking as a predictor to CPC's statistical suite of tools. In particular, we focus on improving CPC's multiple linear regression model (MLR; Harnos *et al.* 2022), whose present version uses as predictors the current state of ENSO and the MJO along with the long-term trend to forecast above or below normal outlooks of Week 3-4 temperature and precipitation. We test several formulations of blocking as a predictor, including traditional indices based on the latitudinal reversal of geopotential height gradients (Tibaldi and Molteni 1990; Barnes *et al.* 2012) along with more common indices such as the North Atlantic Oscillation (NAO) and the Pacific-North American pattern (PNA; Barnston and Livezey 1987), which are known to be associated with blocking (Crocimaspoli *et al.* 2007). Furthermore, we analyze the ability of the dynamical models, including the GEFSv12, to forecast blocking at extended leads with the purpose of using forecasted blocking predictors in a hybrid-statistical-dynamical approach. In section 3, results will focus on the use of forecasted values of the NAO and PNA, as they provided the greatest overall increase in skill compared to the other indices when being tested in the framework of the statistical MLR.

In this framework, the MLR is trained using observed data from the 1981-2010 period and verified during the 2011-2019 period. Observed values of temperature from CPC's Global Temperature data set (Fan *et al.* 2008) and precipitation from CPC's Global Unified Gauge-Based Analysis of Daily Precipitation (Chen *et al.* 2008) are used for both training and verification. Several experiments with the MLR were constructed and tested: the original-MLR, the MLR-NAO, the MLR-PNA, and the merged-MLR. The original-MLR serves as

the control experiment and mimics the MLR currently being used in operations. For predictors, the original-MLR uses the prior two week-averaged value of ENSO3.4, the long-term trend, and recent values of the Real-time Multivariate MJO index (RMM; Wheeler and Hendon 2004). The MLR-NAO uses the same predictors as the original-MLR but also incorporates the NAO. A range of lead-time-specific NAO values are tested as predictors from Day 0 through Day 15. Actual observed values of the NAO are used during the training period, whereas forecasted, ensemble-mean values of the NAO from the GEFS are used during the verification period¹. An analogous experimental design was created for the MLR-PNA. After extensively testing a range of lead-time-specific predictors, it was found that the Day +14 NAO leads to the greatest improvement in skill of the MLR-NAO over the original-MLR, while the Day +12 PNA leads to the greatest improvement in skill of the MLR-PNA over the original-MLR. The merged-MLR makes use of the MLR-NAO and MLR-PNA by using each individual model to forecast for CONUS and AK, respectively. A schematic diagram of the merged-MLR is provided in Figure 1.

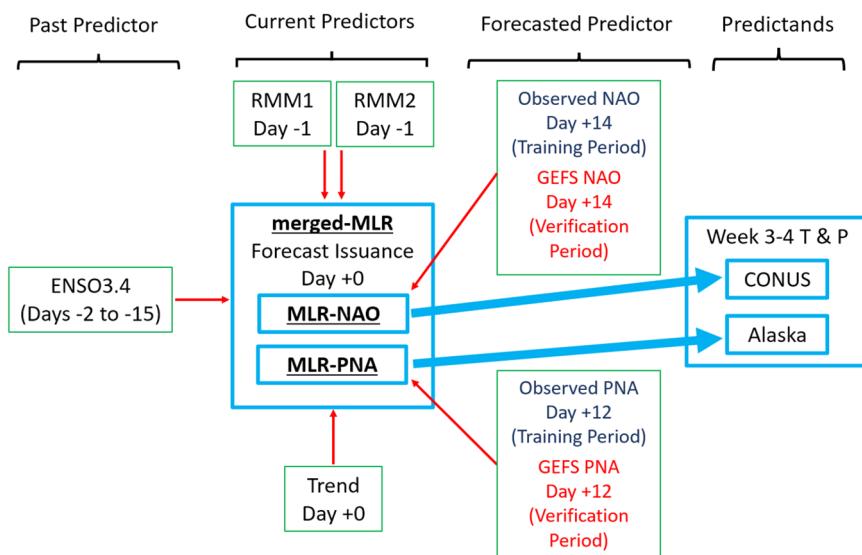


Fig. 1 A schematic diagram of the merged-MLR is shown. The merged-MLR consists of two separate models, the MLR-NAO which is used to forecast for CONUS and the MLR-PNA which is used to forecast for AK. Each of these models is trained over the 1981-2010 period with the following predictors: prior two week-averaged value of ENSO3.4, the most current RMM indices, the long-term trend, the NAO at Day +14 for the MLR-NAO, and the PNA at Day +12 for the MLR-PNA. The observed values of the NAO and PNA are used during the training period, whereas bias- and variance-corrected, ensemble mean values from the GEFSv12 are used during the 2011-2019 verification period. The predictands are anomalous values of Week 3-4 temperature and precipitation, where anomalies for temperature and precipitation are defined against their climatological means and medians, respectively.

3. Results

We verify temperature and precipitation forecasts by the aforementioned MLR models for Thursday initializations during November-April over the 2011-2019 verification period. We only focus on Thursday initializations in order to make a comparison with skill scores derived from the weekly-issued reforecasts of the extended-GEFSv12. Further, we concentrate on November-April initializations, as no improvement in skill was found during the summer months – likely a result of two factors: 1) The teleconnections of ENSO, the MJO, the NAO, and PNA to surface impacts over CONUS are weaker in summer than winter, and 2) We found that the GEFSv12 has less skill in forecasting the NAO and PNA at extended leads during summer. In fact, during winter, correlations between the observed and GEFSv12-forecasted values of the indices exceed 0.5 through lead times of +12 to +14 days, while they were substantially lower during the warmer half of the year (not shown).

An overall summary of two-category temperature and precipitation Heidke Skill Scores (HSS) for each of the MLR statistical models and the GEFSv12 is provided in Tables 1-4. While unfortunately precipitation skill scores were not improved, temperature skill scores from the merged-MLR show notable improvement over the

¹ Before being used as predictors during the verification period, the ensemble-mean values of the NAO from the GEFS have had their lead-time-specific bias and variance corrected to match the observed NAO.

Table 1 Heidke skill scores for temperature for the GEFSv12, the original-MLR, the MLR-NAO, the MLR-PNA, and the merged-MLR aggregated over the domains of CONUS/AK, CONUS, and AK that have been averaged over the verification period of January 6, 2011 through August 29, 2019 for all Thursday issuances of the Week 3-4 forecast during November through April ($n = 224$). Positive values indicate forecast skill and added value as compared to a climatological baseline.

Temperature HSS (All Initializations)

Region	GEFSv12	original-MLR	MLR-NAO	MLR-PNA	merged-MLR
CONUS & AK	24.3	12.7	14.9	10.7	17.2
CONUS	23.9	11.5	15.3	7.6	15.3
AK	26.1	19.2	12.7	26.6	26.6

Table 2 As in Table 1, but for precipitation.

Precipitation HSS (All Initializations)

Region	GEFSv12	original-MLR	MLR-NAO	MLR-PNA	merged-MLR
CONUS & AK	6.9	3.6	2.6	2.6	2.4
CONUS	6.5	2.4	1.4	1.6	1.4
AK	9.1	10.2	9.2	8.0	8.0

Table 3 As in Table 1, but averaged only over forecast issuances when the NAO was amplified on Day 0.

Temperature HSS (Amplified NAO Initializations)

Region	GEFSv12	original-MLR	MLR-NAO	MLR-PNA	merged-MLR
CONUS & AK	22.7	13.2	21.2	11.0	22.7
CONUS	21.9	11.5	21.5	7.5	21.5
AK	26.9	21.9	20.0	29.0	29.0

Table 4 As in Table 3, but for precipitation.

Precipitation HSS (Amplified NAO Initializations)

Region	GEFSv12	original-MLR	MLR-NAO	MLR-PNA	merged-MLR
CONUS & AK	8.4	4.7	5.8	3.4	5.7
CONUS	7.7	3.5	5.0	2.3	5.0
AK	11.8	10.8	10.2	9.3	9.3

original-MLR. Thus, the remaining figures will highlight the results for temperature only. Figure 2 displays the HSS over 224 forecast initializations for the original-MLR, the MLR-NAO, the MLR-PNA, the merged-MLR, and the GEFSv12 for the CONUS/AK domain. The original-MLR scores a 12.7 for the entire CONUS/AK domain. In comparison, the MLR-NAO scores 14.9, an improvement of 17.3%. Upon closer examination, this improvement derives from the MLR-NAO scoring a 15.3 over CONUS compared to the original-MLR's 11.5, an improvement of 33.0%. In contrast, the MLR-NAO's score over AK decreases. With respect to the MLR-PNA, it scores 10.7 over CONUS/AK, which is worse than the original-MLR. However, despite this decrease in skill score, the MLR-PNA improves upon the original-MLR over AK with a score of 26.6 compared to 19.2, an improvement of 38.5%. Thus, because the improved forecast skill of the MLR-NAO and MLR-PNA over the original-MLR are spatially complementary to each other, we combine them into a merged-MLR, which

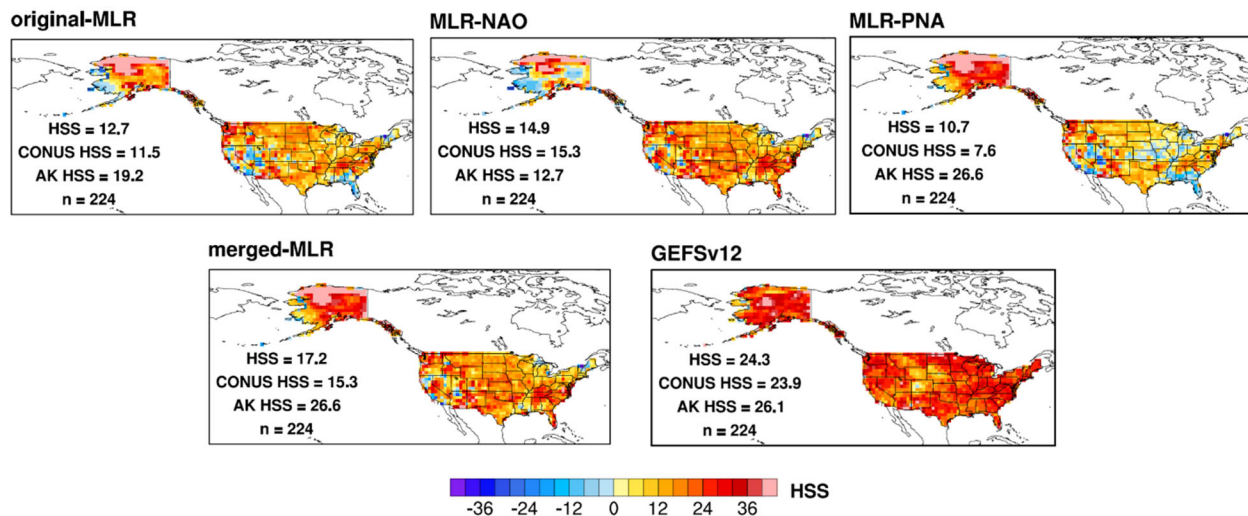


Fig. 2 HSSs for the original-MLR, the MLR-NAO, the MLR-PNA, the merged-MLR, and the GEFSv12 for all Thursday forecast initializations during November-April from 2011-2019 ($n = 224$). Scores averaged over the entire CONUS/AK domain, CONUS only, and AK only are shown for each model.

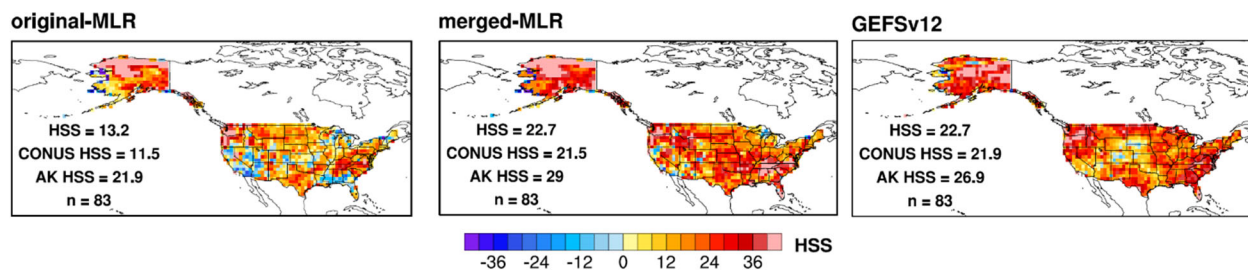


Fig. 3 As in Figure 2, but only forecast initializations with an amplified NAO are considered ($n = 83$).

scores a 17.2 over the CONUS/AK domain, an improvement of 35.4% over the original-MLR's 12.7. Finally, as a point of reference, the GEFSv12 scores a 24.3 over the CONUS/AK domain, which is 41% better than the merged-MLR.

However, the advantage of the GEFSv12 over the merged-MLR considers all forecast initializations, regardless of whether blocking is currently active or not. Indeed, this advantage vanishes when only considering forecast initializations when the NAO is amplified (either greater than 1 standard deviation or less than -1 standard deviation; Fig. 3). While the sample size of these initializations decreases to 83, both the merged-MLR and the GEFSv12 have identical HSSs of 22.7, significantly better than the original-MLR's score of 13.2, an improvement of 72%. The improved skill of the merged-MLR under amplified NAO conditions is likely due to the GEFSv12's ability to more accurately forecast the NAO at extended lead times given a strong NAO at forecast initialization (Feng *et al.* 2021). In a hybrid-statistical-dynamical framework, having an accurate forecast of the NAO at Day +14 by the GEFSv12 leads to an improved forecast of temperature over CONUS by the merged-MLR.

4. Discussion and conclusion

Given the improvement of the merged-MLR over the original-MLR, an experimental, web-based tool has been created and made available to the Week 3-4 operational forecasters at CPC². The tool provides forecasts of Week 3-4 temperature, precipitation, and 500-hPa heights derived from the original-MLR, the MLR-NAO, the MLR-PNA, and the merged-MLR along with their verifications over time. Currently, the tool is being monitored and enhancements are being made given feedback by the forecasters.

² http://cpcintradev.ncep.noaa.gov/~cbaggett/Wk34_Blocking/blocking.html

In conclusion, we list below some features of the merged-MLR that are being monitored and possibly improved upon during its experimental phase:

- 1) While the merged-MLR shows substantial improvement over the original-MLR, most of this improvement occurs during so-called forecasts of opportunity (Mariotti *et al.* 2020). For example, we showed that during amplified NAO conditions, the merged-MLR outperforms the original-MLR by 72.0% and is on-par with the GEFSv12. Thus, the greatest value of the merged-MLR is likely limited to less than one half of forecast initializations during November-April.
- 2) Precipitation skill remains elusive by both statistical and dynamical models. By its nature, precipitation is both temporally and spatially noisy with highly skewed and ill-defined distributions. Thus, linearly fitting non-Gaussian precipitation anomalies to Gaussian predictors within the MLR framework is a challenging statistical endeavor. To tackle this challenge, the experimental tool has been designed to test additional precipitation forecasting schemes as ideas emerge.
- 3) The method via which CPC's teleconnection indices are calculated are undergoing an extensive review at CPC. From this review, it is possible that the method will change. At this time, it is unknown what impact this change will have on the merged-MLR.
- 4) CPC's Week 3-4 outlooks will begin transitioning from a two-category forecast to a three-category forecast over the coming years. The merged-MLR will need to be reformulated to accommodate this change.
- 5) Finally, while this project was conceived to identify "blocking" indices, we settled on using the NAO and PNA as predictors given the enhanced skill they provided over more traditional blocking indices. While both the NAO and PNA are known to be related to blocking (*e.g.*, the negative NAO is related to North Atlantic blocking) they are continuous indices and represent a range of atmospheric flows, some of which constitute the complete absence of blocking. Regardless, using continuous indices such as the NAO and PNA are useful as they facilitate the ability to forecast at each initialization rather than when a rare, strong signal, such as blocking, emerges.

References

- Baggett, C., S. Lee, and S. Feldstein, 2016: An investigation of the presence of atmospheric rivers over the North Pacific during planetary-scale wave life cycles and their role in Arctic warming. *J. Atmos. Sci.*, **73**, 4329–4347, doi:10.1175/JAS-D-16-0033.1.
- Barnes, E. A., J. Slingo, and T. Woollings, 2012: A methodology for the comparison of blocking climatologies across indices, models and climate scenarios. *Clim. Dyn.*, **38**, 2467–2481, <https://doi.org/10.1007/s00382-011-1243-6>.
- Barnston, A., and R. Livezey, 1987: Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon. Wea. Rev.*, **115**, 1083–1126, doi:10.1175/1520-0493(1987)115<1083:CSAPOL>2.0.CO;2.
- Butler, A. H., J. P. Sjöberg, D. J. Seidel, and K. H. Rosenlof, 2017: A sudden stratospheric warming compendium. *Earth Syst. Sci. Data*, **9**, 63–76, doi:10.5194/essd-9-63-2017.
- Chen, M., and Coauthors, 2008: Assessing objective techniques for gauge-based analyses of global daily precipitation. *J. Geophys. Res.*, **113**, D04110, doi:10.1029/2007JD009132.
- Croci-Maspoli, M., C. Schwierz, and H. C. Davies, 2007: Atmospheric blocking: space-time links to the NAO and PNA. *Clim. Dyn.*, **29**, 713–725, doi:10.1007/s00382-007-0259-4.
- Fan, Y., and H. van den Dool, 2008: A global monthly land surface air temperature analysis for 1948–present. *J. Geophys. Res.*, **113**, D01103, doi:10.1029/2007JD008470.
- Feng, P.-N., J. Derome, and T. Merlis, 2021: Forecast skill of the NAO in the subseasonal-to-seasonal prediction models. *J. Climate*, **34**, 4757–4769, doi:10.1175/JCL-D-20-0430.1.
- Harnos, D., N. Johnson, L. Ciasto, A. Allgood, S. Baxter, and M. L'Heureux, 2022: The competitiveness of empirical models in subseasonal temperature and precipitation prediction over the United States. *Wea. Forecasting*, in revision.

- Mariotti, A., and Coauthors, 2020: Windows of opportunity for skillful forecasts subseasonal to seasonal and beyond. *Bull. Amer. Meteor. Soc.*, **101**, E608–E625, doi:10.1175/BAMS-D-18-0326.1.
- Martius, O., L. M. Polvani, and H. C. Davies, 2009: Blocking precursors to stratospheric sudden warming events. *Geophys. Res. Lett.*, **36**, doi:10.1029/2009GL038776.
- Mattingly, K. S., J. T. McLeod, J. A. Knox, J. M. Shepherd, and T. L. Mote, 2015: A climatological assessment of Greenland blocking conditions associated with the track of Hurricane Sandy and historical North Atlantic hurricanes. *Int. J. Climatol.*, **35**, 746–760, doi:10.1002/joc.4018.
- Miller, D. E., and Z. Wang, 2019: Skillful seasonal prediction of Eurasian winter blocking and extreme temperature frequency. *Geophys. Res. Lett.*, **46**, 11530–11538, doi:10.1029/2019GL085035.
- Rex, D. F., 1950: Blocking action in the middle troposphere and its effect upon regional climate. *Tellus*, **2**, 275–301, doi:10.3402/tellusa.v2i4.8603.
- Tibaldi, S., and F. Molteni, 1990: On the operational predictability of blocking. *Tellus*, **42**, 343–365, doi:10.1034/j.1600-0870.1990.t01-2-00003.x.
- Wang, C., H. Liu, and S.-K. Lee, 2010: The record-breaking cold temperatures during the winter of 2009/2010 in the Northern Hemisphere. *Atmos. Sci. Lett.*, **11**, 161–168, doi:10.1002/asl.278.
- Wheeler, M. C., and H. H. Hendon, 2004: An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev.*, **132**, 1917–1932, doi:10.1175/1520-0493(2004)132<1917:AARMMI>2.0.CO;2.
- Wise, E. K., 2016: Five centuries of U.S. West Coast drought: Occurrence, spatial distribution, and associated atmospheric circulation patterns. *Geophys. Res. Lett.*, **43**, 4539–4546, doi:10.1002/2016GL068487.

Dynamical Weighting of the Week 3-4 Models based on Forecasts of Opportunity

Cory F. Baggett,^{1,2,3} Emerson LaJoie,¹ Daniel Collins,¹ Johnna Infanti,¹ and Peitao Peng¹

¹Climate Prediction Center, NOAA/NWS/NCEP, College Park, MD

²Earth Resources Technology, Inc., Laurel, MD

³Innovim, LLC, Greenbelt, MD

1. Introduction

This work seeks to improve the Climate Prediction Center's (CPC) week 3-4 temperature and precipitation outlooks using a new forecast tool based on a dynamically-weighted blend of the subseasonal dynamical models conditioned on climatic states at their initialization. It has been documented that dynamical models have forecast skill that varies as a function of grid point and season along with climatic states at their initialization such as the El Niño-Southern Oscillation (ENSO; Yamagami and Matsueda 2020) and the Madden-Julian Oscillation (MJO; Vitart *et al.* 2017). When climatic conditions arise such that one model has performed historically better than another, this can be considered a forecast of opportunity (Mariotti *et al.* 2020). In our experiment, we seek to identify and take advantage of these forecasts of opportunity to produce dynamically-weighted forecast blends that not only improve upon the individual dynamical models but also improve upon a static, equally-weighted blend.

2. Methods

Figure 1 provides a brief description of the blends and the dynamical models that are used in this experiment. As a control, we create a static, equally-weighted blend where each model is given the same weight regardless of grid point, season, or climatic state at initialization. In our experiment, we test three dynamically-weighted blends: 1) a seasonal blend, where the models are weighted as a function of grid point and season at initialization, 2) a seasonal-ENSO blend, where the models are weighted as a function of grid point, season, and ENSO state at initialization, and 3) a seasonal-ENSO-MJO blend, where the models are weighted as a function of grid point, season, ENSO, and MJO state at initialization.

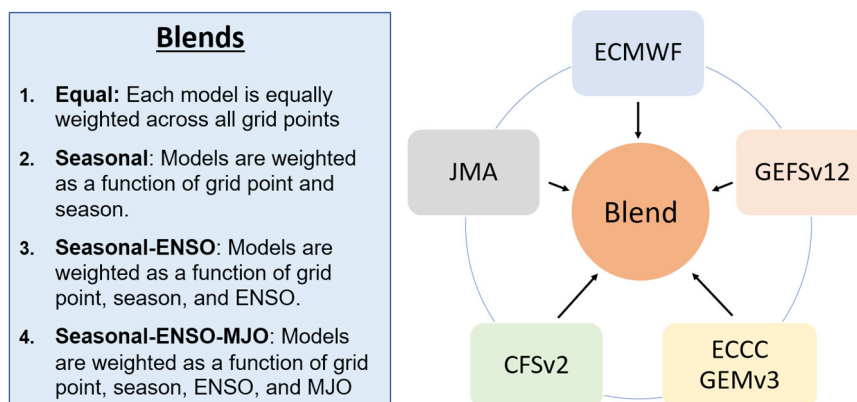


Fig. 1 A schematic diagram showing the dynamical models and the blends that are derived from them.

We derive the weights to be used in the blends from the hindcasts of the dynamical models that CPC uses operationally in real-time: the ECMWF, the GEFSv12, the ECCG, the CFSv2, and the JMA. In the hindcasts, we calculate the two-category Heidke Skill Score (HSS) for each model's week 3-4 probabilistic forecast of temperature and precipitation against observed values of temperature from CPC's Global Temperature data set (Fan *et al.* 2008) and precipitation from CPC's Global Unified Gauge-Based Analysis of Daily Precipitation (Chen *et al.* 2008). Having calculated the HSSs of the models in the hindcasts, these HSS values are then normalized and used to weight the probabilistic forecasts from the dynamical models in real-time to create the

dynamically-weighted blends¹. These real-time forecast blends are then verified against observations to determine their skill relative to the equally-weighted blend and the individual models.

In Section 3, we will provide an example of a real-time week 3-4 temperature forecast issued on February 18, 2022 to demonstrate the weights and the creation of the blends. We will also provide bulk verification statistics for both temperature and precipitation forecasts issued weekly on Fridays from November 6, 2020 through February 18, 2022, totaling 68 forecasts.

3. Results

Figure 2 provides the weights used in the seasonal-ENSO-MJO blend for the week 3-4 temperature forecast issued on February 18, 2022. This particular forecast was issued during January-February-March, during ongoing La Niña conditions, and during an active MJO period in phases 2-3-4. The weights used correspond to the historical, hindcast performance of the models under these conditions. For example, the ECMWF scores well in Alaska and the Central Plains where the JMA does not. Also, the GEFSv12 and ECCC score relatively well from Texas to the Midwest. The real-time Week 3-4 probabilistic forecasts from each dynamical model are provided in Fig. 3. In general, the ECCC, the ECMWF, the GEFSv12, and the CFSv2 were forecasting

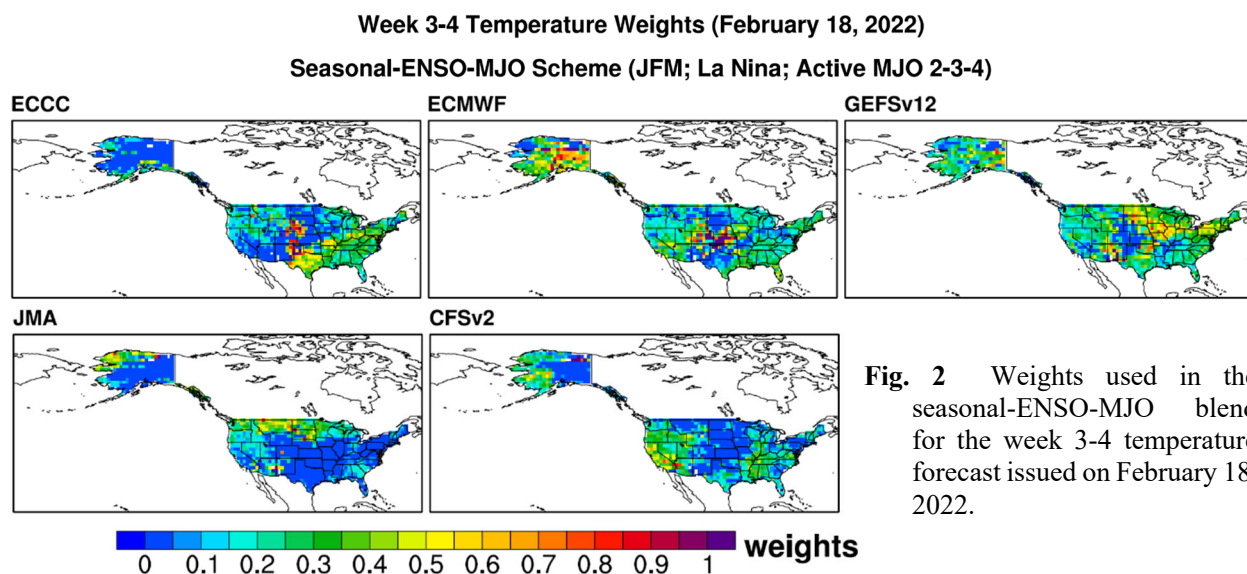


Fig. 2 Weights used in the seasonal-ENSO-MJO blend for the week 3-4 temperature forecast issued on February 18, 2022.

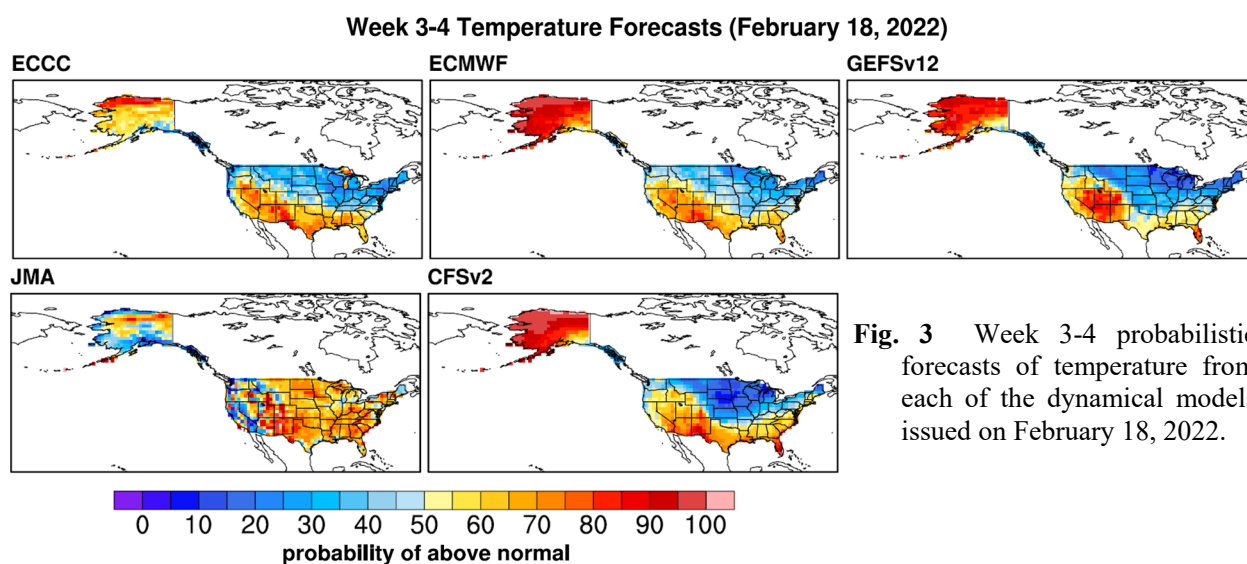


Fig. 3 Week 3-4 probabilistic forecasts of temperature from each of the dynamical models issued on February 18, 2022.

¹ All HSS values that fall below zero are set to zero before calculating the weights.

warmth across Alaska and cold in CONUS across the Northern Plains into the Midwest and Northeast. However, the JMA was not as confident in this depiction, having more cold across Alaska and warmth in CONUS. Applying the weights from Fig. 2 directly to the probabilistic forecasts from Figure 3, one can derive the bottom right panel of Fig. 4, the seasonal-ENSO-MJO blend. The equally-weighted, seasonal, and seasonal-ENSO blends are shown in the remaining panels. In this example, each of the blends looks similar to each other because of the overall consistency amongst four of the five dynamical models. However, in the seasonal-ENSO-MJO blend, Alaska and the Central Plains resemble more closely the ECMWF, which is to be expected given ECMWF's higher weighting in those regions.

Figure 5 provides average HSSs calculated from CONUS and AK for each forecast issuance date from November 6, 2020 through February 18, 2022. In general, the skill scores of the blends are very similar across all forecasts, with skill scores generally falling within an ~15 point range from each other for any given forecast issuance date. When averaged across all forecast issuance dates, the equally-weighted blend has a score of 26.8, which slightly outperforms the dynamically-weighted blends. Thus, a null hypothesis that the equally-weighted blend has higher skill scores than the dynamically-weighted blends cannot be rejected at this point.

Table 1 provides a summary of skill scores for both temperature and precipitation across each of the five dynamical models and the four blends. For temperature, each blend outperforms all of the individual dynamical models except for the GEFSv12, which has a score of 28.1. For precipitation, the equally-weighted blend scores 14.6, which is higher than the dynamically-weighted blends.

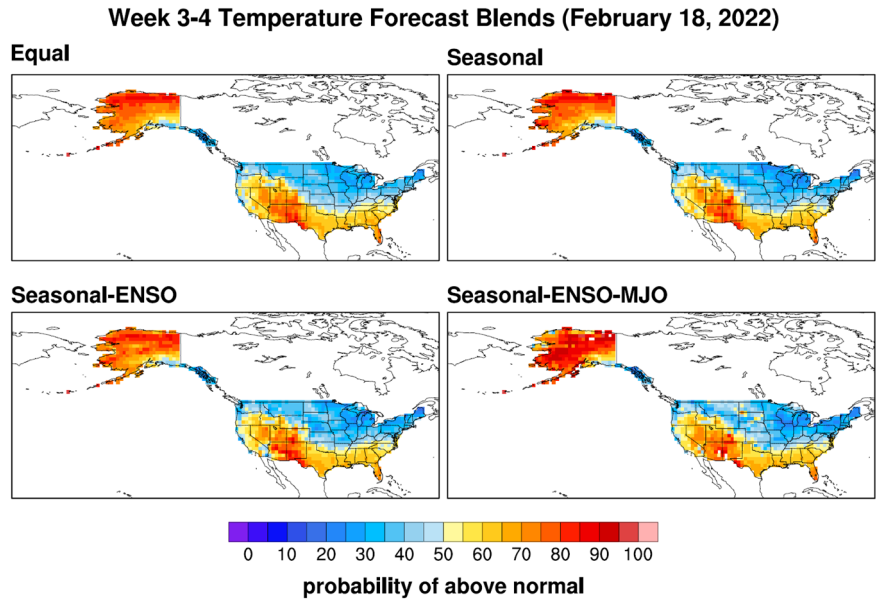


Fig. 4 Week 3-4 probabilistic forecasts of temperature from the equally-weighted and the dynamically-weighted blends issued on February 18, 2022.

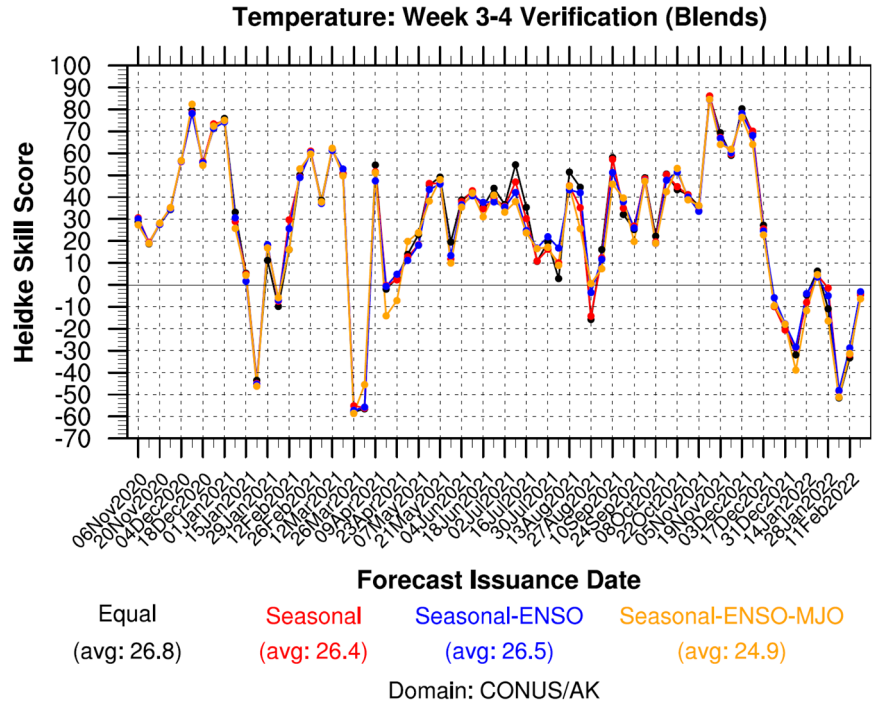


Fig. 5 HSSs of the week 3-4 temperature forecasts from the equally-weighted and dynamically-weighted blends for each forecast issuance date, issued weekly on Fridays from November 6, 2020 to February 18, 2022. HSSs averaged across all forecast issuance dates are provided beneath the figure.

Table 1. HSSs for temperature and precipitation for each of the five dynamical models and each of the blends, averaged across 68 Week 3-4 forecasts issued weekly on Fridays from November 6, 2020 to February 18, 2022.

		Temperature	Precipitation
Dynamical Model	ECCC	16.4	9.5
	ECMWF	24.5	12.6
	GEFSv12	28.1	12.5
	JMA	14.6	8.5
	CFSv2	23.1	10.7
Blend	Equal	26.8	14.6
	Seasonal	26.4	13.1
	Seasonal-ENSO	26.5	12.7
	Seasonal-ENSO-MJO	24.9	13.5

However, in contrast to temperature, each precipitation blend outperforms all of the individual dynamical models.

4. Discussion and conclusion

In this work, we have constructed dynamically-weighted week 3-4 forecasts blends from the real-time probabilistic forecasts of five dynamical models and their corresponding, climate-specific weights derived from their hindcasts. In general, the dynamically-weighted blends outperform the individual models, while they underperform the equally-weighted blend. However several major caveats must be made:

- 1) The real-time sample size is quite small with only 68 forecast issuances. Such a small sample size makes it challenging to determine if a given blend has a statistically significant difference in score compared to the individual dynamical models or the other blends. The difficulty of such a significance test is exacerbated by a large amount of autocorrelation due to the prevalence of one major climatic state, La Niña, during the real-time period.
- 2) The hindcast sample sizes are quite small. For example, the JMA has hindcasts that are issued only twice per month over a 30 year period. One can imagine that conditioning on season, ENSO, and MJO will lead to untenably small sample sizes for the JMA, thus making its weights likely dependent on just a handful of hindcasts.
- 3) Each dynamical model has a different hindcast period with different initialization dates. Because of these differences, it is difficult to make a direct comparison of the skill amongst the models given a particular climate state because we cannot rule out the impact of the initial, non-climate related conditions on the eventual hindcast skill.
- 4) Related to the prior point, in this study, we made an assumption that when a model was more skillful than another model during a certain climatic state, the skill is attributable to the climate state. In fairness, this assumption may not be wholly valid and it should be noted that this assumption was not tested in this study. However, we can offer an interpretation of the results, namely, the fact that the equally-weighted blend scores higher may reflect that the sources of skill in subseasonal forecasting are not solely resting on large-scale climate signals such as ENSO or the MJO. There are "weather signals" in subseasonal forecasting and the results of this study could be indicating that these signals need to be explored further for attributable sources of skill.
- 5) During the real-time period, new model versions for the ECCC and JMA have been released. The ECMWF has also had regular minor updates with a major update occurring during the period as well. Thus, new weights must be derived each time a new model version is released. In this experiment, we

have not derived new weights on-the-fly, opting to use the weights derived from the earlier versions of the models, which may not be optimal.

- 6) As demonstrated by Fig. 4 and implied by Fig. 5, the probabilistic forecasts from the dynamically-weighted blends are nearly indistinguishable from each other. This results from either the dynamical models themselves having similar forecasts or from the weights being nearly identical across models. Regardless, a weighting scheme that actually produces discernibly different forecasts would likely be more desirable to the operational forecaster.

Ideally, the above caveats will be addressed in a future iteration of this work in order to gain more confidence in whether or not a dynamically-weighted blend based on forecasts of opportunity can outperform an equally-weighted blend. However, initial results are encouraging that the dynamically-weighted blends can outperform the individual dynamical models. Given that CPC's operational forecaster is presented with a large amount of information to digest when creating the official week 3-4 outlook, skillful tools that can condense model information into a more digestible quantity are greatly desired.

References

- Chen, M., and Coauthors, 2008: Assessing objective techniques for gauge-based analyses of global daily precipitation. *J. Geophys. Res.*, **113**, D04110, doi:10.1029/2007JD009132.
- Fan, Y., and H. van den Dool, 2008: A global monthly land surface air temperature analysis for 1948-present. *J. Geophys. Res.*, **113**, D01103, doi:10.1029/2007JD008470.
- Mariotti, A., and Coauthors, 2020: Windows of opportunity for skillful forecasts subseasonal to seasonal and beyond. *Bull. Amer. Meteor. Soc.*, **101**, E608–E625, doi:10.1175/BAMS-D-18-0326.1.
- Vitart, F., 2017: Madden-Julian oscillation prediction and teleconnections in the S2S database. *Q. J. R. Meteorol. Soc.*, **143**, 2210–2220, doi:10.1002/qj.3079.
- Yamagami, A., and M. Matsuda, 2020: Subseasonal forecast skill for weekly mean atmospheric variability over the Northern Hemisphere in winter and its relationship to midlatitude teleconnections. *Geophys. Res. Lett.*, **47**, e2020GL088508, doi:10.1029/2020GL088508.

Evaluation of Subseasonal Arctic Sea Ice Hindcasts in an NCEP's UFS-based System

Yanyun Liu,^{1,2} Wanqiu Wang,¹ Weiyu Yang,^{1,2} Jieshun Zhu,¹ Arun Kumar,¹ and David DeWitt¹

¹Climate Prediction Center, NOAA/NWS/NCEP, College Park, MD

²Earth Resources Technology, Inc., Laurel, MD

1. Introduction

In support of NOAA's sea ice forecasts in week 3-4 time-range, the Climate Prediction Center (CPC) has been using an experimental sea ice prediction system, CFSm5, to provide weekly and seasonal Arctic sea ice predictions. The CFSm5 was developed based on the Climate Forecast System (CFS) with the Modular Ocean Model version 5 (MOM5) as oceanic component. Sea ice forecasts from CFSm5 initialized from CPC sea ice system (CSIS) have been shown to significantly improve over that from the operational CFS. In 2021, CPC started to prepare a transition from the use of CFSm5 to a new GFDL Finite-Volume Cubed-Sphere Dynamical Core (FV3) based Unified Forecast System (UFS) framework. In this work, we evaluate the multi-week Arctic sea ice forecast skill of the UFS system during the melt season and investigate the impacts of cloud related parameters for an improved representation of sea ice in the coupled system. Analysis of a suite of retrospective 45-day forecasts with four ensemble members spanning 2012-2020 shows the prediction skill of multi-week Arctic sea ice is generally comparable to CFSm5, and better than the operational CFSv2 hindcasts. The skill comparisons of CFSm5, CFSv2, UFS and Multi-Model Ensemble (MME) hindcasts are also presented.

2. The coupled UFS model

The coupled UFS Subseasonal to Seasonal model (S2S) model in this study consists of an FV3 atmospheric component, a MOM6 oceanic component and a CICE6 sea ice component. All model components are coupled using the Community Mediator for Earth Prediction Systems (CMEPS) infrastructure. The UFS version we use is Prototype 5 (UFS-P5). The atmospheric model has a horizontal resolution of 1degree (C96) and includes 64 vertical levels. The ocean and sea ice model resolutions are 0.25°. The UFS-P5 uses atmospheric initial conditions from the Climate Forecast System Reanalysis (CFSR, Saha *et al.* 2010). The ocean and sea ice initial conditions come from CSIS, which assimilates NASA Team sea ice concentration (SIC) and National Centers for Environmental Information (NCEI) sea surface temperature (SST) analyses. The UFS-P5 hindcasts are

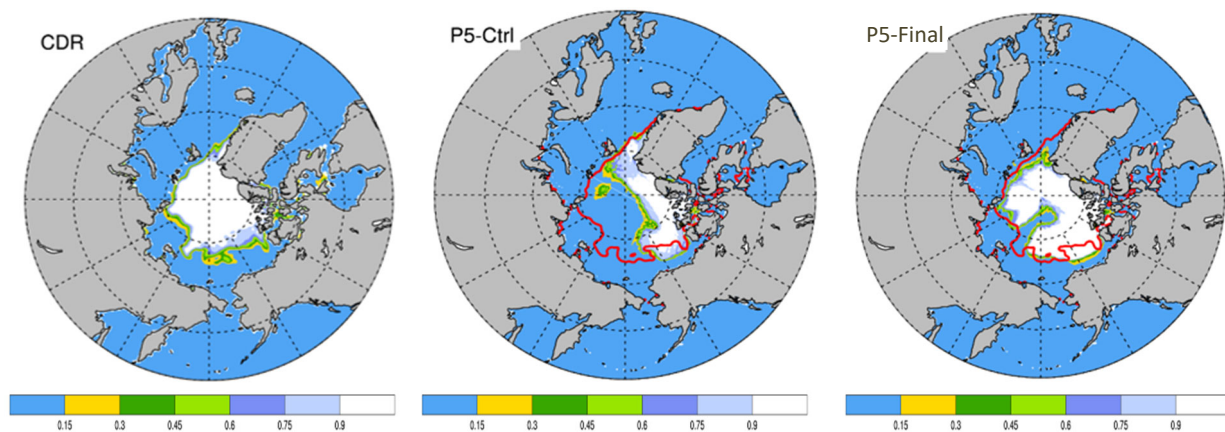


Fig.1 September SIC hindcasts initialized from June 1, 2015 for UFS-P5 (Control run and Final setting with cloud parameters adjusted and using constant freezing temperature) and CFSm5 compared with NSIDC CDR SIC.

initialized from each day of April-September, 2012-2020 and cover 45 target days with four ensemble members. The sea ice prediction skill of UFS-P5 is compared with that of CFSm5, CFSv2 and MME hindcasts. The National Snow and Ice Data Center (NSIDC) Climate Data Record (CDRv4) sea ice concentration (SIC) is used for verification.

3. Initial comparison and parameter adjustments

Figure 1 compares a 4-month hindcast of September SIC from UFS-P5 using the default configuration (Ctrl) with the observational analyses from NSIDC CDR SIC. The hindcasts were initialized from June 1, 2015. As shown in Fig. 1, there is a large negative sea ice bias in parts of the central Arctic with the default configuration (P5-Ctrl). Further analysis indicated that this bias is related to the positive downward SW radiation bias due to the negative cloud fraction bias in the central Arctic (not shown). To reduce this negative sea ice bias, three cloud parameters are adjusted, including the critical cloud drop radius (r_{thresh}), cloud condensation nuclei over ocean (ccn_o) and auto conversion coefficient from cloud water to rain (c_{paut}). A series of twenty-three experiments were performed for different combinations of these three cloud parameters. The three adjusted parameter values for r_{thresh} , ccn_o , and c_{paut} are selected as $12.0 \mu\text{m}$, 120.0 cm^{-3} , and 0.45 , compared to $10.0 \mu\text{m}$, 100 cm^{-3} , and 0.5 in the Ctrl configuration, respectively. Additionally, the constant freezing temperature option is used in the CICE6 to further reduce the bias. With those parameter adjustments, September sea ice cover is largely improved compared with the Ctrl run (Fig. 1).

The downward shortwave radiation (DSW) is further examined for the UFS-P5 experiments. As shown in Fig. 2, the positive DSW bias in P5-Final is reduced around the Bering Strait

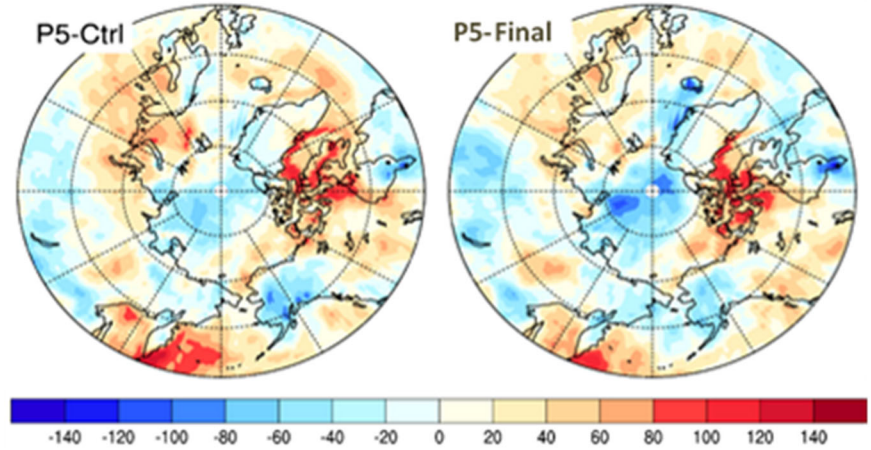


Fig. 2 June Downward shortwave radiation bias relative to EBAF for P5-Ctrl and P5-Final hindcasts (Jun 1, 2015, initial conditions)

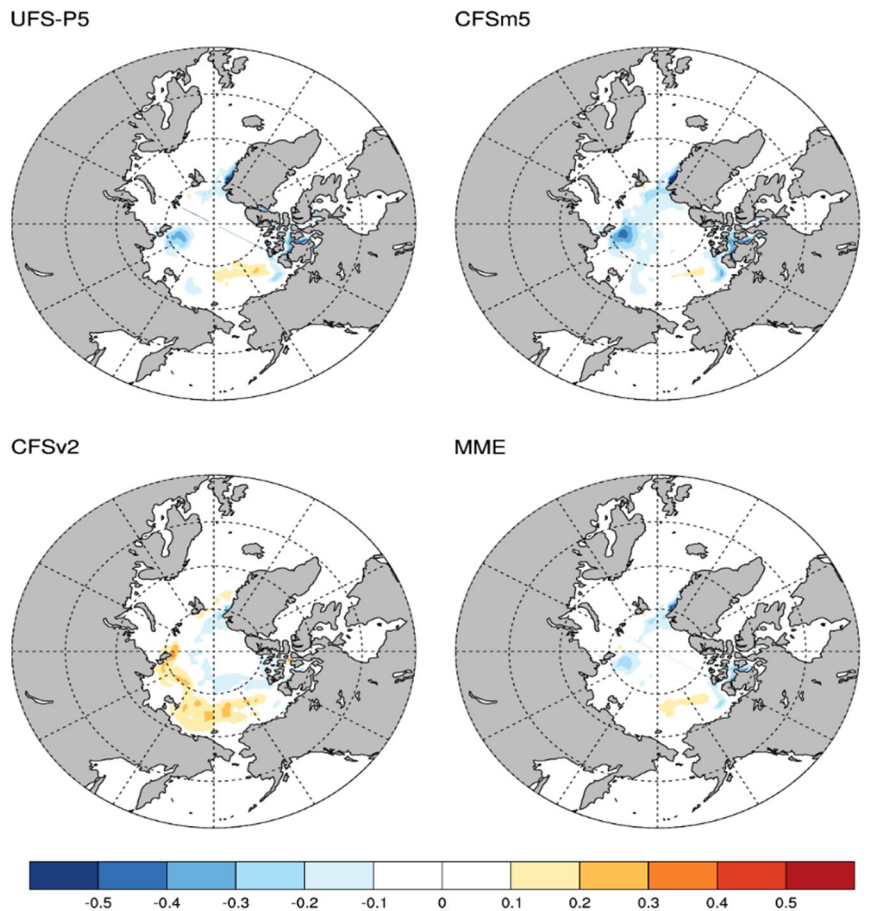


Fig. 3 Climatology week 4 SIC mean bias (initial dates = Aug 1-31, 2012-2020) for UFS-P5, CFSm5, CFSv2 and MME hindcasts.

and the North Pacific compared with the P5-Ctrl run. The error reductions are confirmed in the experiments for other years (not shown). Therefore, the UFS-P5 configuration with cloud parameter adjustments and constant water freezing temperature (P5-Final) is selected for the final configuration for UFS 45- day hindcasts.

4. Sea ice prediction skill assessment

The mean bias of UFS-P5 hindcasts is compared with that of CFSm5, CFSv2 and MME hindcasts. As shown in Fig. 3, for August 1-31 2012-2020 initial conditions, the climatology ensemble mean week 4 SIC in UFS-P5 and CFSm5 have comparable mean bias. There is relatively larger bias in CFSv2 hindcasts. The MME hindcasts (ensemble mean of UFS-P5, CFSm5 and CFSv2) has smallest bias. The UFS-P5 hindcasts initialized from other months suggest the similar conclusions (not shown).

The Heidke skill score (HSS) is used to assess the Arctic sea ice forecast performance. The HSS is calculated based on the forecast of presence or absence of sea ice. Sea ice is considered to exist in the forecast or observation if the SIC is greater than 15%. The HSS is defined as:

$$HSS = \frac{AC - AC_e}{AT - AC_e}$$

where AC is the total area of correct forecast, AC_e the total area of expected correct forecast based on observed climatology, and AT the total area of all grid boxes being considered. It is shown that for the melt seasons, there is comparable skills for the UFS-P5 and CFSm5 (Fig.4). The UFS-P5 and CFSm5 have significantly higher skills than the operational CFSv2 hindcasts and persistent forecasts. MME hindcasts generally have the highest skill.

For winter seasons, the Arctic sea ice cover in UFS-P5 is much closer to the observed estimates than CFSm5, especially around the Bering Sea, and in Atlantic (Fig.5).

5. Summary and discussions

There are biases in the UFS-P5 control configuration in the downward shortwave radiation and cloud fraction, causing less sea ice coverage in the central Arctic during the boreal summer. The adjustments to the three cloud parameters and the use of constant freezing temperature reduce the model bias in terms of DSW and SIC. The selected configuration (P5-Final) shows comparable or better performance than CFSm5 for

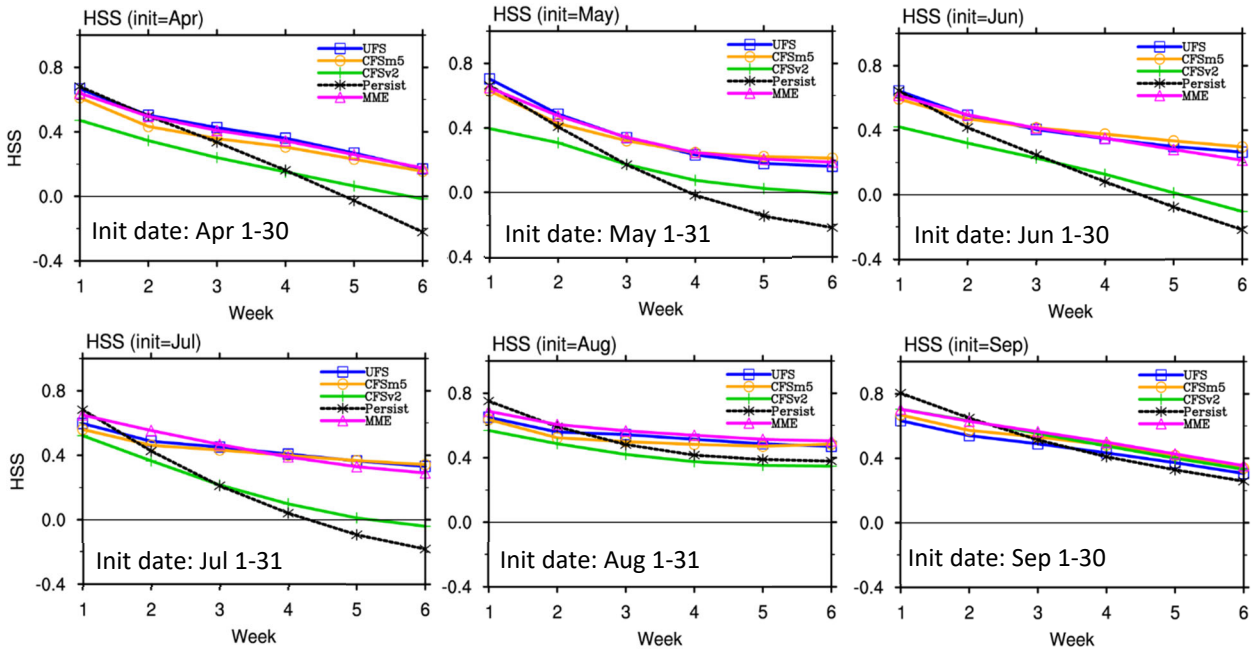


Fig. 4 Arctic sea ice Heidke Skill Score from 1 week to 6 weeks for selected initial dates. Each panel is for the initial date from May 1, June 1, July 1, August 1, September 1 and November 1 of 2012-2019.

selected initial dates for sea ice melt/freeze up seasons. In particular, a reduction in SIC mean bias in Bering Sea and Atlantic for winter seasons is found. We will continue to perform 45-day hindcasts from 2012-2020 for sea ice freeze-up seasons, and compare with CFSm5 and CFSv2 hindcasts. We will also develop the bias correction algorithms (*e.g.* mean bias correction or cumulative distribution function mapping) for UFS based real-time sea ice weekly forecasts. The current CSIS only assimilates the observed NASA Team SIC and the NCEI SST. Additional information of observational estimates of NSIDC CDRv4 SIC and Operational Sea Surface Temperature and Ice Analysis (OSTIA) SST may provide more accurate initial sea ice and ocean conditions for the sea ice predictions.

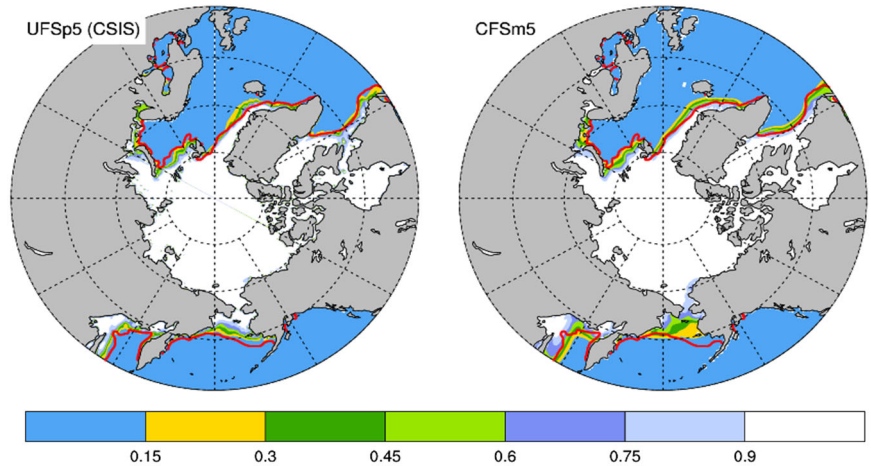


Fig. 5 Climatology month-3 SIC (January) initialized from Nov01 2012-2020. Left: UFS-P5. Right: CFSm5. The 2012-2020 January climatology of 15% CDR SIC is contoured in red.

References

Saha, S., S. Moorthi, H. Pan, and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015-1067, doi: 10.1175/2010BAMS3001.1.

Regional and Global Climate Drivers of Marine Heat Waves and Related Atmosphere-Ocean Anomalies in the Eastern North Pacific

Mitch Porter,¹ Katie Kohlman,² Seth Madden,³ and Tom Murphree⁴

¹*Johns Hopkins University, Baltimore, MD*

²*Pennsylvania State University, State College, PA*

³*Santa Clara University, Monterey, CA*

⁴*Naval Postgraduate School, Monterey, CA*

ABSTRACT

Persistent sea surface temperature anomalies in the eastern North Pacific (ENP) have strong impacts on marine ecosystems, regional climate, and public health. We investigated negative and positive sea surface temperature anomalies events in the ENP, which we refer to as cool and warm events. We studied 17 cool and 17 warm events (the lower and upper tercile events) during 1970-2020, with a focus on summer events. Our objectives were to: (1) characterize the spatial and temporal anomalies prior to and during cool and warm events; and (2) determine the regional and global processes involved in generating sea surface temperature anomalies in the ENP. The processes that lead to cool events are approximately opposite to those leading to warm events, and appear to be driven by: 1) sea surface temperatures and atmospheric convective anomalies in the tropical Indian-Pacific oceans region; 2) atmospheric wave trains that teleconnect the tropics to the ENP; and 3) resulting anomalous dipoles in ENP sea level pressure and wind forcing of the upper ocean.

1. Introduction

Marine heat waves (MHWs) and their impacts have been increasing in frequency and intensity globally in the last several decades (*e.g.*, Holbrook *et al.* 2019; Hayashida *et al.* 2020; Kohlman *et al.* 2020; Laufkötter *et al.* 2020; Sen Gupta *et al.* 2020). It is important to understand the dynamical processes leading to cool and warm because they impact marine ecosystems, nutrient transport, regional climate, and public health (*e.g.*, Amaya *et al.* 2016; Rogers-Bennett and Catton 2019; Smale *et al.* 2019; Holbrook *et al.* 2020). We studied 17 periods of persistent anomalously cool SSTs and 17 periods of persistent anomalously warm SSTs in the eastern North Pacific (ENP) during June-August 1970-2020 (Fig. 1), including some events identified as MHWs in prior studies (*e.g.*, Bond *et al.* 2015, Amaya *et al.* 2016, 2020). We refer to the periods with persistent summer negative (positive) SSTAs in the ENP as cool (warm) events. We examined both regional and global atmosphere-ocean anomalies to characterize the set of processes that lead to cool and warm events.

2. Data and methods

The primary data used were monthly mean values of atmospheric and oceanic variables from the NCEP/NCAR Reanalysis (R1; Kalnay *et al.* 1996) for 1970-2020 at a 2.5-degree resolution. Our focus study area in the ENP is located between 43 - 53°N and 215 - 228°W (black box in Fig. 2). The

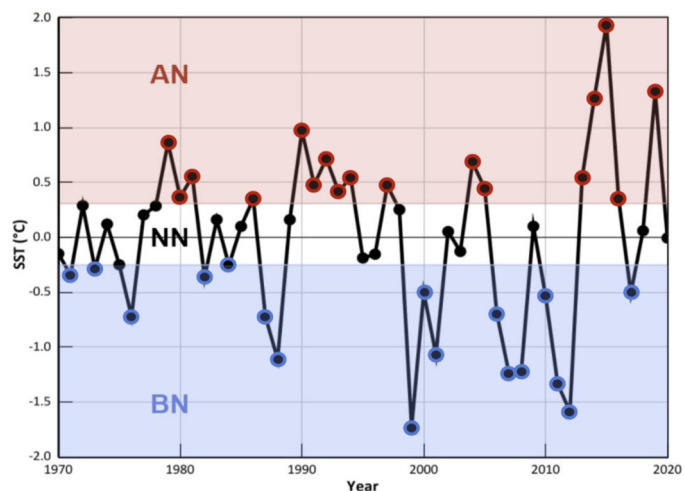


Fig. 1 Detrended SST anomalies (SSTAs; C°) for Jun - Aug 1970-2020 in the ENP box. Cool (warm) events are identified by the blue (red) shading.

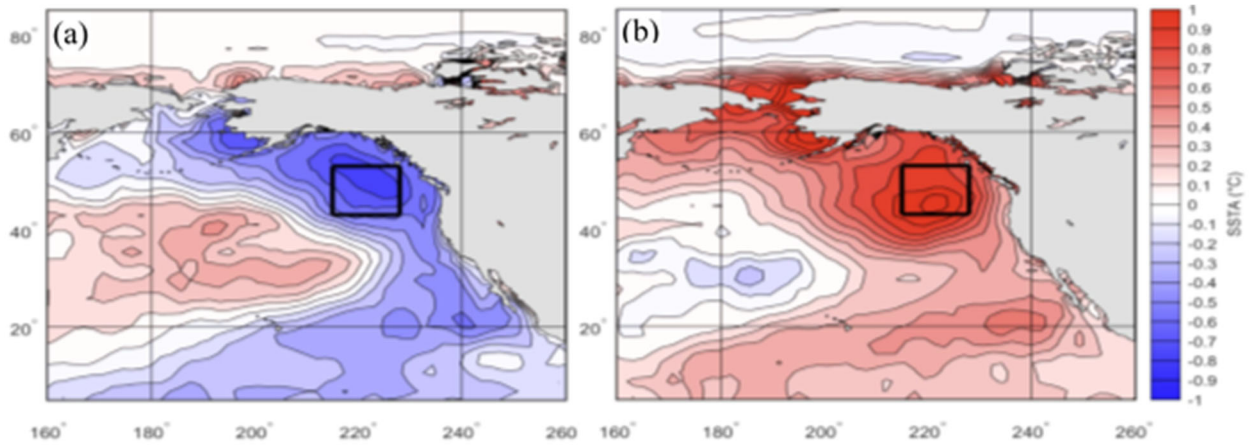


Fig. 2 Composite SSTAs (°C) for the (a) 17 coolest events and (b) the 17 warmest events during June-August 1970-2020. The black box shows our study region within the ENP (43-53°N, 215-228°E). This box encompasses the largest magnitude SSTAs for both ENP cool events and warm events.

location of this area was optimized to represent the most common location for extremes in SSTAs in the ENP. Cool (warm) events were characterized as the 17 coolest (warmest) June, July, and August (JJA) periods in the ENP study box based on detrended SSTAs between 1970 and 2020 (Fig. 1).

3. Results

Figures 1 shows that, after removing the multidecadal trend, the magnitude of the SSTAs associated with both cool and warm events increased from 1970 to 2020. Figure 2 shows that cool and warm events have spatial patterns that are approximately opposite over most of the North Pacific, with the most extreme SSTAs occurring within the ENP study box for both events.

Figure 3a (Figure 3b) shows the composite sea level pressure anomalies (SLPAs) in December - May prior to JJA of the 17 cool (warm) event years. Winter and spring SLPAs for cool and warm events have opposite north-south dipoles that are centered on the ENP box. The SLPa gradients for cool (warm) events indicate increased (decreased) eastward winds in the ENP focus area compared to the long term mean eastward flow in the winter and spring (Peixoto and Oort 1992). Increased (decreased) eastward wind flow for cool (warm) events located to the east and through the ENP box creates optimal conditions for 1) anomalously strong (weak) sensible and latent heat fluxes from the ocean and 2) increased (decreased) wind driven ocean mixing. Both of these processes are favorable for the development of negative (positive) SSTAs, corresponding to cool (warm) events.

Not shown are the conditional composites of global 200 mb eddy geopotential height anomalies (ZA200) for the winter and spring preceding cool and warm events, which show approximately the same opposite dipoles centered on the ENP box as the SLPa. The ZA200 dipole creates positive (negative) wind speed anomalies that result in strong (weak) eastward flow when compared to normal (Peixoto and Oort 1992). These ZA200 dipoles are part of global atmospheric wave trains that constructively interfere over the ENP. Atmospheric wave trains for warm and cool events are approximately opposite. One wave train is an arcing wave train that appears to originate from near the central tropical Pacific (CTP), and the other is an approximately zonal wave train that spans the entire northern hemisphere and appears to originate from near the Maritime Continent region. We inferred the origins of these wave trains from the ZA200 patterns and from corresponding outgoing longwave radiation anomalies (not shown).

4. Discussion and conclusion

We have found that cool and warm events in the ENP during the summer are a result of complex interactions involving: 1) SST and atmospheric convective anomalies in the Maritime Continent and CTP regions; 2) winter and spring atmospheric wave trains that transport energy from the tropics to the ENP; and 3) anomalous dipoles in upper and lower tropospheric geopotential heights in the ENP that create anomalous ENP surface winds. The dynamical processes associated with cool and warm events are approximately opposite, illustrated by opposite

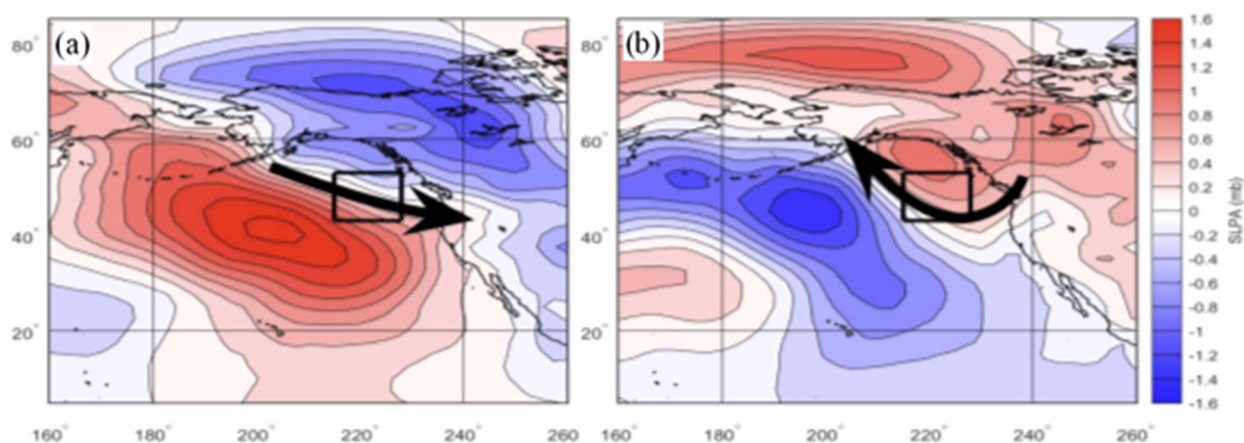


Fig. 3 Sea level pressure anomalies (SLPAs; mb) in December-May preceding (a) Cool events and (b) Warm events. Implied surface wind anomalies are shown schematically by the black arrows.

anomalies, spatial patterns, and temporal patterns in all variables discussed. Additionally, the dynamical processes that lead to these events begin at least six months prior, indicating that anomalous tropical and ENP conditions in winter and spring strongly determine summer SSTAs in the ENP. We also found that both cool and warm events have become more common and extreme, with frequent interannual variations between the two events. In on-going research, we are investigating: (1) methods for monitoring ENP warm and cool events; (2) the use of predictors in the tropical Indo-Pacific region to predict ENP SSTs at subseasonal to seasonal lead times; and (3) the atmospheric-oceanic dynamics in the ENP associated with the development of cool and warm events.

References

- Amaya, D. J., N. E. Bond, A. J. Miller, and M. J. Deflorio, 2016: The evolution and known atmospheric forcing mechanisms behind the 2013–2015 North Pacific warm anomalies. *US CLIVAR*, **14**, 1–6.
- , A. J. Miller, S. P. Xie, and Y. Kosaka, 2020: Physical drivers of the summer 2019 North Pacific marine heatwave. *Nat. Commun.*, **11**, 1903, <https://doi.org/10.1038/s41467-020-15820-w>
- Bond, N. A., M. F. Cronin, H. Freeland, and N. Mantua, 2015: Causes and impacts of the 2014 warm anomaly in the NE Pacific. *Geophys. Res. Lett.*, **42**, 3414–3420.
- Hayashida, H., R. J. Matear, P. G. Strutton, and X. Zhang, 2020: Insights into projected changes in marine heatwaves from a high-resolution ocean circulation model. *Nat. Commun.*, **11**, 4352, <https://doi.org/10.1038/s41467-020-18241-x>
- Holbrook, N. J., and co-authors, 2019: A global assessment of marine heatwaves and their drivers. *Nat. Commun.*, **10**, 2624, <https://doi.org/10.1038/s41467-019-10206-z>
- Kalnay, E., and co-authors, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**, 437–470.
- Kohlman, K., S. Madden, and T. Murphree, 2021: Marine heat waves in the Eastern North Pacific: Characteristics and causes. *Climate Prediction S&T Digest, 45th NOAA Climate Diagnostics and Prediction Workshop*, Virtual Online, DOC/NOAA, 113–119, doi: 10.25923/tpfe-4n87.
- Laufkötter, C., T. L. Frölicher, and J. Zscheischler, 2020: High-impact marine heatwaves attributable to human-induced global warming. *Science*, **369**, 1621–1625.
- Peixoto, J. P., and A. H. Oort, 1992: *Physics of Climate*, Springer, 520 pp.
- Rogers-Bennett, L., and C. A. Catton, 2019: Marine heat wave and multiple stressors tip bull kelp forest to sea urchin barrens. *Sci. Rep.*, **9**, 15050.
- Sen Gupta, A., and Coauthors, 2020: Drivers and impacts of the most extreme marine heatwaves events. *Sci. Rep.*, **10**, 19359.
- Smale, D. A., and Coauthors, 2019: Marine heatwaves threaten global biodiversity and the provision of ecosystem services. *Nat. Clim. Change*, **9**, 306–312.

Mechanism of the Centennial Subpolar North Atlantic Cooling Trend in the FGOALS-g2 Historical Simulation

Yifei Fan,^{1,2} Jianhua Lu,³ and Laifang Li^{1,4,5}

¹*Department of Meteorology and Atmospheric Science, the Pennsylvania State University, University Park, PA*

²*School of Atmospheric Sciences, Sun Yat-sen University, Zhuhai, China*

³*School of Atmospheric Sciences, Guangdong Province Key Laboratory for Climate Change and Natural Disaster Studies, Sun Yat-sen University, Guangzhou, China*

⁴*Institute of Computational and Data Sciences, the Pennsylvania State University, University Park, PA*

⁵*Environmental and Earth Science Institute, the Pennsylvania State University, University Park, PA*

ABSTRACT

A cold blob, manifested as a centennial cooling trend in sea surface temperature (SST), is observed in the mid-latitude North Atlantic. The presence of the cold blob is hypothesized as an evidence of a slowdown of Atlantic Meridional Overturning Circulation (AMOC), based on paleoclimate proxies and global climate models (GCMs). However, the performance of GCMs in simulating the cold blob remains unsatisfactory in terms of the SST cooling rate and the spatial extent. This study investigates the forcing mechanism of the cold blob using the Flexible Global Ocean-Atmosphere-Land System Model Grid-point version 2 (FGOALS-g2), a GCM that reasonably simulates the observed cooling trend in the subpolar North Atlantic and its spatial pattern. Surface heat budget analysis suggests that the cold blob is largely a result of the imbalance between changes in the heat storage and surface turbulent heat fluxes, exhibiting a cooling and warming effect, respectively. Investigation of ocean heat content indicates that heat advection into the cold blob region has decreased, mainly due to ocean circulation changes. However, in the FGOALS-g2, both the AMOC slowdown and the reduced

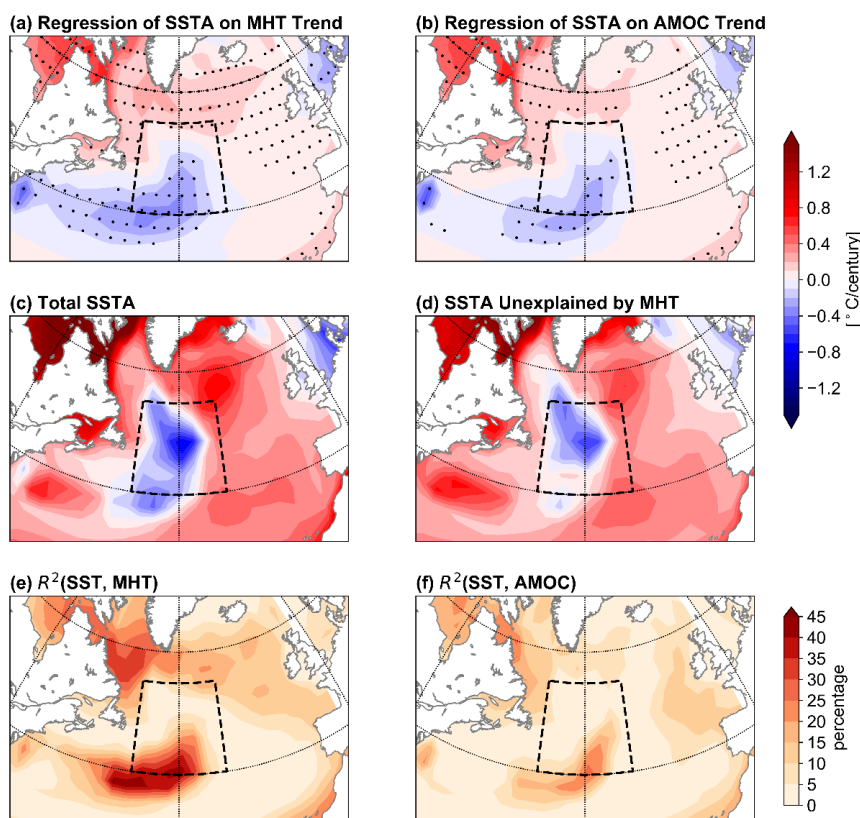


Fig. 1 Sea surface temperature anomaly (SSTA) regressed on the trends of (a) meridional heat transport (MHT) and (b) AMOC. Statistically significant ($p < 0.01$) regression coefficients are dotted. (c) Total SSTA trend simulated by FGOALS-g2. (d) The difference between (c) and (a). (e) and (f) Coefficient of determination between MHT and SST and between AMOC and SST, respectively.

meridional heat transport explain a limited portion (less than 50%) of the cold blob sea surface temperature anomaly trend (Fig. 1). Overall, complementing existing studies that attribute the cold blob to an AMOC slowdown, our results suggest that additional processes, including subpolar gyre circulation and a synergy between the atmosphere and the ocean, are at work in the formation of the cold blob.

This study was published in *the Journal of Geophysical Research: Oceans* in 2021.

References

Fan, Y., J. Lu, and L. Li, 2021: Mechanism of the centennial subpolar North Atlantic cooling trend in the FGOALS-g2 historical simulation. *J. Geophys. Res.: Oceans*, **126**, doi:10.1029/2021JC017511

Initialized and Uninitialized ENSO Predictability in Year 2+

Nathan Lenssen,¹ Lisa Goddard,^{1†} Simon Mason,¹ and Yochanan Kushnir²

¹*International Research Institute for Climate and Society (IRI), Columbia University, Palisades, NY*

²*Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY*

1. Introduction

Extending global climate prediction past the currently operational 9-month outlooks is of significant scientific and societal importance. Currently, the most promising source of multi-year predictive skill comes from the El Niño-Southern Oscillation (ENSO), the dominant source of seasonal and interannual climate variability. The 2021 food crisis in the Horn of Africa highlights the need for, and potential of, multi-year ENSO predictions. Consecutive La Niña events often cause drought in the Horn of Africa (Hoell and Funk 2014, Lenssen *et al.* 2020) leading to reduced crop yields (Iizumi *et al.* 2014), which contributes to food crises (FEWSNet 2021). In this example, translation of the known predictability of consecutive La Niña events (DiNezio *et al.* 2017) into a skillful operational forecast would have provided critical early information to humanitarian organizations responding to the food crisis.

Key features of ENSO, particularly its duration, are theoretically predictable several years in advance (Gonzalez and Goddard 2016, DiNezio *et al.* 2017, Ham *et al.* 2019, Dunstone *et al.* 2020, Wu *et al.* 2021). Potential physical mechanisms leading to multi-year ENSO predictability include the subsurface heat content in the tropical Pacific Ocean (McPhaden 2003, Zhao *et al.* 2021), basin-scale Pacific Ocean dynamics (Vimont *et al.* 2003; Joh and DiLorenzo 2019), and cross-basin interactions with the Indian (Mayer and Balmaseda 2021) and Atlantic Oceans (Ham *et al.* 2013). Despite our understanding of the dynamical processes leading to extended ENSO predictability, multi-year skill in traditionally initialized dynamical forecast systems has remained elusive.

Possible reasons why multi-year predictability does not always translate into predictive skill in initialized dynamical forecasts include model bias and errors due to initialization. Nearly all coupled general circulation models (CGCMs) exhibit bias in the tropical Pacific mean state and variability, leading to unrealistic ENSO behavior (Li and Xie, 2014, Planton *et al.* 2021). The link between ENSO bias and predictive skill is still poorly understood with some studies finding no relationship (Scaife *et al.* 2019), while others find a conclusive relationship (Ding *et al.* 2020). In addition, initialization shock, or rapidly increasing forecast error due to the initial observed state being incompatible with the CGCM's dynamics, results in drifts in the mean state and variability of ENSO (Mullholland *et al.* 2015), including a westward shift of the predicted ENSO anomaly that results in poor forecast skill in the western tropical Pacific (Newman and Sardeshmukh 2017). This loss of skill due to initialization shock can be seen through the superior long-lead skill of empirical-dynamical forecast methods that take advantage of CGCM simulations, but do not use traditional initialization techniques, such as model-analogues (Ding *et al.* 2018), neural networks (Ham *et al.* 2019), and linear inverse models (Penland and Sardeshmukh 1995). Understanding, reducing, and correcting initialization shock is necessary to improve our climate forecast systems. However, there has not been a comprehensive study on the effect of initialization on ENSO predictability in initialized dynamical prediction systems.

Here, the potential multi-year ENSO skill currently not captured in initialized prediction systems due to CGCM bias and initialization shock is investigated. This work identifies systemic biases in CGCM ENSO dynamics that must be reduced to improve long-lead prediction by comparing metrics of CGCM bias with predictive skill across models. The potential skill to be gained in initialized forecast systems will be assessed

[†] Deceased

through a comparison of traditionally initialized forecasts with model-analogue forecasts, an empirical-dynamical model based on CGCMs output.

2. Data and methodology/experimental design

Observed monthly sea-surface temperature (SST) is taken from HadISST1.1 (Rayner *et al.* 2003). Observed sea-surface height (SSH) is from the ECMWF Ocean Reanalysis System 4 (ORAS4) (Balmaseda *et al.* 2013). All model and observational data is regridded to a 2×2 degree grid prior to any analyses. ENSO events are defined as seasonal Niño3.4 anomalies that exceed the upper or lower quartile of the seasonal Niño3.4 index following (Gonzalez and Goddard 2016).

The initialized predictions come from the CMIP6 Decadal Climate Prediction Project (DCPP) Component A hindcasts with annual initializations from 1960 - 2016 (Boer *et al.* 2016). This study was limited to the three models with complete data and corresponding control runs of sufficient length on the Google Cloud CMIP6 archive: CanESM5, CESM1-1-CAM5-CMIP5 (CESM1.1), MIROC6. CESM1.1 and MIROC6 are both initialized in November and CanESM5 is initialized in January. The Niño3.4 indices for the initialized hindcasts are bias corrected to account for both lead-dependent mean-state bias as well as trend bias following Kharin *et al.* (2012). In addition, ENSO event thresholds in the initialized models are defined as the upper and lower quartile for each lead following Gonzalez and Goddard (2016). After bias correcting, probabilistic forecasts of ENSO state are made for JFM for years 0-5 for each model and lead time.

Model-analogue hindcasts are made over the same 1960-2016 period using control runs from the same CGCM configuration as the traditionally initialized forecasts. A model-analogue forecast is made by determining which state in a library of model output best matches the observed climate state. The forecast is the evolution of these closest matching library states, following the assumption that a pair of states that is initially similar will evolve along similar trajectories (Lorenz 1969). Here, ENSO forecasts are made by matching observed Indo-Pacific SST and SSH to states in a library of CGCM control run output following Ding *et al.* (2018). Though the model-analogue forecasts are able to issue forecasts with CGCM output while avoiding initialization shock, they have the disadvantage of larger error than traditionally initialized forecasts in the first month due to the analogues not perfectly matching the observed state.

Hindcasts are verified against the observed evolution of ENSO as calculated through the Niño3.4 Index from HadISST1.1. The forecasts of interest are probabilistic forecasts of ENSO events for which forecast skill is assessed by the area under the receiver operating characteristic (ROC) curve (Mason 1982; Hogan and Mason 2012). The ROC area is a measure of forecast discrimination, a skill measure relative to a climatological forecast, and can be generalized as a *U*-Statistic providing useful statistical properties (Mason and Graham 2002). In particular, statistical significance of skillful ROC scores is assessed using the assumption that the *U*-statistic is Gaussian, which holds for large sample sizes (Mason and Graham 2002).

3. Results and discussion

The two major goals of this preliminary work are to quantify the GCM ENSO climatologies and assess the ENSO prediction skill at leads of multiple years. First, the ENSO evolution in mean state and variability is assessed in the uninitialized pre-industrial control experiment (piControl) runs of the three CGCMs (Fig. 1). The seasonal cycle of each of the models roughly matched that of observations with temperature peaking in boreal summer and reaching minimum during boreal winter. However, the timing of the winter minimum is early in all three models suggesting a mismatch in ENSO timing when compared to observations. Here, each of the three models shows an ENSO climatological amplitude that is larger than observations, in agreement with other studies of CMIP5 and CMIP6 ENSO amplitude (Bellenger *et al.* 2014, Planton *et al.* 2021). The evolution of the seasonal cycle of the initialized models shows the signature of errors arising from initialization (Fig. 1). In particular, the initialized CanESM5 model shows a large and growing departure from observations in the critical winter months.

The variability of Niño3.4 temperatures is assessed to determine if the ENSO variability is properly represented in the models. CESM1.1 and MIROC6 both show quite realistic variability in the uninitialized piControl runs, but CanESM5 shows dramatically reduced ENSO variability in the key boreal winter months

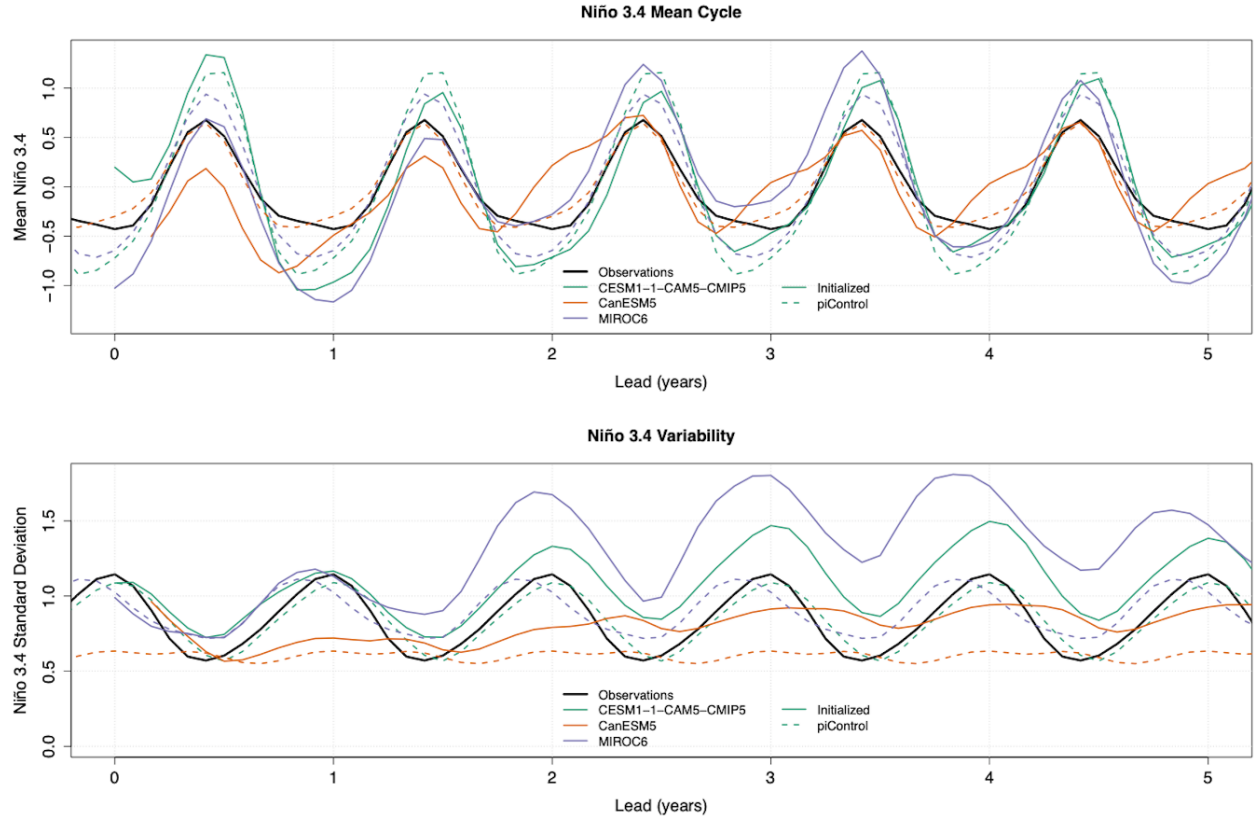


Fig. 1 The climatology of the initialized and uninitialized ENSO simulations as compared with observations with the lead time zeroed at January of the first year. Shown is (top) the mean cycle of monthly Niño3.4 mean absolute temperature with the 12-month running mean removed and (bottom) variability of monthly Niño 3.4 absolute temperature. The observations (solid black line) are calculated over 1960-2016 and do not depend on lead time. The piControl climatologies (colored dashed lines) are calculated over the entire length of the piControl and also do not depend on lead time. The initialized model climatologies (colored solid lines) vary with lead time, reflecting the lead-dependent biases in mean and variability.

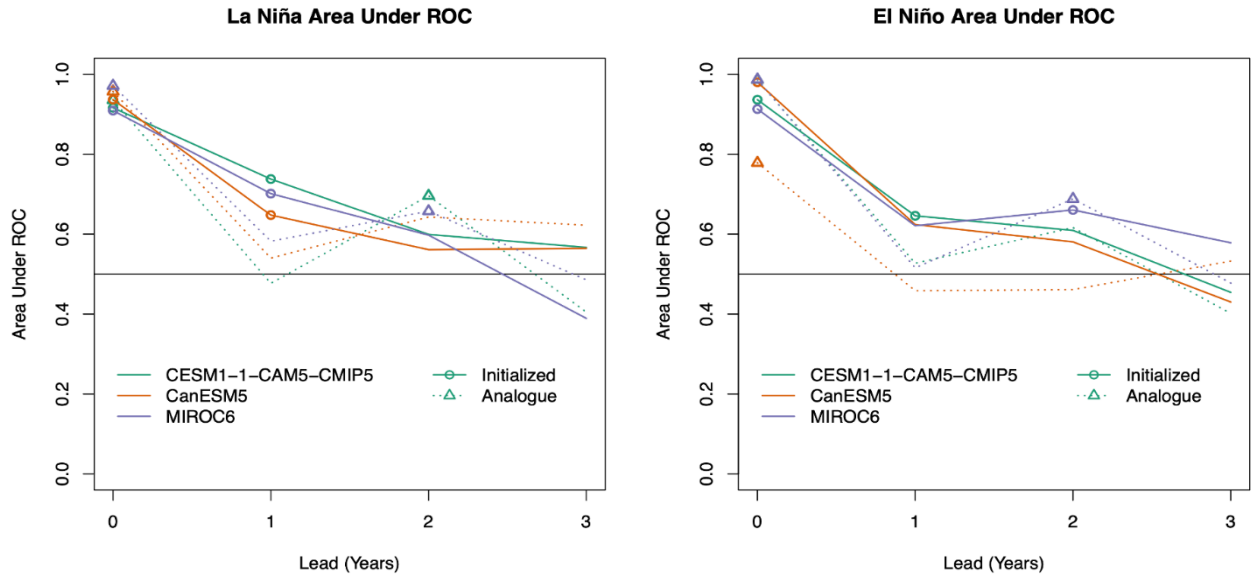


Fig. 2 The ROC skill for ENSO event detection at 0-3 year leads. Climatological skill is 0.5 and marked by the solid black line. Statistically significant positive skill is marked with a circle for initialized forecasts or triangle for model-analogue forecasts.

suggesting that CanESM5 struggles to simulate a diverse range of ENSO amplitudes in agreement with Planton *et al.* (2021). When the models are initialized, each of their ENSO variabilities grow as a function of lead time while preserving a similar seasonal cycle to that of the uninitialized versions. This behavior is potentially a signature of initialization shock-caused dynamical corrections in the ENSO system and is to be investigated further.

The initialized models all show significant skill in predicting La Niña events at 1-year leads, extending the findings of DiNezio *et al.* (2017) to multiple CGCMs (Fig. 2). However, the uninitialized model analogues do not show this same skill at 1 year. Curiously, these model analogue forecasts are more skillful at 2 years than at 1 year, a finding which warrants further investigation. There is less skill for El Niño events at 1 year, with only the initialized CESM1.1 showing significant discrimination when compared to climatology (Fig. 2). Again, the increase in skill from lead 1 to lead 2 is observed in many of the initialized and model-analogue predictions, though the individual increases are not statistically significant. Notably, both initialized and model-analogue MIROC6 show significant skill in lead-2 El Niño despite not showing skill in lead 1. No link between the CGCM ENSO climatology and predictive skill was found, but such investigations will be revisited as more models are added into the analysis

4. Conclusions and future work

This work is a promising initial investigation of multi-year prediction of ENSO events using CMIP6-class initialized models as well as model-analogue predictions using existing control run simulations. Here, evidence for the effect of initialization on CGCM ENSO dynamics is presented. However, the hypothesis that these growing errors due to initialization shock would make model-analogue predictions superior at longer leads was not shown. While the initialized forecasts show more discrimination in forecasting ENSO events at most leads, there is promise in the significant two year lead predictive skill shown by model-analogue forecasts for both La Niña and El Niño events.

Looking forward, this work will continue by expanding the number of models analyzed up to the 11 CGCMs that plan to submit runs to the Decadal Climate Prediction Project (DCPP), allowing for a more rigorous investigation of the link between model ENSO dynamics and predictive skill. In addition, the model-analogue methods will be compared with the predictions from linear inverse models (LIMs), which have comparable skill to initialized predictions (Newman and Sardeshmukh 2017). Finally, an investigation of successful long-lead forecasts will be conducted to determine if there are certain characteristics of ENSO events and/or CGCM systems that lend themselves to greater predictive skill.

References

- Balmaseda, M.A., K. Mogensen, and A. T. Weaver, 2013: Evaluation of the ECMWF ocean reanalysis system ORAS4. *Q. J. R. Meteorol. Soc.*, **139**, 1132–1161.
- Bellenger, H., E. Guilyardi, J. Leloup and Coauthors, 2014: ENSO representation in climate models: from CMIP3 to CMIP5. *Clim. Dyn.*, **42**, 1999–2018.
- Boer, G. J., D. M. Smith, C. Cassou, and Coauthors, 2016: The Decadal Climate Prediction Project (DCPP) contribution to CMIP6. *Geosci. Model Dev.*, **9**, 3751–3777
- DiNezio, P. N., C. Deser, Y. Okumura, and Coauthors, 2017: Predictability of 2-year La Niña events in a coupled general circulation model. *Clim. Dyn.*, **49**, 4237–4261.
- Ding, H., M. Newman, M. A. Alexander, and A. T. Wittenberg, 2018: Skillful climate forecasts of the tropical Indo-Pacific Ocean using model-analogs, *J. Climate*, **31**, 5437–5459.
- , —, —, and —, 2020: Relating CMIP5 model biases to seasonal forecast skill in the tropical Pacific. *GRL*, **47**(5).
- Dunstone, N., D. Smith, S. Yeager, and Coauthors, 2020: Skillful interannual climate prediction from two large initialised model ensembles, *Environ. Res. Lett.*, **15**, 094083.
- FEWS NET, 2021: Worsening drought threatens Horn of Africa as conflict-driven emergency persists in northern Ethiopia. <https://fewsn.net/east-africa/alert/october-27-2021>
- Gonzalez, P. L., and L. Goddard, 2016: Long-lead ENSO predictability from CMIP5 decadal hindcasts. *Clim. Dyn.*, **46**, 3127–3147.

- Ham, Y.-G., J.-H. Kim, and J.-J. Luo, 2019: Deep learning for multi-year ENSO forecasts. *Nature*, **573**, 568–572.
- , J.-S. Kug, and J.-Y. Park, 2013: Two distinct roles of Atlantic SSTs in ENSO variability: North Tropical Atlantic SST and Atlantic Niño. *Geophys. Res. Lett.*, **40**, 4012–4017.
- Hoell, A., and C. Funk, 2014: Indo-Pacific sea surface temperature influences on failed consecutive rainy seasons over eastern Africa. *Clim. Dyn.*, **43**, 1645–1660.
- Hogan, R. J., and I. B. Mason, 2012: Deterministic forecasts of binary events. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, I. T. Jolliffe and D. B. Stephenson, Eds., Wiley, 31–59.
- Iizumi, T., J. J. Luo, A. J. Challinor, and Coauthors, 2014: Impacts of El Niño Southern Oscillation on the global yields of major crops. *Nat. Commun.*, **5**, 3712, <https://doi.org/10.1038/ncomms4712>
- Joh, Y., and E. Di Lorenzo, 2019: Interactions between Kuroshio Extension and Central Tropical Pacific lead to preferred decadal-timescale oscillations in Pacific climate. *Sci. Rep.*, **9**, 13558, doi:10.1038/s41598-019-49927-y.
- Kharin, V. V., G. J. Boer, W. J. Merryfield, and Coauthors, 2012: Statistical adjustment of decadal predictions in a changing climate. *Geophys. Res. Lett.*, **39**, L19705, doi:10.1029/2012GL052647.
- Lenssen, N. J., L. Goddard, and S. Mason, 2020: Seasonal forecast skill of ENSO teleconnection maps. *WAF*, **35**, 2387–2406.
- Li, G., and S.-P. Xie, 2014: Tropical biases in CMIP5 multimodel ensemble: The excessive equatorial Pacific cold tongue and double ITCZ problems. *J. Climate*, **27**, 1765–1780.
- Lorenz, E. N., 1969: Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.*, **26**, 636–646.
- Mason, I., 1982: A model for assessment of weather forecasts. *Aust. Meteorol. Mag.*, **30**, 291–303.
- Mason, S. J., and N. E. Graham, 2002: Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. *Q. J. R. Meteorol. Soc.*, **128**, 2145–2166.
- Mayer, M., and M. A. Balmaseda, 2021: Indian Ocean impact on ENSO evolution 2014–2016 in a set of seasonal forecasting experiments. *Clim. Dyn.*, **56**, 2631–2649.
- McPhaden, M. J., 2003: Tropical Pacific Ocean heat content variations and ENSO persistence barriers. *Geophys. Res. Lett.*, **30**, 1480, doi:10.1029/2003GL016872.
- Mulholland, D. P., P. Laloyaux, K. Haines, and Coauthors, 2015: Origin and impact of initialization shocks in coupled atmosphere–ocean forecasts. *Mon. Wea. Rev.*, **143**, 4631–4644.
- Newman, M., and P. D. Sardeshmukh, 2017: Are we near the predictability limit of tropical Indo-Pacific sea surface temperatures?, *Geophys. Res. Lett.*, **44**, 8520–8529.
- Planton, Y. Y., E. Guilyardi, A. T. Wittenberg, and Coauthors, 2021: Evaluating climate models with the CLIVAR 2020 ENSO metrics package. *Bull. Amer. Meteor. Soc.*, **102**, E193–E217.
- Rayner, N. A., D. E. Parker, E. B. Horton, and Coauthors, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108**, No. D14, 4407, doi:10.1029/2002JD002670.
- Scaife, A. A., L. Ferranti, O. Alves, and Coauthors, 2019: Tropical rainfall predictions from multiple seasonal forecast systems. *Int. J. Climatol.*, **39**, 974–988.
- Vimont, D. J., J. M. Wallace, and D. S. Battisti, 2003: The seasonal footprinting mechanism in the Pacific: Implications for ENSO, *J. Climate*, **16**, 2668–2675.
- Wu, X., Y. M. Okumura, C. Deser, and P. N. DiNezio, 2021: Two-year dynamical predictions of ENSO event duration during 1954–2015, *J. Climate*, **34**, 4069–4087.
- Zhao, S., F. F. Jin, and M. F. Stuecker, 2021: Understanding lead times of warm water volumes to ENSO sea surface temperature anomalies. *Geophys. Res. Lett.*, **48**, e2021GL094366, doi:10.1029/2021GL094366.

Seasonal Tropical-Extratropical Teleconnections Originating from Tropical Rainfall Modes Beyond Canonical ENSO for Northern Winters

Peitao Peng, Wanqiu Wang and Arun Kumar

Climate Prediction Center, NOAA/NWS/NCEP, College Park, MD

1. Introduction

Since the pioneering work by Horel and Wallace (1981), the tropical-extratropical atmospheric teleconnection excited by El Niño-Southern Oscillation (ENSO) has been recognized as a predominant source of short-term climate predictability (Anderson *et al.* 1998; Shukla *et al.* 2000; Hoerling and Kumar 2002).

In the past two decades, efforts extended to finding possible tropical-extratropical teleconnections originating from other sectors of tropical oceans, such as the impact of tropical Indian Ocean warming onto the positive trend in the North Atlantic Oscillation (NAO) (Bader and Latif 2003; Hurrell *et al.* 2004), and rainfall anomalies over the western and central Indian Ocean onto height anomalies over Alaska and North Atlantic regions (Moltoni *et al.* 2015).

These studies have advanced our knowledge on the relationships between extratropical atmospheric anomalies and tropical SST and rainfall variability. Despite that, there are still issues needing to be resolved. For example, what teleconnection pattern originates from Indian Ocean and Maritime Continent if ENSO is not involved? More generally, are there any teleconnections from the tropics that were not identified in previous studies?

This study aims at addressing these questions. The analysis starts with mode decomposition for the tropical rainfall and then calculates the atmospheric patterns associated with the rainfall modes. The reason for choosing rainfall rather than SST in the analysis is that the rainfall represents vertically integrated latent heating, which is the direct forcing of the tropical atmosphere. In addition, there seems no definite relationship between rainfall and SST in the west Pacific and Indian Ocean (Chen *et al.* 2012; Kumar *et al.* 2013; Moltoni *et al.* 2015).

A straightforward method of mode decomposition is the empirical orthogonal function (EOF) analysis (Wallace and Gutzler 1981; Barnstone and Livezey 1987). However, when it is applied to tropical rainfall, the obtained modes are found not quite clean, that is, the ENSO signal tends to be mixed with other modes. The improvement is limited even when using the rotated EOF (REOF). Following the study of Peng *et al.* (2014), in which the Pacific and North America (PNA) pattern was successfully separated from the ENSO teleconnection pattern by applying REOF analysis to the data with ENSO signal removed, here we apply the same technique to tropical rainfall data. For the sake of simplicity, we focus on the seasonal mean of northern winters when teleconnection patterns have the strongest amplitude in the Northern Hemisphere (NH) (Barnston and Livezey 1987).

2. Data and analysis procedures

2.1 Data

The data used in this study are northern winter [December-January-February (DJF)] mean precipitation, 200hPa geopotential height (Z200) and sea surface temperatures (SSTs). The precipitation data are from the CPC merged analysis of precipitation (CMAP, Xie and Akin 1997), Z200 data from NCEP/NCAR reanalysis (Kalnay *et al.* 1996), and SSTs are based on Hurrell *et al.* (2008). Because of the availability of rainfall data over the oceans, the analysis period is from 1979/80 to 2020/21, thus, there are 42 DJF seasons in total. Seasonal mean anomalies are computed with respect to the climate mean over the whole data period. The Nino 3.4 SST index (Nino 3.4 index hereafter), widely used as an ENSO proxy, is calculated as the averaged SST anomalies over the Nino 3.4 region (120°W-170°W, 5°S-5°N).

2.2 Analysis procedures

a) Mode decomposition for tropical rainfall

To avoid the mixing of ENSO signal with other variability, the mode decomposition is done in two steps. The first step removes the ENSO signal from the rainfall data. This is done by regressing the Niño 3.4 index to the rainfall anomalies, and then removing the product of the obtained regression pattern and the Niño 3.4 index from the rainfall anomalies for each DJF season. The second step applies the EOF analysis to the residual data in the deep tropics (10°S - 10°N). Confining the EOF analysis to the deep tropics is to avoid mixing middle latitude variability. The EOF analysis is based on the covariance matrix to have more variance explained with fewer leading modes. REOFs are calculated with seven EOF modes. The cutoff number is chosen for an optimal use of REOFs (O'Lenic and Livezey 1988). The time series associated with the REOF patterns are referred to as the rotated principal components (RPCs).

b) Teleconnection patterns associated with the tropical rainfall modes

The global teleconnection patterns associated with these tropical rainfall modes are obtained by regressing their time series to the global fields of Z200 and other variables.

3. Results

3.1 Dominant modes of tropical rainfall residual and associated SST patterns

Figure 1 shows the four leading non-ENSO modes of the DJF mean rainfall in the deep tropics. Starting from the top, these modes are ranked with their explained percentages of the total variance in the deep tropics.

The first mode, explaining 13.0% of the total variance, has a dipole pattern near the dateline. The variation in the time series (RPC1) is basically interannual. Interestingly, the two largest negative values in the time series are in 1997/98 and 1982/83 winters, corresponding to two of the three strongest El Niño events in the data period. As will be seen later (Fig. 2), this rainfall mode is associated with a tropical SST mode called El Niño Modoki (Ashok *et al.* 2007).

The second mode, explaining 6.2% of the total variance, has its spatial loading mainly in the Indo-Pacific warm pool region, but also with some features of opposite sign in the western and central tropical Pacific and the western tropical Indian Ocean. Its time series (RPC2) shows an upward trend in addition to the interannual variability.

The third mode, explaining 5.3 % of the total variance, has its spatial loading mainly in the southern tropical Indian Ocean. Its time series (RPC3) contains interannual variability and a downward trend.

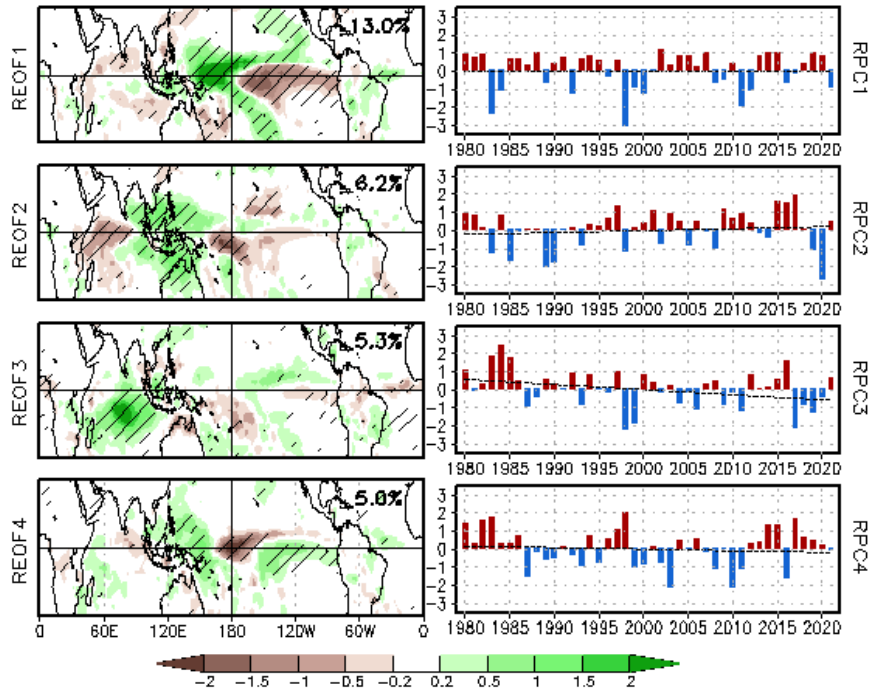


Fig. 1 Dominant non-ENSO modes of DJF mean rainfall in the deep tropics (15°S - 15°N). The spatial patterns (left column) are the regressions of the rainfall data (30°S - 30°N) to the time series of the modes (right column). Percentage numbers are the fractions of total variance explained by the corresponding modes. Hatches indicate the significant level exceeding 95% in the t-test. Dotted lines show the linear trend of the time series.

The fourth mode, explaining 5.0% of the total variance, has a tripole structure in the tropical Pacific, with the central pole on the dateline. Its time series (RPC4) has a small downward trend and some inter-decadal variability, with positive values dominating the earlier and recent periods and more negative values in between.

To examine the SST anomalies associated with the rainfall patterns, Fig. 2 presents the regressions of DJF SSTs for the four rainfall modes. The pattern in the first panel looks the same as the El Niño Modoki, that is, the second EOF mode of the tropical SSTs, reflecting differences between the eastern Pacific (EP) ENSO and the central Pacific (CP) ENSO (Ashok *et al.* 2007; Kao and Yu 2009). Its corresponding rainfall pattern (Fig.1) is westward shifted with respect to SST, implying that SST anomalies mostly influence rainfall over climatologically high SST regions (He *et al.* 2018).

The SST anomalies corresponding to the second rainfall modes are in the Pacific warm pool region, while the rainfall anomalies extend to the entire Maritime Continent, this may be due to the thermal and topographical effect of the Maritime Continent (Yang *et al.* 2019).

For the third mode, the SST and rainfall patterns in the Indian Ocean are only partially matched, suggesting the SST pattern may not be an only forcing of the rainfall pattern.

For the fourth mode, the SST pattern is consistent with its corresponding rainfall pattern in a tripole structure.

Overall, the relationship between the non-ENSO seasonal SST and rainfall patterns is not quite definite. The reason is that SST variation is not the only factor influencing the atmospheric stability in the tropics, upper tropospheric temperature anomalies also have influence on gross moist stability (Neelin and Held 1987; Izumo *et al.* 2019). In addition, a seasonal mean rainfall anomaly can be a residual of intraseasonal variability (Moltoni *et al.* 2015), which may originate either from the tropics or from the extratropics (Stan *et al.* 2017).

3.2 Teleconnection patterns

To give a conventional view of the teleconnection patterns in middle and high latitudes, Fig. 3 displays regression patterns of DJF Z200 for the four tropical rainfall modes on the north-polar projection maps. A positive anomaly over the polar region and Alaska and a tripole structure over North America and North Atlantic correspond to the ENSO Modoki (Fig. 3a); a positive anomaly over the polar region, a PNA-like pattern and a

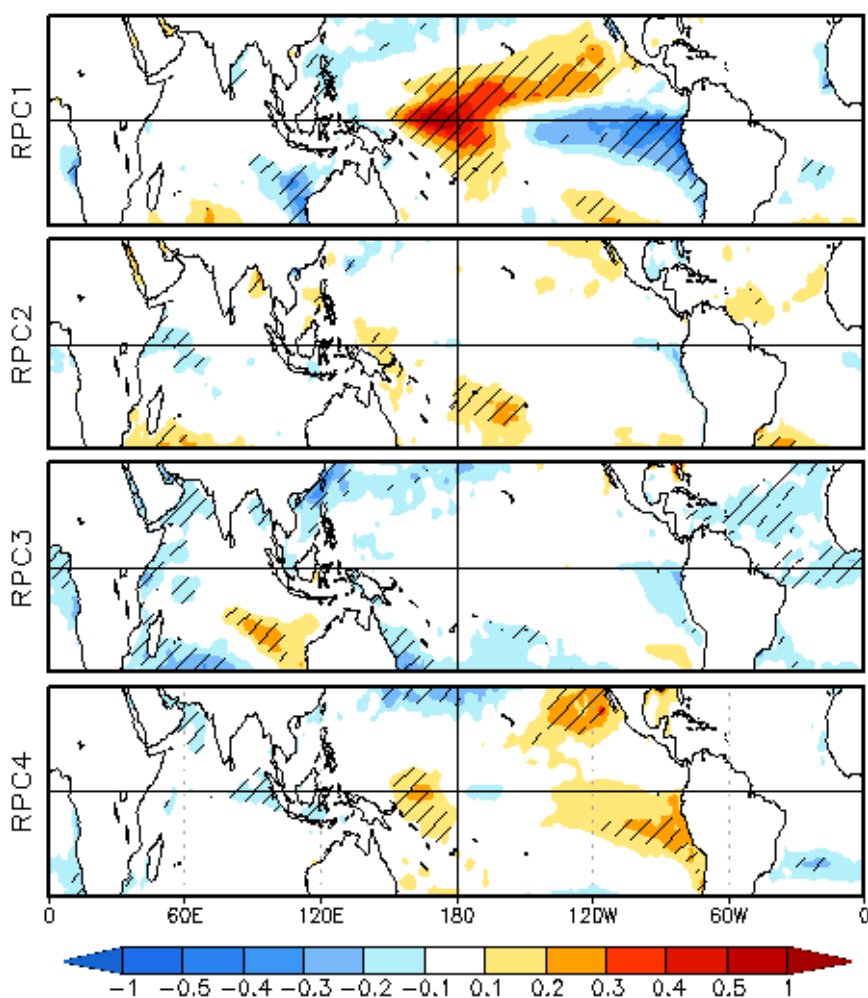


Fig. 2 Regression of tropical (30°S-30°N) SST (K) onto the time series of the rainfall modes shown in Fig. 1. Hatches indicate the significant level exceeding 95% in the t-test.

trough over the Europe are associated with the Indo-Pacific warm pool rainfall pattern (Fig. 3b); a ridge over the west coast of North America and a positive NAO are related to the rainfall pattern in the tropical Indian Ocean south of the equator (Fig. 3c); a WPO pattern and a similar tripole structure over North Atlantic are related to the tripole rainfall pattern in Pacific (Fig. 3d).

4. Summary

This study clarifies seasonal tropical-extratropical atmospheric teleconnection patterns beyond ENSO using tropical rainfall modes for northern winters. The mode decomposition is done by first removing the ENSO signal from tropical rainfall data and then applying the REOF analysis to the residual variability. This procedure effectively separates ENSO signal from other variability. The four dominant teleconnection patterns obtained by regressing global atmospheric fields against the time series of the rainfall modes include those related to El Niño Modoki and other rainfall patterns in various sectors of tropical oceans, respectively. The examination of the relationship between the rainfall and SST patterns helps understanding the sources.

The advantage of using rainfall rather than SST in the mode analysis is that rainfall represents vertically integrated latent heating, which is the direct forcing of the tropical atmosphere, while SST may have no definite relationship with rainfall in the western Pacific and Indian Ocean (Chen *et al.* 2012; Kumar *et al.* 2013; Molteni *et al.* 2015). Therefore, tropical SST modes beyond ENSO and ENSO Modoki may not be important for representing tropical forcing.

Most teleconnection patterns exhibited in Figs. 3 consist of some familiar patterns (*i.e.* PNA, NAO and WPO) found in previous analyses of mid-latitude height field (Wallace and Gutzler 1981; Barnston and Livezey 1987) and known as originating from atmospheric internal dynamics (Straus and Shukla 2002; Linkin and Nigam 2008). Their occurrence in the tropical-extratropical teleconnection patterns implies these intrinsic modes of the mid-latitude atmosphere can be tropically triggered as well.

The results of this study are applicable to the analyses of climate prediction and attribution. Specifically, one can decompose the forecasted or observed tropical rainfall anomalies into the independent modes, obtain the corresponding teleconnection patterns, and identify the modes contributing significantly to the target climate anomalies through a reconstruction procedure (Peng *et al.* 2018). In addition, climate models can be evaluated by comparing their tropical rainfall modes and the associated teleconnection patterns with that from observational data.

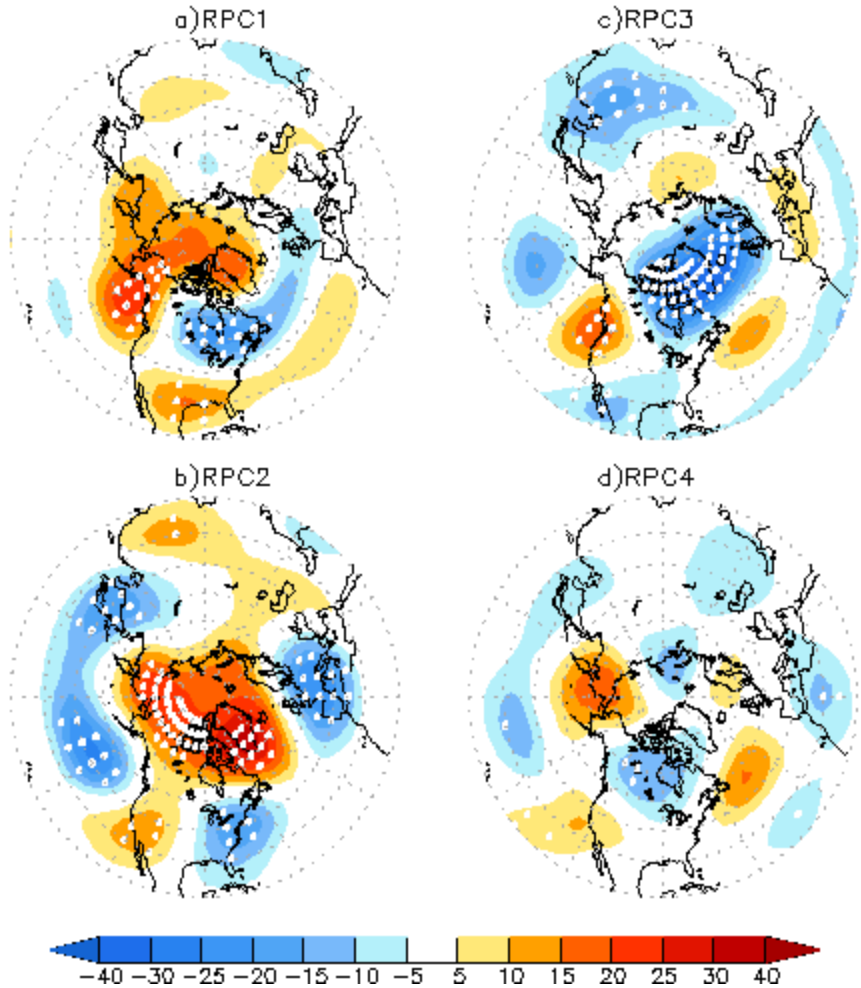


Fig. 3 Regression of 200hPa geopotential height (m) onto the time series of the rainfall modes shown in Fig. 1. White dots indicate the significant level exceeding 95% in the t-test.

A more practical question is that if these teleconnections can help improve seasonal predictions for the extratropics. To answer this, the predictability of these tropical rainfall modes needs to be investigated.

References

- Anderson, D. L. T., E. S. Sarachik, P. J. Webster, and L. M. Rothstein, 1998: Preface [to special section on The TOGA Decade: Reviewing the Progress of El Niño Research and Prediction]. *J. Geophys. Res.*, **103**(C7), 14167-14510, doi:10.1029/98JC00112.
- Ashok, K., S. K. Behera, S. A. Rao, H. Weng, and T. Yamagata, 2007: El Niño Modoki and its possible teleconnection. *J. Geophys. Res.*, **112**, C11007.
- Bader, J., M. Latif, 2003: The impact of decadal-scale Indian Ocean sea surface temperature anomalies on Sahelian rainfall and the North Atlantic Oscillation. *Geophys. Res. Lett.*, **30**, doi:10.1029/2003gl018426.
- Barnston, A. G., and R. E. Livezey, 1987: Classification, seasonality, and persistence of low-frequency atmospheric circulation patterns. *Mon. Wea. Rev.*, **115**, 1083-1126.
- Chen, M., W. Wang, A. Kumar, H. Wang, and B. Jha, 2012: Ocean surface impacts on the seasonal-mean precipitation over the tropical Indian Ocean, *J. Climate*, **25**, 3566-3582.
- He, J., N. C. Johnson, G. A. Vecchi, B. Kirtman, A. T. Wittenberg, and S. Sturm, 2018: Precipitation sensitivity to local variations in tropical sea surface temperature. *J. Climate*, **31**, 9225-9238, doi:10.1175/JCLI-D-18-0262.1
- Hoerling, M. P., and A. Kumar, 2002: Atmospheric response patterns associated with tropical forcing. *J. Climate*, **15**, 2184-2203.
- Horel, J. D., and J. M. Wallace, 1981: Planetary-scale atmospheric phenomena associated with the Southern Oscillation. *Mon. Wea. Rev.*, **109**, 813-829.
- Hurrell, J. W., J. Hack, D. Shea, J. Caron, and J. Rosinski, 2008: A new sea surface temperature and sea ice boundary dataset for the community atmosphere model. *J. Climate*, **21**, 5145-5153.
- , M. P. Hoerling, A. S. Phillips, and T. Xu, 2004: Twentieth Century North Atlantic climate change. Part I: Assessing determinism, *Clim. Dyn.*, **23**, 371-389, doi: 10.1007/s00382-004-0432-y.
- Izumo, T., J. Vialard, M. Lengaigne, and I. Suresh, 2019: Relevance of relative sea surface temperature for tropical rainfall interannual variability. *Geophys. Res. Lett.*, doi:10.1029/2019gl086182.
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437-472, doi:10.1175/1520-0477(1996)077<0437:TNYP>2.0.CO;2.
- Kao, H., and J. Yu, 2009: Contrasting Eastern-Pacific and Central-Pacific types of ENSO. *J. Climate*, **22**, 615-632, doi:10.1175/2008JCLI2309.1.
- Kumar, A., M. Chen, and W. Wang, 2013: Understanding prediction skill of seasonal mean precipitation over the tropics, *J. Climate*, **26**, 5674-5681.
- Linkin, M. E., S. Nigam, 2008: The North Pacific Oscillation and western North Pacific teleconnection pattern: Mature-phase structure and winter impacts. *J. Climate*, **21**, 1979-1997, doi:10.1175/2007JCLI2048.1.
- Molteni, F., T. N. Stockdale, and F. Vitart, 2015: Understanding and modelling extra-tropical teleconnections with the Indo-Pacific region during the northern winter. *Clim. Dyn.*, **45**, 3119-3140. doi:10.1007/s00382-015-2528-y.
- Neelin, J. D., and I. M. Held, 1987: Modeling tropical convergence based on the moist static energy budget. *Mon. Wea. Rev.*, **115**, 3-12, doi:10.1175/1520-0493(1987)115,0003:MTCBOT.2.0.CO;2,1.
- O'Lenic, E., and R. E. Livezey, 1988: Practical considerations in the use of rotated principal components analysis (RPCA) in diagnostic studies of upper-air height fields. *Mon. Wea. Rev.*, **116**, 1682-1689.
- Peng, P., A. Kumar, and B. Jha, 2014: Climate mean, variability and dominant patterns of the Northern Hemisphere wintertime mean atmospheric circulation in the NCEP CFSv2. *Clim. Dyn.*, **42**, 2783-2799.
- , —, and Z.-Z. Hu, 2018: What drove the Pacific and North America climate anomalies in winter 2014/15? *Clim. Dyn.*, **51**, 2667-2679, doi:10.1007/s00382-017-4035-9.

-
- Shukla, J., and Coauthors, 2000: Dynamical seasonal prediction. *Bull. Amer. Meteor. Soc.*, **81**, 2593-2606, doi:10.1175/1520-0477(2000)081<2593:DSP>2.3.CO;2.
- Stan, C., D. M. Straus, J. S. Frederiksen, H. Lin, E. D. Maloney, and C. Schumacher, 2017: Review of tropical-extratropical teleconnections on intraseasonal time scales. *Rev. Geophys.*, **55**, 902-937, doi:10.1002/2016RG000538.
- Straus, D. M., and J. Shukla, 2002: Does ENSO force the PNA? *J. Climate*, **15**, 2340-2358, doi:10.1175/1520-0442(2002)015<2340:DEFTP>2.0.CO;2
- Wallace, J. M., and D. S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Wea. Rev.*, **109**, 784-812.
- Xie, P., and P. A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.*, **78**, 2539-2558.
- Yang, S., T. T. Zhang, Z. N. Li, and Coauthors, 2019: Climate variability over the Maritime Continent and its role in global climate variation: A review. *J. Meteor. Res.*, **33**, 993-1015, doi:10.1007/s13351-019-9025-x.



2. Extremes and Extreme Events



Global Seasonal Forecasts of Marine Heatwaves

Michael G. Jacox,^{1,2} Michael A. Alexander,² Dillon Amaya,² Emily Becker,³ Steven J. Bograd,¹
Stephanie Brodie,^{1,4} Elliott L. Hazen,¹ Mercedes Pozo Buil,^{1,4} and Desiree Tommasi^{4,5}

¹NOAA Southwest Fisheries Science Center, Monterey, CA

²NOAA Physical Sciences Laboratory, Boulder, CO

³University of Miami, Miami, FL

⁴University of California Santa Cruz, Santa Cruz, CA

⁵NOAA Southwest Fisheries Science Center, La Jolla, CA

ABSTRACT

Marine heatwaves (MHWs) – periods of exceptionally warm ocean temperature lasting weeks to years – are now widely recognized for their capacity to disrupt marine ecosystems. The dramatic ecological and socioeconomic impacts of these extreme events present significant challenges to marine resource managers, who would benefit from forewarning of MHWs to facilitate proactive decision making. However, despite extensive research into the physical drivers of MHWs, there has been no comprehensive global assessment of our ability to predict these events. Here, we use a large multi-model ensemble of global climate forecasts to

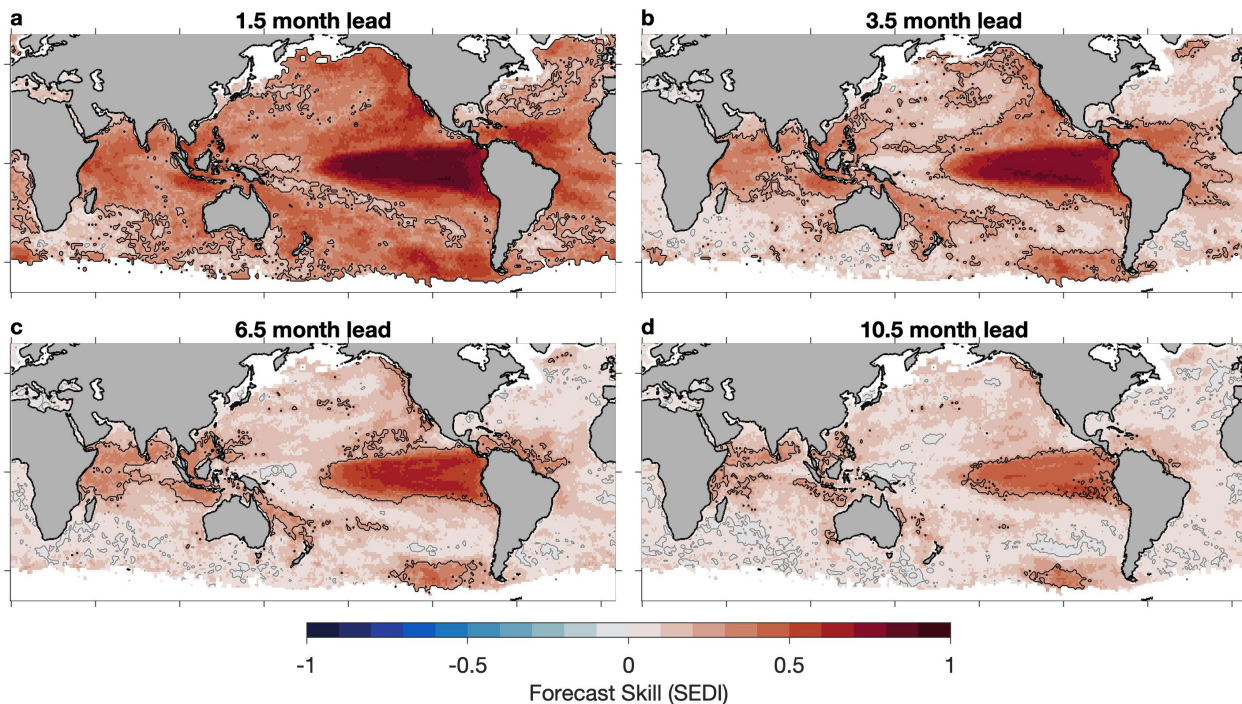


Fig. 1 Skill of global marine heatwave (MHW) forecasts. Maps indicate MHW forecast skill, as measured using the Symmetric Extremal Dependence Index (SEDI), for the 73-member ensemble of forecasts obtained from six global climate forecast systems for the period 1991-2020. SEDI scores range from -1 (no skill) to 1 (perfect skill). Scores above (below) zero, indicated by gray contours, indicate skill better (worse) than chance, and skill that is significantly better than random forecasts at the 95% confidence level is indicated by black contours. MHW forecasts were initialized every month, with lead times up to 11.5 months; a subset of lead times is shown here: (a) 1.5 months, (b) 3.5 months, (c) 6.5 months, and (d) 10.5 months. Areas with permanent or seasonal sea ice coverage are masked in white.

develop and assess MHW forecasts that cover the world's oceans with lead times of up to a year. Using 30 years of retrospective forecasts, we show that the onset, intensity, and duration of MHWs are often predictable, with skillful forecasts possible from 1 to 12 months in advance depending on region, season, and the state of large-scale climate modes such as the El Niño-Southern Oscillation (ENSO). We discuss considerations for setting decision thresholds based on the probability that a MHW will occur, empowering stakeholders to take appropriate actions based on their risk profile. These results highlight the potential for operational MHW forecasts, analogous to forecasts of extreme weather phenomena, to promote climate resilience in global marine ecosystems.

The paper for this study has been published in *Nature* in 2022.

References

Jacox, M. G., M. A. Alexander, D. Amaya, E. Becker, S. J. Bograd, S. Brodie, E. L. Hazen, M. Pozo Buil, and D. Tommasi, 2022: Global seasonal forecasts of marine heatwaves. *Nature*, **604**, 486–490, doi:10.1038/s41586-022-04573-9.

Was the February 2021 Cold Air Outbreak over the Central U.S. a Subseasonal Forecast of Opportunity?

John R. Albers,^{1,2} Matthew Newman,^{1,2} Andrew Hoell,² Sam Lillo,^{1,2}
 Yan Wang,^{1,2} and Melissa Breeden^{1,2}

¹*Cooperative Institute for Research in the Environmental Sciences,
 University of Colorado Boulder, Boulder, CO*

²*NOAA Physical Sciences Laboratory, Boulder, CO*

ABSTRACT

The sources of predictability for the February 2021 cold air outbreak (CAO) over the central United States that led to power grid failures and water delivery shortages are diagnosed using a machine learning-based approach called a linear inverse model (LIM). The flexibility and low computational cost of the LIM allows it to be used as a forecast for identifying and assessing the predictability of key physical processes, and as a climate model used for sensitivity and risk analysis. As a forecast model, the LIM indicates that the February 2021 CAO was a forecast of opportunity, as it accurately predicted both the onset and duration of the CAO four weeks in advance, up to two weeks earlier than leading initialized numerical forecast models (Fig. 1). A LIM-based ‘dynamical filter’ indicates that the February 2021 CAO was principally caused by predictable La Niña teleconnections and unpredictable internal atmospheric variability, with nominally predictable contributions from the Madden-Julian Oscillation and the sudden stratospheric warming the month prior (Fig. 2). As a climate model, the LIM estimates that the February 2021 CAO was in the top 1% of CAO severity, and suggests that similarly extreme CAOs are expected to occur approximately every 20-30 years.

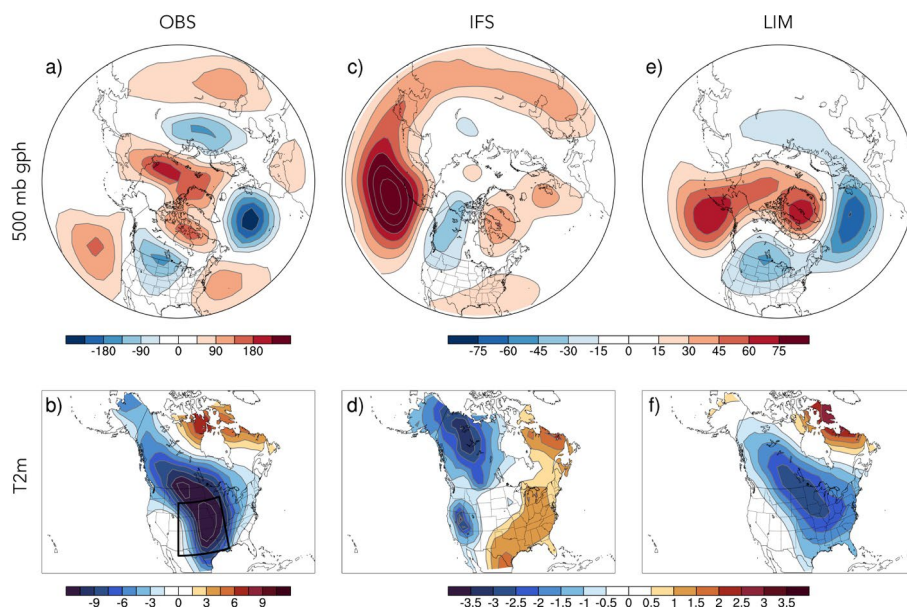


Fig. 1 JRA-55 verifications (left column), ECMWF IFS forecasts (middle column), and LIM forecasts (right column) of 500 mb geopotential height anomalies (top row) and 2m temperature anomalies (bottom row). The JRA-55 verification is shown for 8-21 February 2021. The LIM forecast is initialized on 24 January 2021 and verifies 8-21 February 2021. The IFS forecast (model version CY47R1, operational 2021) is initialized on 25 January 2021 and verifies 9-22 February 2021. Note that the

verification color scale is three times larger than the LIM and IFS color scales. Units are geopotential meters for geopotential height and degrees Celsius for 2m temperature.

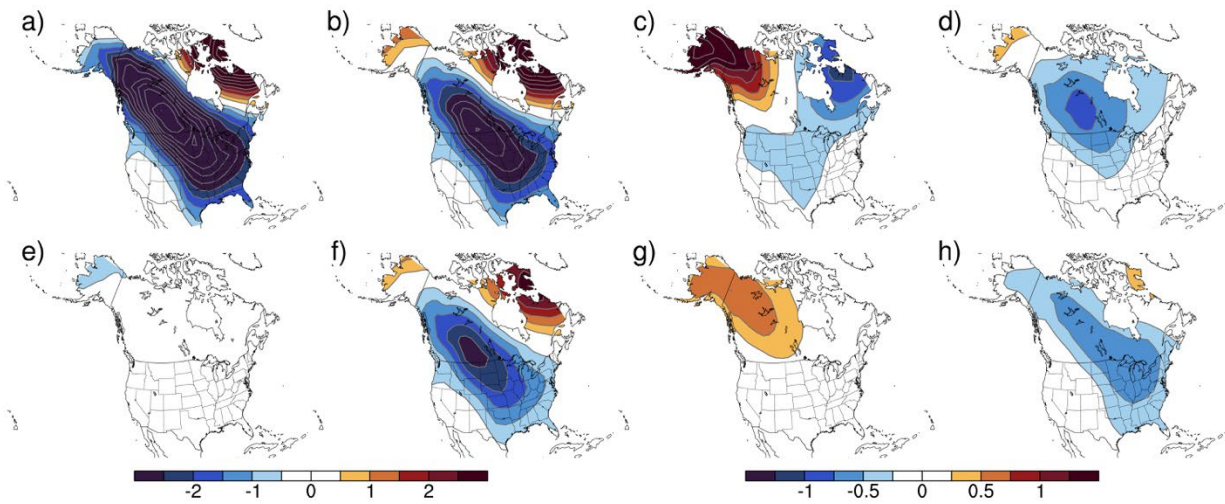


Fig. 2 JRA-55 verifications (top row) and LIM forecasts (bottom row) filtered into contributions from (left to right): internal variability, tropical SSTs, the MJO, and downward propagating stratospheric anomalies. Note the color scale for the internal variability and tropical SST anomalies is two times larger than that of the MJO and stratospheric anomaly color scales. Units are in degrees Celsius.

A Diagnosis of 2018/19 and 2019/20 Winter CPC Outlooks and Model Forecasts

Arun Kumar, Zeng-Zhen Hu, Bhaskar Jha, and Mingyue Chen

Climate Prediction Center, NOAA/NWS/NCEP, College Park, MD

1. Introduction

The verification of seasonal outlooks for December-January-February (DJF) for 2018/19 and 2019/20 made by Climate Prediction Center (CPC) against the observed seasonal mean anomalies presented a contrasting picture. While the Heidke Skill Score (HSS) for DJF 2018/19 was -20, for DJF 2019/20 it was +72 (Fig. 1). Despite the fact that the sea surface temperature (SST) anomalies in the equatorial tropical Pacific were similar, and both were categorized as weak El Niño conditions, the extratropical circulation anomalies (Fig. 2), and surface temperature anomalies over the CONUS had appreciable differences. Upper-level circulation anomalies for DJF 2018/19 in the extratropical latitudes were weak and had a wavy structure; anomalies for DJF 2019/20 had a zonal structure and projected strongly on the Arctic Oscillation (AO) (Fig. 2).

Given similarities between the SST anomalies in the equatorial tropical Pacific across two winters, and differences in the performance of CPC's seasonal surface temperature outlooks raises following questions: (i) What was the role of ocean SST anomalies in shaping the observed seasonal mean anomalies? (ii) What guidance led to differences in CPC outlooks (Fig. 1)? (iii) What was the guidance from multi-model ensemble forecast systems? and (iv) What was the role of the influence of atmospheric initial conditions in DJF mean seasonal forecasts? To explore these questions, a suite of model simulations and forecasts from initialized predictions are used.

2. Analysis procedure

To quantify the contribution of SST anomalies, an ensemble of AMIP simulations forced with observed SSTs are analyzed. As the ensemble mean of AMIP simulations is the atmospheric response to SSTs, this analysis addresses the question of the role of ocean (SST) anomalies in shaping the observed seasonal mean anomalies. To quantify the role of initial conditions in determining the DJF seasonal mean, multi-model ensemble forecasts from the WMO Lead Center for Long-Range Forecasts Multi-Model Ensembles (LC-LRFMME) (<https://wmo.org/>) are used.

3. Results

Shown in Fig. 3 are the atmospheric responses for two DJFs inferred from the ensemble

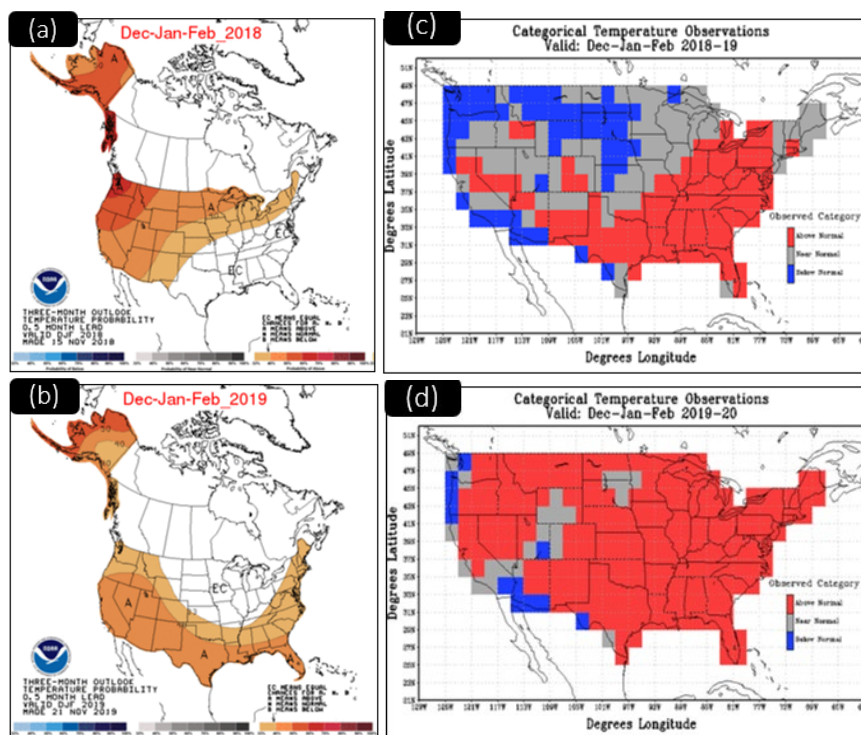


Fig. 1 CPC's outlook for surface temperature for (a) DJF 2018/19, and (b) DJF 2019/20. The category for the verifying observed seasonal means for (c) DJF 2018/19, and (d) 2019/20.

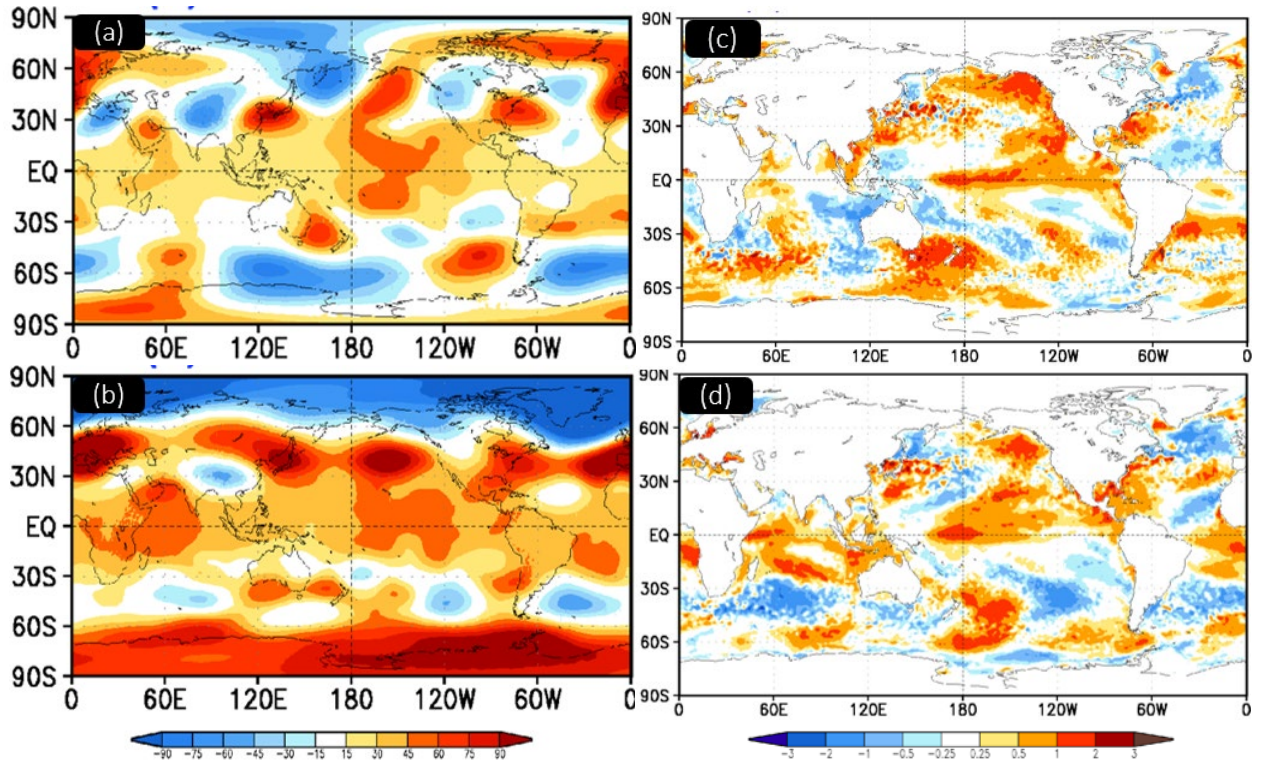


Fig. 2 Observed 200 hPa height anomalies for (a) DJF 2018/19, and (b) 2019/20. Units are in meters. Observed SST anomalies for (c) DJF 2018/19, and (d) 2019/20. Units are in °C.

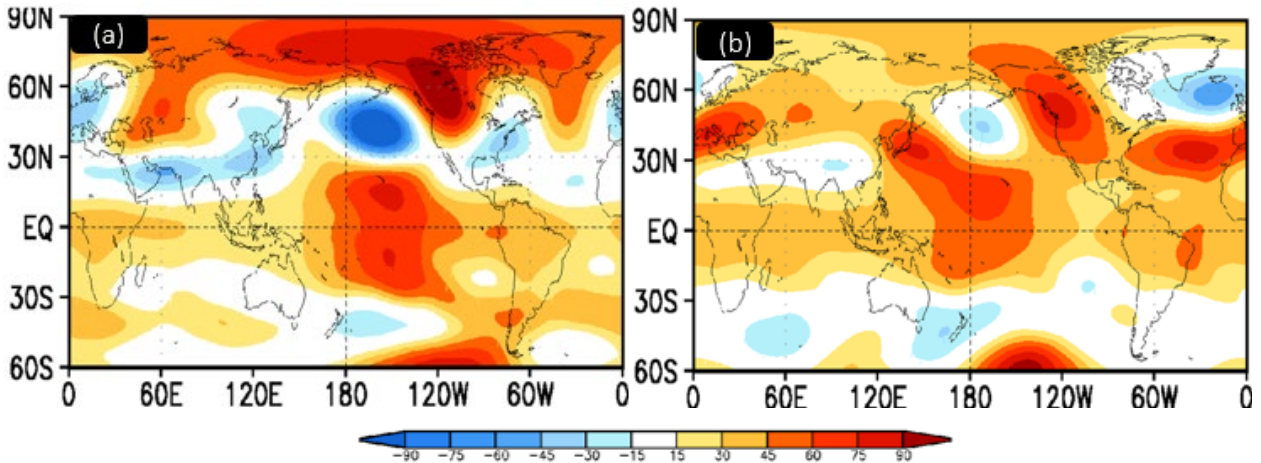


Fig. 3 Ensemble mean 200 hPa height anomalies based on AMIP simulations for (a) DJF 2018/19, and (b) 2019/20. Units are in meters.

mean of 18-member AMIP simulations with NCEP's GFS atmospheric model. For DJF 2018/19, the height response is like one expected during El Niño winter (Fig. 3, left panel). The response, however, differs from the observed anomalies (Fig. 2). Resemblance of atmospheric response during DJF 2019/20 with the expected response during El Niño is less so but the elements of the El Niño signature can still be discerned. A comparison with the observed anomalies (Fig. 2), however, once again shows no similarities in extratropical latitudes. This analysis indicates that atmospheric responses during two El Niño winters, if used as the seasonal prediction, would not have been a good predictor for upper-level heights.

Possible contribution of atmospheric initial conditions can be inferred from initialized seasonal predictions. For this purpose, seasonal mean predictions from the WMO LC-LRFMME are used. These predictions are

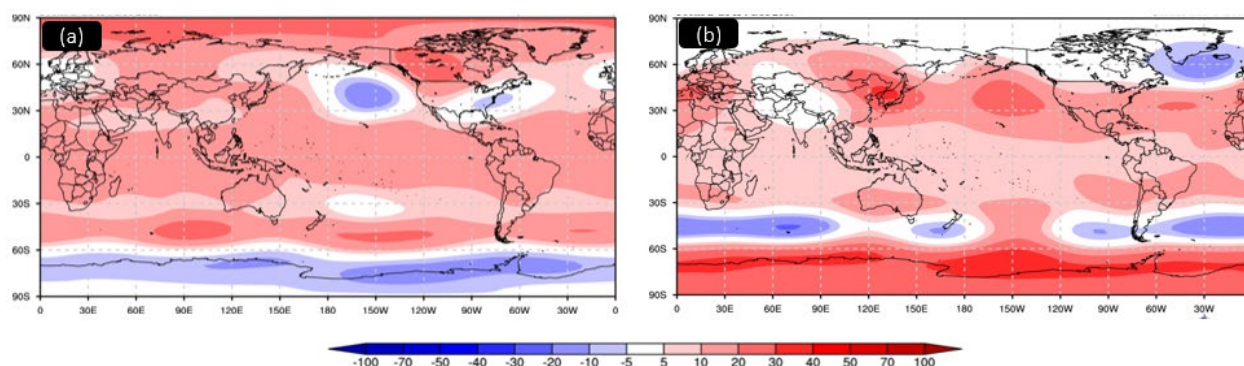


Fig. 4 Same as for Fig. 3 but for initialized MME forecasts from the WMO LC-LRFMME.

based on averaging forecasts from 14 seasonal prediction systems, each one of which itself is an ensemble prediction system. It is noted that predictions for DJF are initialized at the beginning of November.

The multi-model ensemble (MME) predicted anomalies for DJF 2018/19 (Fig. 4, left panel) has large similarities with the atmospheric response in the AMIP simulations (Fig. 3, left panel). This indicates that the role of atmospheric initial conditions for this winter was not significant. As was the case for AMIP simulations, upper-level height anomalies from initialized predictions would not have been a good prediction either.

For DJF 2019/20 MME predicted anomalies (Fig. 4, right panel) differed significantly from the corresponding anomalies in the AMIP simulations (Fig. 3, right panel) indicating possible influence of atmospheric initial conditions contributing to seasonal mean anomalies. Further, there is also a good similarity between predicted upper-level height anomalies and the observed anomalies. Therefore, although the predicted anomalies did not resemble the El Niño response, they were a good forecast for the observed height anomalies in DJF 2019/20.

CPC's seasonal outlooks are based on the various tools that include guidance from MME predictions as well as guidance from various empirical prediction tools. A prognostic discussion also accompanies the release of seasonal outlooks and summarizes the rationale behind the outlook. The prognostic discussion for DJF 2018/19 noted that “—THIS SET OF OUTLOOKS UTILIZED TYPICAL IMPACTS ASSOCIATED WITH EL NINO EVENTS_” The basis of developing the DJF 2018/19 seasonal outlook, therefore, was consistent with the atmospheric response to SSTs (both in the AMIP simulations and in the initialized predictions). This, however, was not the case for the observed anomalies and led to a poor performance for the seasonal outlook.

For DJF 2019/2020 the prognostic discussion noted that “...THE DJF 2019-2020 TEMPERATURE OUTLOOK IS INFORMED BY THE OBJECTIVE CONSOLIDATION, BUT HEAVILY ADJUSTED TO ACCOUNT FOR THE LATEST CFSv2 RUNS, THE INTERNATIONAL MODEL SUITE...” CPC outlook for this winter, therefore, relied heavily on the MME predictions, which were influenced by the atmospheric initial conditions. Predicted anomalies, however, departed appreciably from the expected atmospheric response to El Niño, but had a good resemblance with the observed anomalies. As a consequence, the skill of the CPC's surface temperature outlook during this winter was much better than it was for DJF 2018/19.

4. Discussion

A summary of the inferences drawn from the analysis is shown in Table 1. For DJF 2018/19, AMIP simulations, initialized predictions, and CPC surface temperature outlook were consistent with the El Niño response. However, seasonal mean observed anomalies, likely due to the contribution from the atmospheric internal variability, deviated from the El Niño response. This led to low forecast skill. For DJF 2019/20, while AMIP simulations were consistent with the El Niño response, initialized predictions, either because of the lingering influence of initial conditions or because of the influence of SST anomalies in ocean basins other than the tropical Pacific, deviated from the El Niño response. CPC outlook for surface temperature relied heavily on MME guidance. The observed anomalies for DJF 2019/20 had good resemblance with the initialized predictions leading to a better performance for the CPC outlook.

Table 1 A table with the synthesis of various analyses.

Analysis	2018/19	2019/20
AMIP ensemble mean	El Niño response	Somewhat like El Niño response.
Initialized MME predictions	El Niño response Small influence from initial conditions	Not like El Niño response but has similarity with observed anomalies. Initial condition information or influence of SSTs in another ocean basin, e.g., Indian Ocean?
CPC outlooks	El Niño factored in the outlook ...also the MME outlook	El Niño did not factor in the outlook... informed by MME.
Observations	Not like El Niño response. Likely due to atmospheric noise/internal variability.	Not like El Niño response BUT like MME forecasts
CPC outlook performance	Worse	Better

The analysis also raises following questions: (i) What the potential role of SST anomalies was in the ocean basins other than in the equatorial tropical Pacific, *e.g.*, in the Indian Ocean where SST anomalies over two winters had appreciable differences? (ii) Were the initial conditions in 2019/20, particularly related to the positive phase of the AO, unusually persistent that led to a better prediction?

The 2020-21 Southwestern U.S. Drought: A Preliminary Study on Its Causes and Prediction

Hailan Wang, Arun Kumar, Brad Pugh, Mingyue Chen, and Bala Narapusetty
 Climate Prediction Center, NOAA/NWS/NCEP, College Park, MD

1. Introduction

Historic drought conditions developed across the southwestern United States (U.S.) in late summer 2020 and persisted through summer 2021. According to the U.S. Drought Monitor (USDM, <https://droughtmonitor.unl.edu/>) (Fig. 1), the U.S. Southwest experienced a rapid drought intensification and expansion from July to November 2020. For the subsequent months through summer 2021, nearly the entire (~97%) Southwest was in drought (D1-D4), with ~70% of the area having severe and exceptional drought conditions (D3-D4), making the event the most severe southwestern drought in the USDM record. The rapid drought development during July-August-September (JAS) 2020 is particularly noteworthy, as the Southwest experienced the record driest and warmest conditions since 1895 (NCEI, <https://www.ncei.noaa.gov/access/monitoring/us-maps/>), due to the record low North American monsoon rainfall and extreme heat. In the following months through summer 2021, below-normal precipitation and above-normal temperature anomalies continued, which, while not as severe as those in JAS 2020, helped sustain the severe drought conditions (Mankin *et al.* 2021). The 2020-21 southwestern drought has been reported to cause substantial socioeconomical impacts, including those on local water supplies, energy production, farming, livestock operations and increased wildfire risks.

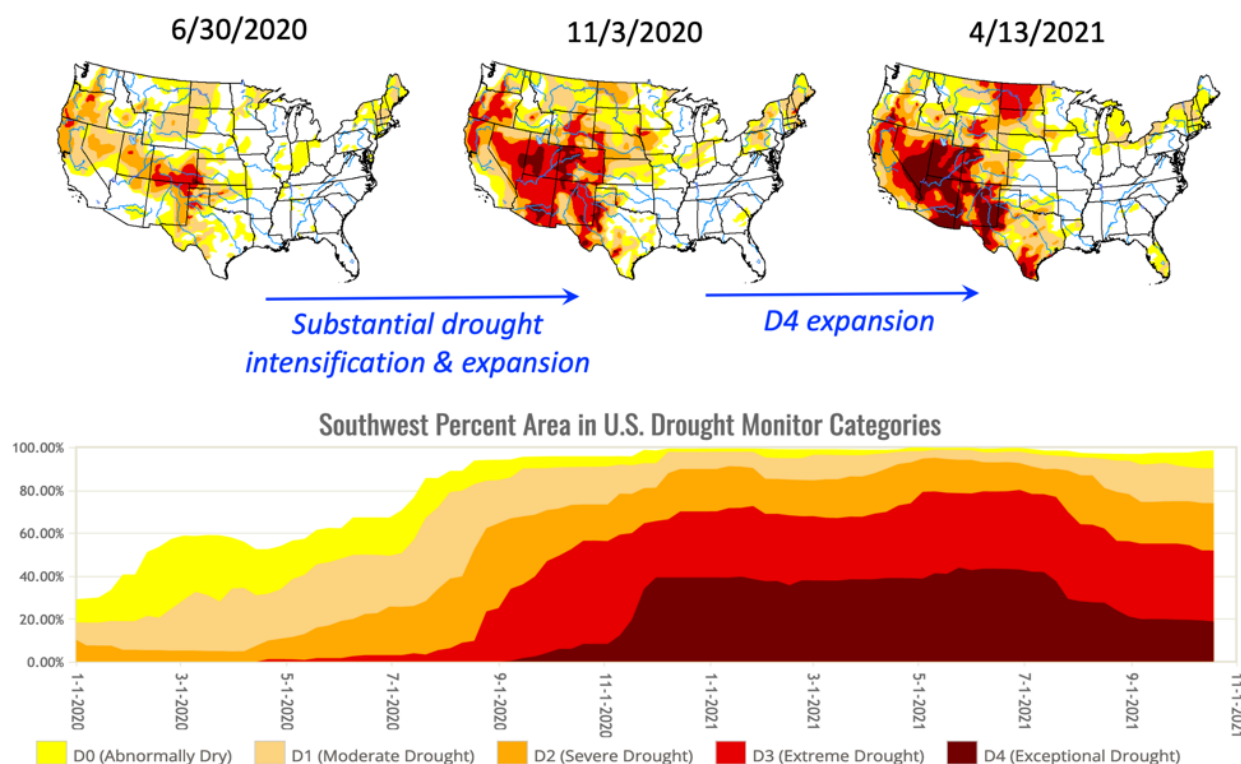


Fig. 1 The 2020-21 southwestern U.S. drought in the U.S. Drought Monitor (USDM, <https://droughtmonitor.unl.edu/>)

In this study, we performed a preliminary investigation on the causes and prediction of the 2020-21 southwestern U.S. drought, with the focus on the rapid drought development during July-October 2020. The findings are applied to better understand the predictability of the drought and the performance of CPC drought outlooks (DOs).

2. Data and method

The proximate causes of the drought were investigated by assessing the separate effects of observed precipitation (P) and temperature (T) anomalies during July-October 2020 on concurrent and subsequent evolution of land surface anomalies. Three land surface model experiments, consisting of a control experiment and two anomaly experiments, were performed using the Variable Infiltration Capacity (VIC) hydrology model (v4.0.6, Wood *et al.* 2002; Mo *et al.* 2012). The control experiment (PT_Anom) was forced with observed daily P and T from CPC unified observations (Chen *et al.* 2008), and ran from January 1, 1979 through March 31, 2021. The two anomaly experiments (P_Anom, T_Anom) started from June 30, 2020, initialized using the land state of the same date from the control experiment, and ran through March 31, 2021. The experiment P_Anom was forced with observed daily P and daily climatological T, whereas the experiment T_Anom was driven by observed daily T and daily climatological P. The comparison of the three experiments focuses on their daily soil moisture percentiles (SMPs) during July 2020-March 2021. Here the SMPs were computed using a 15-day moving window of soil moisture values, they were based on the empirical probability density function (PDF) built using the control simulation for the period of 1979-2019.

The causes of the southwestern U.S. drought were further investigated by diagnosing the maintenance of observed drought-inducing atmospheric circulation anomalies during JAS 2020 and identifying key regional forcings, using a stationary wave model (SWM) and the NCEP/NCAR reanalysis. Here the SWM is based on the three-dimensional (3-D) primitive equations in *sigma* coordinates. It is time-dependent and resolves stationary nonlinearity (Ting and Yu 1998). The model has rhomboidal wavenumber-30 truncation in the horizontal and 14 unevenly spaced *sigma* levels in the vertical. The NCEP/NCAR reanalysis was used to compute observed atmospheric circulation anomalies as well as basic state and stationary wave forcing inputs for the SWM.

For understanding contributions from different factors, stationary wave forcing anomalies, consisting of those of diabatic heating and transient flux convergences, were first computed (Wang and Ting 1999). The SWM was then used to diagnose the relative roles of regional diabatic heating and transient forcing anomalies in maintaining the drought-inducing circulation anomalies. Additionally, an optimal forcing pattern analysis was performed to identify forcing regions and patterns that are most effective in driving the observed drought-inducing circulation anomalies (Schubert *et al.* 2011). This proceeded by performing a set of 1188 SWM runs, in which idealized heating sources were introduced every 10° longitudes and 5° latitudes across the globe. The horizontal structure of the idealized heating anomaly has a sine-squared function form, with horizontal scales of 20° longitudes and 10° latitudes and its vertical profile having a peak of 1.4 K/day in the middle troposphere. The three-dimensional JAS climatological flow was used as the background state in the SWM runs. To obtain the optimal forcing pattern, the inner products between each of the low-level SWM responses and the observed low-level meridional wind anomaly in the monsoon region was computed and placed at the forcing location.

To study prediction of the southwestern U.S. drought, we focused on NOAA official drought and precipitation outlooks produced and issued by the Climate Prediction Center (CPC). These outlook products are produced by integrating short-term and long-term dynamical model forecasts, statistical tools, climatologies and analogs, as well as feedback from stakeholders. The verifications of the outlooks are routinely performed using their respective observations. It is worth noting that the CPC DOs are initialized using the weekly U.S. drought monitor (USDM). They are presented using national maps showing drought tendency at monthly and seasonal lead times for four categories: drought

persists, drought remains but improves, drought removal likely, and drought development likely. The skill performance of the DOs is evaluated using the temporal changes of the USDM.

3. Results

The causes of the 2020-21 southwestern U.S. drought were first investigated by assessing the separate effects of precipitation deficits and extreme heat during July-October 2020 on the evolution of soil moisture percentiles. It is evident from Fig. 2 that the precipitation deficits play a predominant role in rapidly drying local soil moisture, intensifying and shaping the southwest U.S. drought. While the temperature warming contributes as well, its effect is modest.

Given the importance of the precipitation deficits in driving the drought, we next focused on low-level atmospheric circulation anomalies, as they play a dominant role in regulating atmospheric moisture transport and hence precipitation. Figure 3a shows that JAS 2020 features substantially weakened low-level southerlies in the North American monsoon region. These southerlies are known to be crucial in transporting abundant moisture northward inland from the Gulf of California and the Gulf of Mexico as well as nearby oceans, facilitating the formation of thunderstorms in the southwestern U.S. The substantial weakening of the southerlies during JAS 2020 would reduce the frequency of occurrence of the thunderstorms, leading to monsoon rainfall deficits.

Because of the importance of the low-level southerlies, we used the SWM to diagnose their maintenance and assess the separate roles of regional stationary wave forcings. Figure 3 shows that the main forcing contributors are the diabatic heating anomalies in the monsoon and nearby regions. By comparison, the heatings in the rest of the global regions and transient flux convergence play secondary roles.

Through the optimal forcing pattern analysis, Fig. 4 further identifies heating regions and patterns that are most effective in driving the weakened southerlies. Figure 4a shows that the most effective regional heating anomalies consist of the cooling anomalies along the southwest coast of North America and the heating anomalies in the south-central US and the Gulf of Mexico. Since the optimal forcing pattern analysis assumes the magnitudes of idealized heating anomalies in

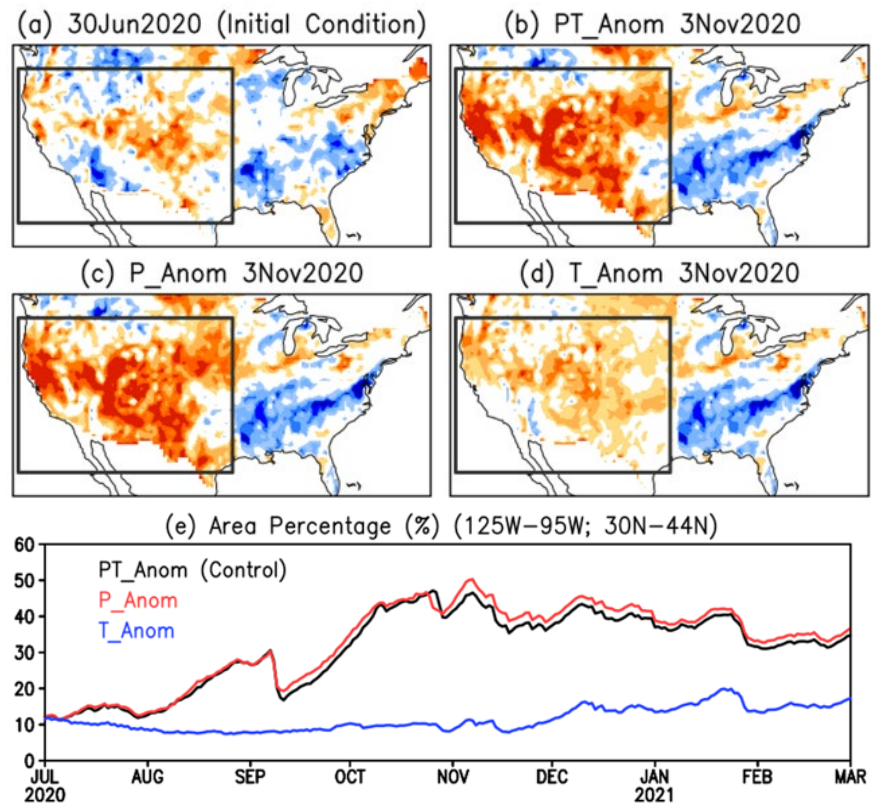


Fig. 2 VIC land surface model experiments to assess the separate roles of observed precipitation and temperature anomalies during July-October 2020 in contributing to the 2020-21 southwestern U.S. drought development. (a) Soil Moisture Percentile (SMP) for June 30, 2020 in the control experiment (PT_Anom) forced with observed precipitation and temperature. (b) SMP for November 3, 2020 in the control experiment (PT_Anom). (c) Same as (b) but for the VIC experiment forced with observed precipitation for July 1-October 31, 2020 (P_Anom). (d) Same as (b) but for the VIC experiment forced with observed temperature for July 1-October 31, 2020 (T_Anom). (e) The comparison of the three VIC experiments in the percent area of the western U.S. (125°W-95°W, 28°N-46°N, black box in panel a) for D2-D4 drought categories (SMP ≤ 10%).

the SWM runs to be identical for all forcing locations, the optimal forcing pattern in Fig. 4a was further scaled using the observed diabatic heating anomalies (Fig. 3f) to identify the actual regional heating anomalies that contribute (Fig. 4b). Figure 4b highlights the dominant importance of diabatic heating anomalies in the monsoon region, eastern tropical Pacific off the west coast of U.S. southwest and Mexico, as well as the Gulf of Mexico, in driving the low-level weakened southerlies. Work is currently underway to investigate the observed subseasonal physical processes that led to these regional diabatic heating anomalies.

Turning to drought prediction, an examination of CPC Seasonal Drought Outlooks (SDOs) shows that the biggest challenge lies in predicting the development phase of the drought. The drought development was primarily driven by the record low monsoon precipitation and excessive heat. Its prediction thus critically depends on the forecast skills for these extreme precipitation and temperature anomalies. The DOs would not be able to capture the rapid drought development if the precipitation deficits and warming anomalies are not skillfully forecast.

As exemplified by the SDO for August-October 2000, the SDO missed the rapid drought development in Arizona and New Mexico (*cf.* Fig. 5a with Fig. 5b, also see the SDO verification in the right panel of Fig. 5a), due to the limited skills in capturing the precipitation deficits in these regions, as shown in the seasonal precipitation outlook (Fig. 5c). The limited forecasting skills of the monsoon rainfall deficits could be in part due to their driving physical processes being inherently unpredictable at seasonal lead-time, and in part

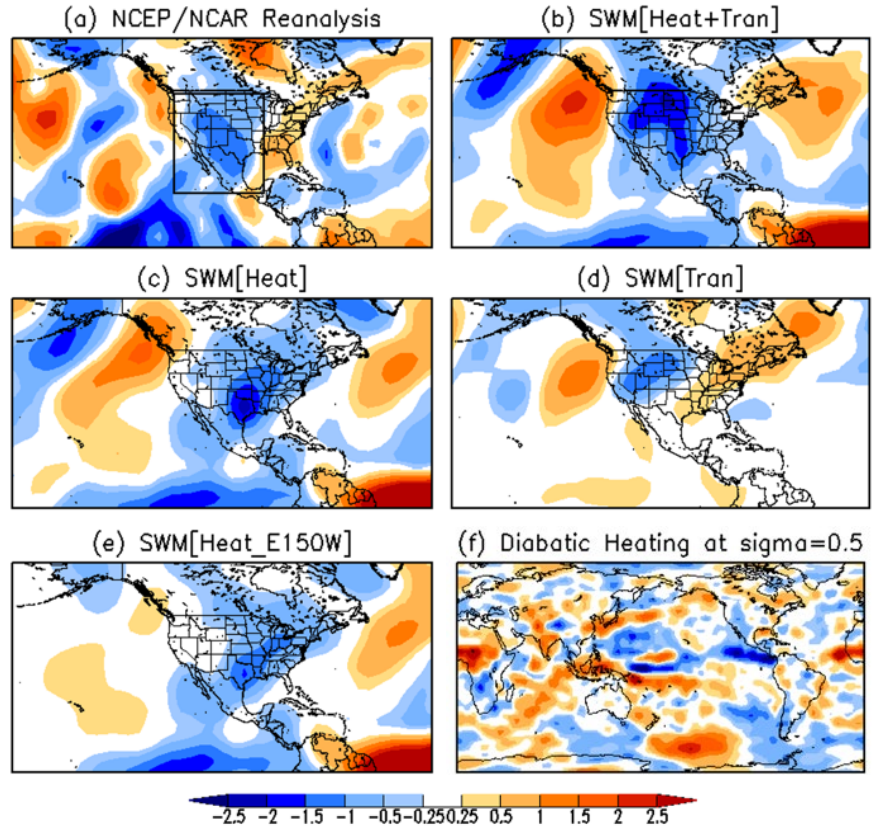


Fig. 3 Stationary wave modeling diagnosis of the low-level weakened southerly in the monsoon region during July-August-September (JAS) 2020. (a) Zonally asymmetric meridional wind anomaly (m/s) at $\sigma=0.866$ in the NCEP/NCAR reanalysis. (b) Stationary wave model (SWM) response at $\sigma=0.866$ to the sum of anomalous diabatic heating and transient flux convergences. (c) Same as (b) but for the SWM response to anomalous heating only. (d) Same as (b) but for the SWM response to anomalous transient forcing only. (e) Same as (b) but for the SWM response to the anomalous heating east of 150°W . (f) Residually derived diabatic heating (K/day) at $\sigma=0.5$.

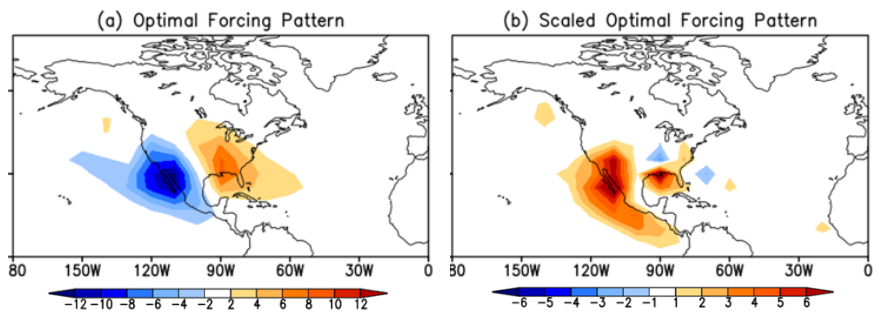


Fig. 4 (a) Optimal pattern of diabatic heating forcing for the observed JAS 2020 low-level weakened southerly in the monsoon region (black box in Fig. 3a). (b) Scaled optimal forcing pattern, obtained by multiplying (a) with the residual diabatic heating anomaly at $\sigma=0.5$ in Fig. 3f. Units are arbitrary.

in part due to their driving physical processes being inherently unpredictable at seasonal lead-time, and in part

due to the limitations of forecast tools (*e.g.*, biases in the NMME forecast models). The predictability of the JAS 2020 monsoon rainfall deficits and their driving physical processes requires further investigation. By comparison, after drought conditions are established, the SDOs became more skillful, as they can draw a considerable portion of the skill from the persistence of dry land initial states (Fig. 5d).

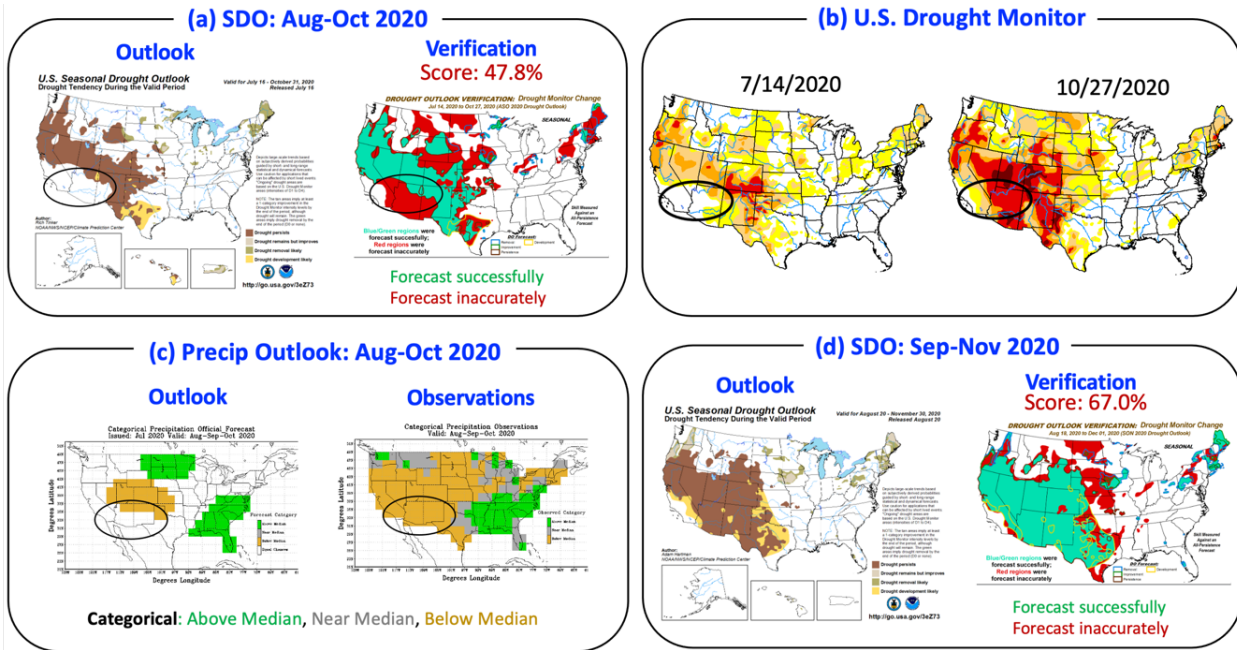


Fig. 5 (a) NOAA CPC Seasonal Drought Outlook (SDO) for August-October 2020, and its verification using the U.S. Drought Monitor (USDM). (b) The comparison of USDM maps between July 14, 2020 and October 27, 2020. (c) NOAA CPC Precipitation Outlook for August-October 2020, and its categorical verification using observations. (d) Same as (a) but for September-November 2020. The CPC outlook maps are taken from CPC website. The USDM maps are downloaded from <https://droughtmonitor.unl.edu/>

4. Summary and discussion

This study performed a preliminary investigation of the causes and prediction of the 2020-21 southwestern U.S. drought. The focus is on the rapid drought intensification in July – October 2020, during which the region experienced record driest conditions and extreme heat. The rapid drought development, as manifested in soil moisture decline, was found to be primarily driven by the record low North American monsoon rainfall and secondarily by the extreme heat. The strong precipitation deficits are closely linked to the considerably weakened low-level southerlies in the monsoon region, which reduced the northward inland moisture transport from the Gulf of California and the Gulf of Mexico, and hence, monsoon rainfall. Through a stationary wave modeling diagnosis, we found that the weakened southerlies are mainly driven by diabatic cooling anomalies in the monsoon region and regions off the southwest coast of North America as well as diabatic heating anomalies in the Gulf of Mexico.

Turning to drought prediction, the main challenge of CPC SDOs lies in predicting the rapid drought development in JAS 2020. The limited forecast skill for the drought development results from the difficulty in skillfully predicting its meteorological drivers, particularly the record low monsoon precipitation. While the low skill for precipitation may be due in part to the limitations of forecast tools (*e.g.*, biases in dynamical forecast models), it is likely that some fraction of monsoon precipitation variability is inherently unpredictable at seasonal lead-time. With the existing outlook product resources at CPC, the forecast challenges of SDOs can be in part remedied by considering drought and precipitation forecasts at shorter lead-time (*e.g.*, Monthly

Drought Outlook - MDO). Work is also currently underway at CPC to develop probabilistic DOs to provide users and stakeholders with probability information of drought forecasts and facilitate their decision-making.

References

- Chen, M., W. Shi, P. Xie, V. B. S. Silva, V. E. Kousky, R. W. Higgins, and J. E. Janowiak, 2008: Assessing objective techniques for gauge-based analyses of global daily precipitation, *J. Geophys. Res.*, **113**, D04110, doi:10.1029/2007JD009132.
- Mankin, J. S., I. Simpson, A. Hoell, R. Fu, J. Lisonbee, A. Sheffield, and D. Barrie, 2021: *NOAA Drought Task Force Report on the 2020-2021 Southwestern U.S drought*. NOAA Drought Task Force, MAPP, and NIDIS, 20 pp. <https://cpo.noaa.gov/MAPP/DTF4SWReport>.
- Mo, K., S. Shukla, D. P. Lettenmaier, and L. Chen, 2012: Do Climate Forecast System (CFSv2) forecasts improve seasonal soil moisture prediction? *Geophys. Res. Lett.*, **39**, L23703.
- Schubert, S., H. Wang and M. Suarez, 2011: Warm season subseasonal variability and climate extremes in the Northern Hemisphere: The role of stationary Rossby waves. *J. Climate*, **24**, 4773–4792.
- Ting, M. and L. Yu, 1998: Steady response to tropical heating in wavy linear and nonlinear baroclinic models. *J. Atmos. Sci.*, **55**, 3565–3582.
- Wang H. and M. Ting, 1999: Seasonal cycle of climatological stationary waves in the NCEP/ NCAR Reanalysis. *J. Atmos. Sci.*, **56**, 3892-3919.
- Wood A., E. P. Maurer, A. Kumar, and D. P. Lettenmaier, 2002: Long-range experimental hydrologic forecasting for the eastern United States. *J. Geophys. Res.*, **107**, 4429.

Impact of the MJO on the Forecast Skill of Week-2 Severe Weather over the United States

Hui Wang, Arun Kumar, Wanqiu Wang, David DeWitt, and Matthew Rosencrans
 Climate Prediction Center, NOAA/NWS/NCEP, College Park, MD

1. Introduction

The Madden-Julian oscillation (MJO) is a dominant mode of tropical subseasonal variability (Madden and Julian 1971) and is also known to influence severe weather over the United States (*e.g.*, Baggett *et al.* 2018). A dynamical-statistical model has been developed at the NOAA Climate Prediction Center (CPC) for forecasting week-2 severe weather (Wang *et al.* 2021). Here we examine the impact of the MJO on both weekly severe weather activity over the U.S. and week-2 forecast skill.

The goals of this study are (1) to document weekly (7-day total) severe weather activity conditional to the phase of the MJO, (2) to examine the forecast skill of week-2 severe weather during different phases of the MJO, and (3) to identify the regions where tropical heating associated with the MJO may favor severe weather over the U.S.

2. Data and methods

The data used in this study include both observational data and model forecasts. For observations, the NCEP Climate Forecast System Reanalysis (CFSR) and NWS local storm reports (LSRs) are used. The LSR consists of hail, tornado, and damaging wind reports, as well as their location, time, and intensity. The sum of the LSRs for hail, tornado and damaging wind are referred to as LSR3 hereafter. All data are re-gridded to a $0.5^\circ \times 0.5^\circ$ grid.

The forecast model used in this study is a hybrid dynamical-statistical model (Wang *et al.* 2021). It uses the dynamical model (NCEP GEFSv12) predicted environmental variables as a predictor to forecast severe weather (LSR3) based on the statistical relationship between the predictor and actual LSR3 in historical records. We use the GEFSv12 16-day hindcasts to train the model for week-2 severe weather. The hindcast period is from 2000 to 2018 with 5 members daily. The

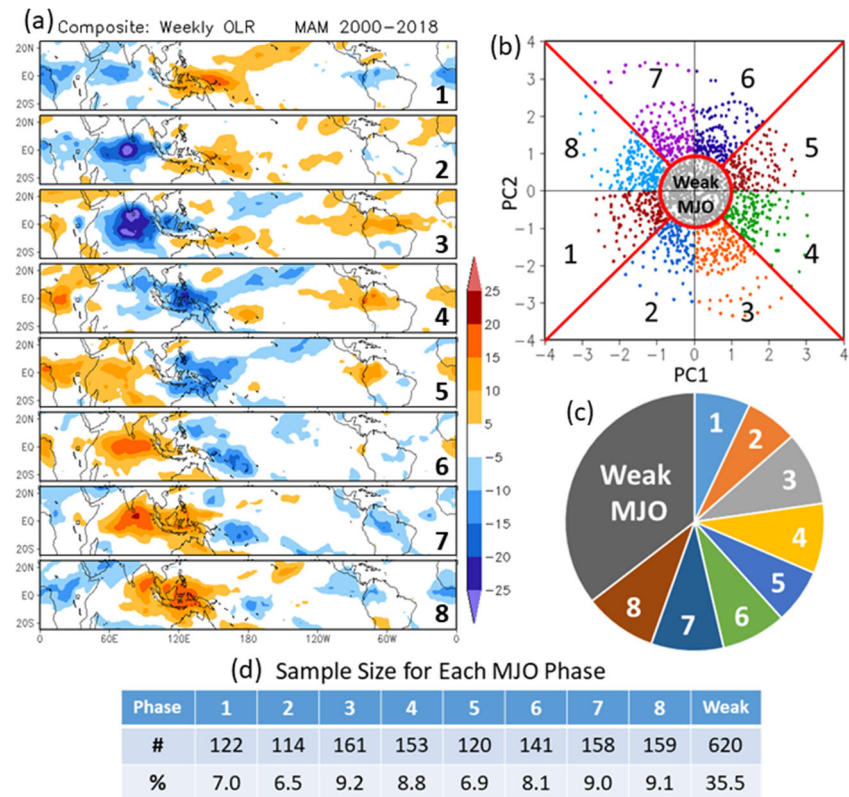


Fig. 1 (a) Composites of weekly OLR anomalies for eight MJO phases, (b) phase-space diagram of the PC time series of two leading EOFs with combined OLR, U200 and U850 averaged over 15°S – 15°N along the equator, (c) pie chart of sample size for each MJO phase, also listed in (d). The numbers (1–8) denote the MJO phases.

analysis is performed using the 5-member ensemble mean forecasts. Following Carbin *et al.* (2016), the predictor, supercell composite parameter (SCP), is defined as

$$\text{SCP} = (\text{CAPE}/1000 \text{ J kg}^{-1}) \times (\text{SRH}/50 \text{ m}^{-2} \text{ s}^{-2}) \times (\text{BWD}/20 \text{ m s}^{-1}),$$

where CAPE is convective available potential energy, SRH storm-relative helicity, and BWD bulk wind difference. The three constants are used to normalize SCP such that when SCP is greater than one, severe weather is likely to occur. Both forecasts and analysis focus on the spring season from March to May (MAM).

The MJO is represented by two leading EOFs of combined OLR and 200 and 850-hPa zonal winds averaged over 15°S–15°N along the equator (*e.g.*, Kessler 2001), taken from the CFSR. A time-dependent linear barotropic model (*e.g.*, Ting 1996) is applied to the 200-hPa level to test the sensitivity of the extratropical circulation response to tropical heating associated with the MJO.

3. Results

The characteristics of the MJO are examined first. Figure 1a shows the composites of weekly OLR anomalies associated with each phase of the MJO during MAM. A negative OLR anomaly indicates enhanced convection, which is related to a heating anomaly for the atmosphere. As the MJO propagates eastward from phase 1 to phase 8, the location of the heating also shifts towards the east. The sample size of each MJO phase is summarized in both a phase-space diagram (Fig. 1b) and a pie chart (Fig. 1c), as well as in a table (Fig. 1d), and ranges from 6.5% to 9.2% of the total events. Among all cases, 35.5% are weak MJO events, which are excluded from the analysis.

Figure 2 shows the composites of LSR3 anomalies associated with the eight MJO phases. Overall, there is enhanced severe weather activity in the eastern and central U.S. from phase 1 to phase 4 and suppressed storm activity from phase 5 to 8. The anomaly correlation skill of the hybrid model for week-2 severe weather is shown in Fig 3 (top panel) for the entire MAM 2000–2018. Relatively high skill values are found along or near the Tornado Alley. The bottom panels are the changes in the skill from the top panel for each MJO phase. Compared to Fig. 2, the forecast skill is generally increased (decreased) over regions with the enhanced (suppressed) severe weather activity linked to the MJO.

Figure 4 shows the 200-hPa streamfunction in the linear barotropic model response to heating over the tropical Indian Ocean from day 1 to day 15. The tropical heating (Fig. 4, the separate panel on the right) is prescribed as an anomalous divergence field in the barotropic vorticity equation. It takes about 10 days for the

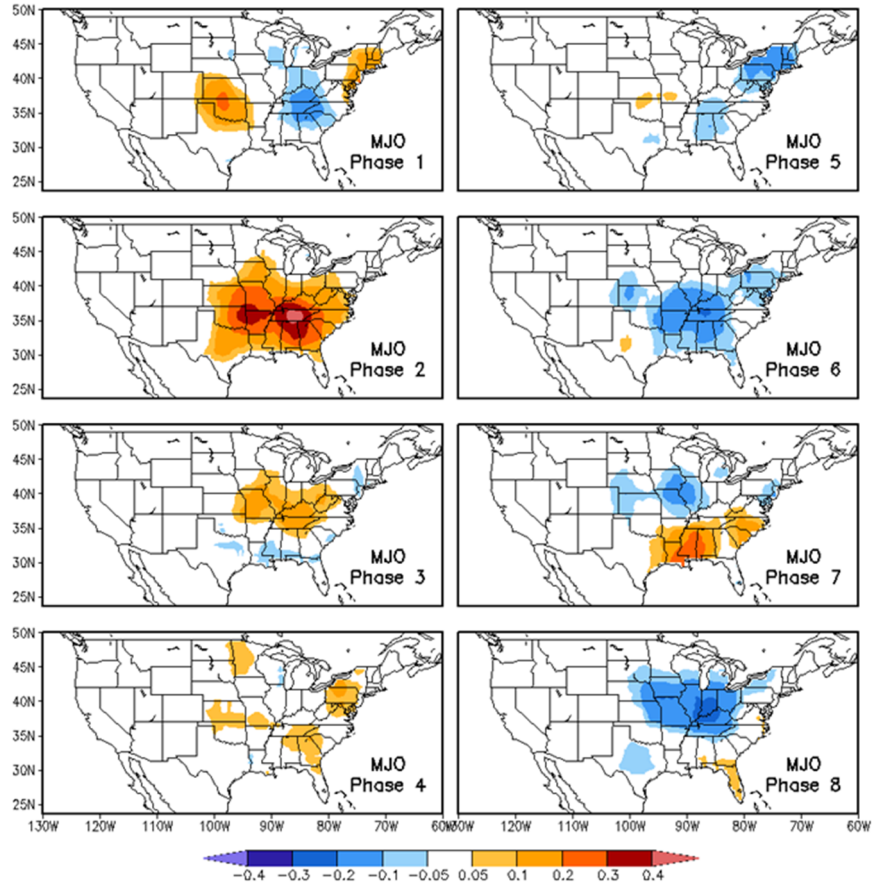


Fig. 2 Composites of weekly LSR3 anomalies for each MJO phase during MAM 2000–2018.

It takes about 10 days for the

extratropical circulation to reach a nearly steady response over North America. Therefore, an *extratropical* circulation concurrent with MJO phase 2 may be forced by MJO phase 1, because of this delayed response.

The circulation anomalies associated with the enhanced severe weather activity (Fig. 2) are examined in Fig. 5, which shows the composites of weekly LSR3 during MJO phase 2 (Fig. 5a) and the associated 200-hPa height anomalies (Fig. 5b). A well-defined wave train crosses East Asia and North America. Downstream, there is an anomalous ridge near the U.S. southeast coast (Fig. 5b, red box), which can enhance a low-level jet and transport moisture from the Gulf of Mexico, favorable for severe weather over the central and eastern U.S. (Fig. 5a).

Whether the tropical heating associated the MJO can generate such a circulation pattern is tested using the barotropic model forced by an idealized heating prescribed in the model. Figures 6a and 6b are two examples of the circulation response to a point heating (red dot). The response over the red box area near the southeast coast is positive when the heating anomaly is over the tropical Indian Ocean (Fig. 6a), but negative when the heating is over the eastern tropical Pacific (Fig. 6b).

An influence function is used to determine the locations where tropical heating may generate circulation anomalies favorable for severe weather over the central and eastern U.S. The influence function is a spatial map derived from individual circulation responses to heating at each grid point. The value of the influence function at a grid point is the value of the circulation anomalies over a specific region in response to the point heating at the same grid point. Here, the specific region is the area near the southeast coast (red box in Figs. 5b, 6a, 6b). To obtain a global distribution of the influence function, the barotropic model is run each time with a point heating specified at one grid point and repeated for all grid points.

Figure 6c shows the influence function for the area we are interested in (red box in Figs. 5b, 6a, 6b). The influence function indicates that heating over the western Indian Ocean can generate positive streamfunction anomalies near the southeast coast but heating over the eastern Indian Ocean and western Pacific generates negative anomalies. Based on the distribution of the influence function (Fig. 6c), both the tropical heating (negative OLR anomalies, warm color in Fig. 6d) and cooling (positive OLR anomalies, cold color in Fig. 6d)

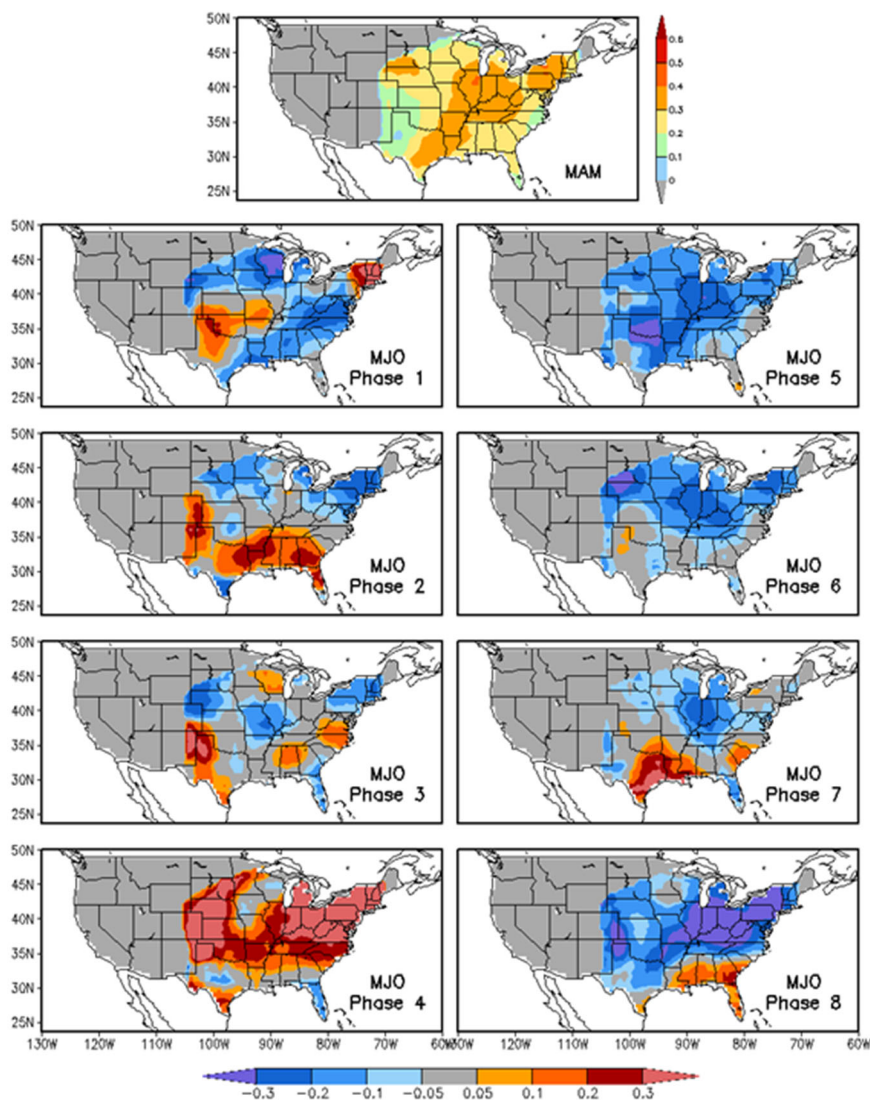


Fig. 3 Anomaly correlation (AC) skill of week-2 severe weather for entire MAM 2000–2018 (top panel) and changes in the AC skill from the top panel for each MJO phase (bottom panels).

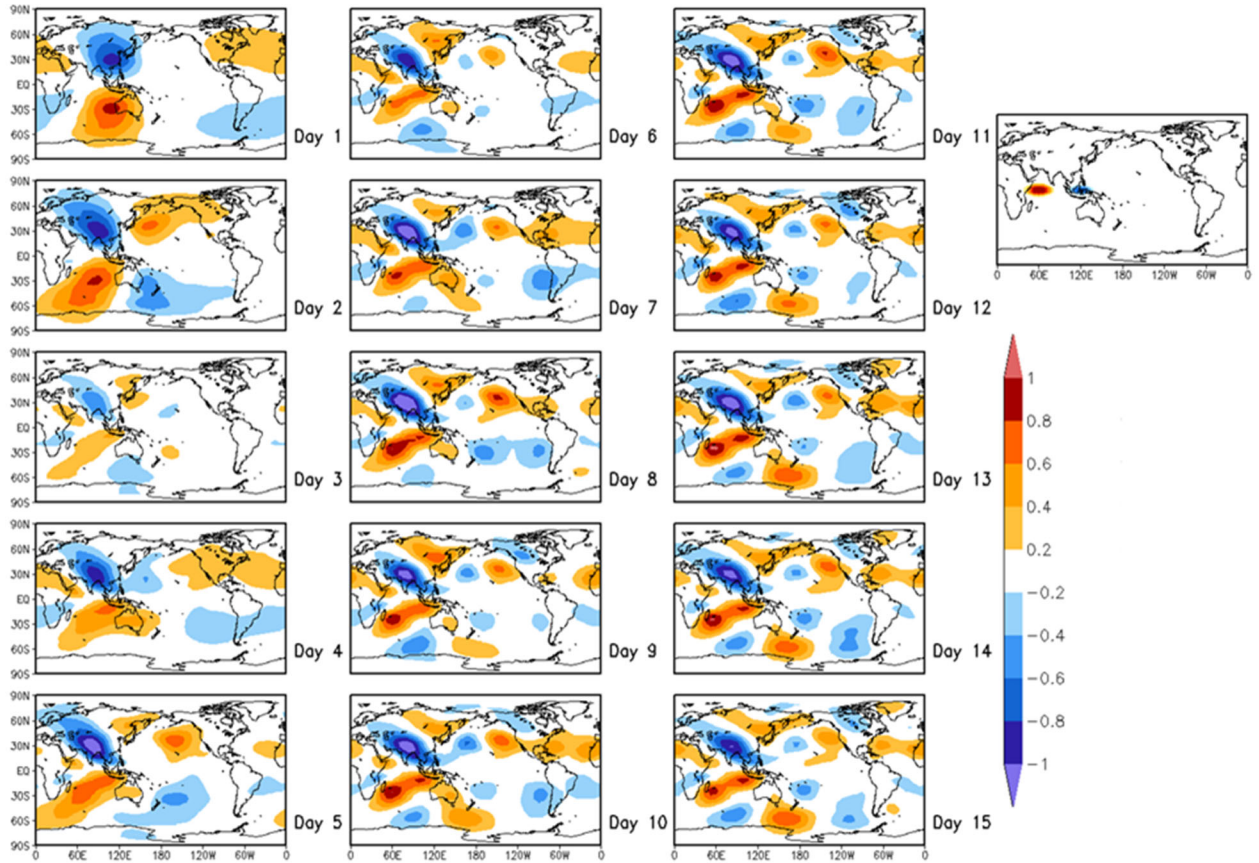


Fig. 4 200-hPa streamfunction response to the tropical heating from day 1 to day 15. The heating is prescribed as divergence anomalies in the linear barotropic model.

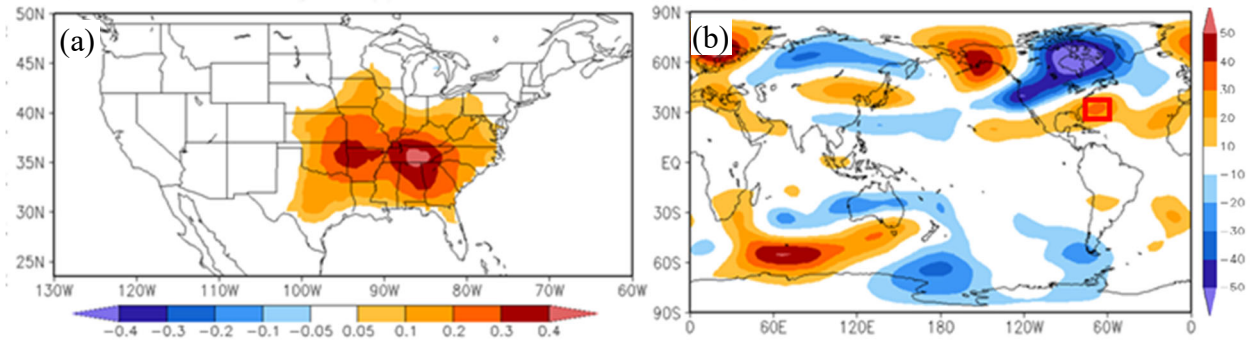


Fig. 5 Composites of (a) weekly LSR3 anomalies and (b) 200-hPa height anomalies for MJO phase 2 during MAM 2000–2018. The red box denotes the region of an anomalous ridge.

associated MJO phase 1 can produce positive height anomaly near the southeast coast, favorable for the severe weather over the central and eastern U.S.

4. Summary

The impact of the MJO on weekly severe weather activity and week-2 forecast skill was examined in this study. The composite analysis indicates that weekly severe weather over the eastern and central U.S. is enhanced during MJO phases 1–4. The week-2 forecast skill can be increased in the areas of enhanced severe weather related to the MJO. A linear barotropic model is used to illustrate that it takes about 10 days for the extratropical circulation to reach a steady response to tropical heating. The influence function identifies the specific regions where tropical heating can force the extratropical circulation pattern that is favorable for severe

weather over the U.S. The results are consistent with both the relationship between severe weather activity and MJO phase and the location of the tropical heating associated with the MJO phase.

References

- Baggett, C. F., K. M. Nardi, S. J. Childs, S. N. Zito, E. A. Barnes, and E. D. Maloney, 2018: Skillful subseasonal forecasts of weekly tornado and hail activity using the Madden-Julian oscillation. *J. Geophys. Res. Atmos.*, **123**, 12,661–12,675.
- Carbin, G. W., M. K. Tippett, S. P. Lillo, and H. E. Brooks, 2016: Visualizing long-range severe thunderstorm environment guidance from CFSv2. *Bull. Amer. Meteor. Soc.*, **97**, 1021–1032.
- Kessler, W. S., 2001: EOF representations of the Madden-Julian oscillation and its connection with ENSO. *J. Climate*, **14**, 3055–3061.
- Madden, R. A., and P. R. Julian, 1971: Description of a 40–50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.*, **28**, 702–708.
- Ting, M., 1996: Steady linear response to tropical heating in barotropic and baroclinic models. *J. Atmos. Sci.*, **53**, 1698–1709.
- Wang, H., A. Kumar, A. Diawara, D. DeWitt, and J. Gottschalck, 2021: Dynamical–statistical prediction of week-2 severe weather for the United States. *Wea. Forecasting*, **36**, 109–125.

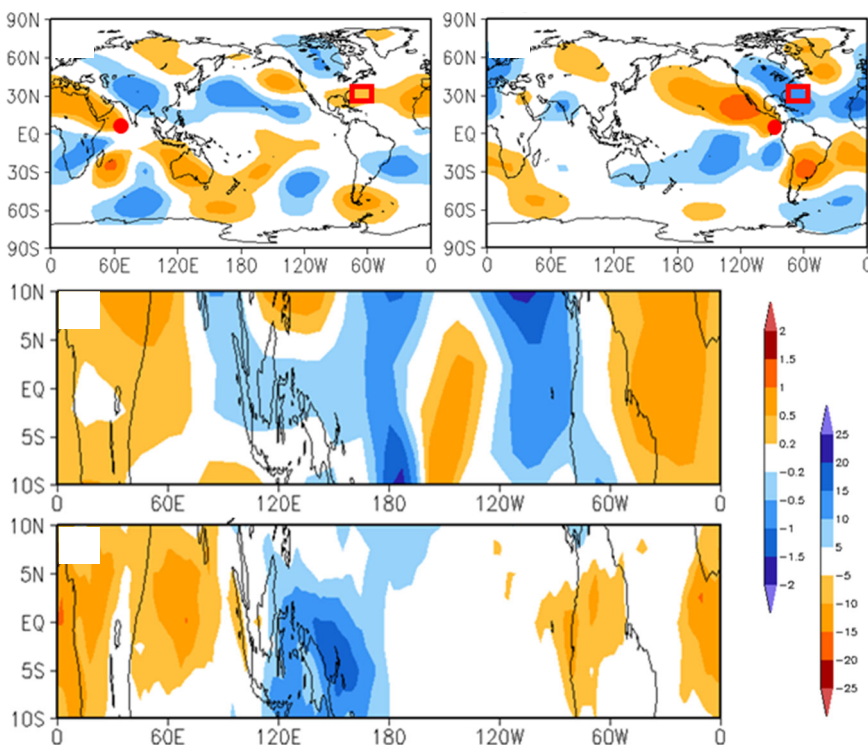


Fig. 6 (a, b) Responses of 200-hPa streamfunction to a point heating (red dot), (c) influence function for the red box in (a, b) and Fig. 5b derived from individual circulation responses to 1920 points of heating globally (R15 resolution: 48×40 grid points) and (d) composite of weekly OLR anomalies for MJO phase 1.

Fewer Troughs, Not More Ridges, Have Led to a Drying Trend in the Western United States

Wei Zhang,¹ Vittal Hari,² Simon S-Y Wang,^{1,3} Matthew D. LaPlante,^{1,4} Gregg Garfin,⁵
 Grace Affram,¹ and Rohini Kumar²

¹Department of Plants, Soils and Climate, Utah State University, Logan, UT

²UFZ-Helmholtz Centre for Environmental Research, Leipzig, 04318, Germany

³Utah Climate Center, Utah State University, Logan, UT

⁴Department of Journalism and Communication, Utah State University, Logan, UT

⁵School of Natural Resources and the Environment, University of Arizona, Tucson, AZ

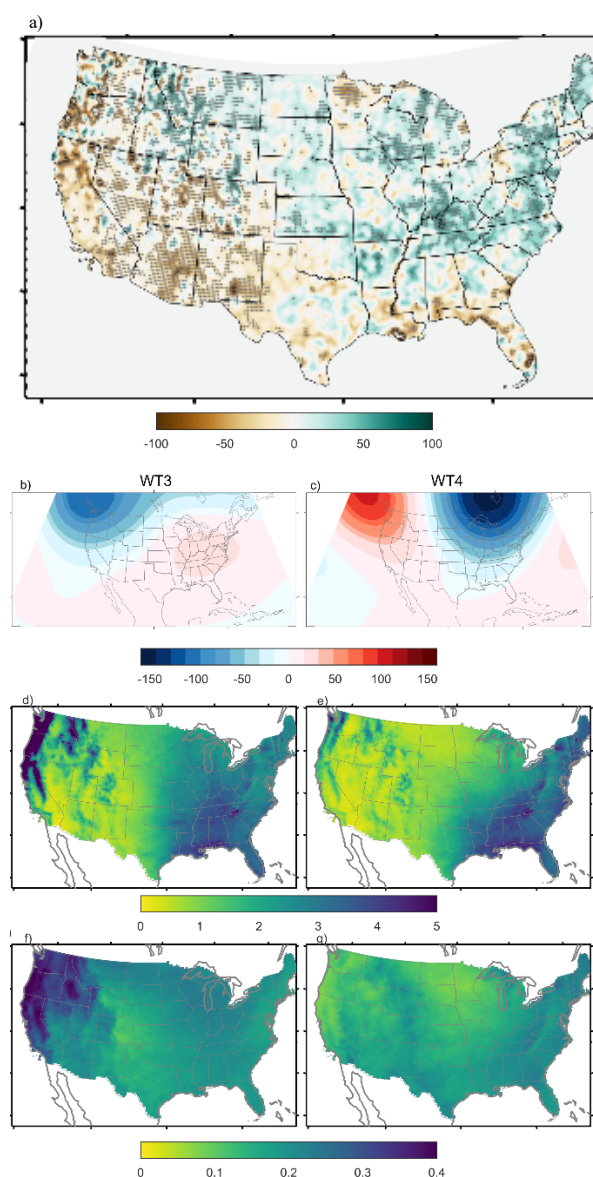
1. Introduction

High-amplitude ridges can enforce heat-trapping systems that persist through an entire season, contributing to drought events. While the strength of the seasonal-mean atmospheric ridge has increased, as previously documented, the impact of day-to-day weather system changes at the meso- and synoptic-scale on drought events are not yet clear. Here, we examine the day-to-day ridge and trough patterns from the perspective of weather types within a transient framework (Zhang *et al.* 2022).

2. Data and methodology

Daily precipitation data over the continental United States is obtained from the Climate Prediction Center at 0.25 spatial resolution. The 500-hPa geopotential height data for deriving the weather types/regimes are obtained from the National Aeronautics and Space Administration (NASA)'s Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2) at $0.5^\circ \times 0.625^\circ$ spatial resolution for 1980-present; the European Centre for Medium-Range Weather Forecasts (ECMWF)'s ERA-5 at $0.25^\circ \times 0.25^\circ$ spatial resolution (1979-present); the National Centers for Environmental

Fig. 1 a) trend in annual total precipitation (shading: unit: mm/decade) over the continental US during 1980-2018. The hatched regions are statistically significant at the 0.05 level. b) 500-hPa geopotential height anomaly (unit: gpm), d) composite precipitation (unit: mm/day) and f) fractional contribution of WT3 (the western trough pattern) precipitation to total precipitation. c) 500-hPa geopotential height anomaly, e) composite precipitation and g) precipitation fraction for WT4 (the western ridge pattern).



Prediction and National Center for Atmospheric Research (NCEP/NCAR) reanalysis data at $2.5^\circ \times 2.5^\circ$ spatial resolution (1948-present); and the Japanese 55-year ReAnalysis (JRA-55, 1958-present) at $1.25^\circ \times 1.25^\circ$ spatial resolution (Kobayashi *et al.*, 2015). The daily temperature data set is obtained from Berkeley Earth (<http://berkeleyearth.org/data/>). We use daily 500-hPa geopotential height data from the Coupled Model Intercomparison Project Phase 6 (CMIP6) project and CESM large ensemble run with 42 members. We use k-means cluster analysis to obtain weather regimes.

3. Results

Following previous studies that examine weather types (WT) (Zhang and Villarini 2019, 2021), we apply K-means clustering to daily 500-hPa geopotential height anomalies based on reanalysis data, resulting in five distinct WT over the continental United States. Each WT represents a unique, large-scale pattern that modulates precipitation and temperature on a day-to-day basis. Weather Type 3 (WT3) is associated with a deep trough in the western U.S. (Fig. 1b), while WT4 represents a strong ridge pattern similar to the amplified and semi-permanent ridge associated with severe drought conditions in California (Fig. 1c). Five WTs exhibit distinct characteristics of composite precipitation and fractional contribution in the western U.S. and this is also true across different seasons. WT3 exhibits higher precipitation and fractional precipitation contribution than WT4 (Fig. 1d-g), as well as colder temperatures, an effect that is particularly pronounced in the winter season, which is similar to the Central and Midwestern regions of the U.S. (Zhang and Villarini 2019, 2021).

Using Poisson regression, we detect a significant decreasing trend (Beta = -0.0147, p-value < 0.01) in the annual frequency of WT3 from 1980 to 2018 that is consistent across different reanalysis data sets. No such trend can be detected for WT4 (Fig. 2). The results are consistent at the seasonal scale, exhibiting significant decreasing trends for WT3 and insignificant trend for WT4 during March-May (MAM), June-August (JJA), September-November (SON) and December-February (DJF). Assessing the relative roles of the different

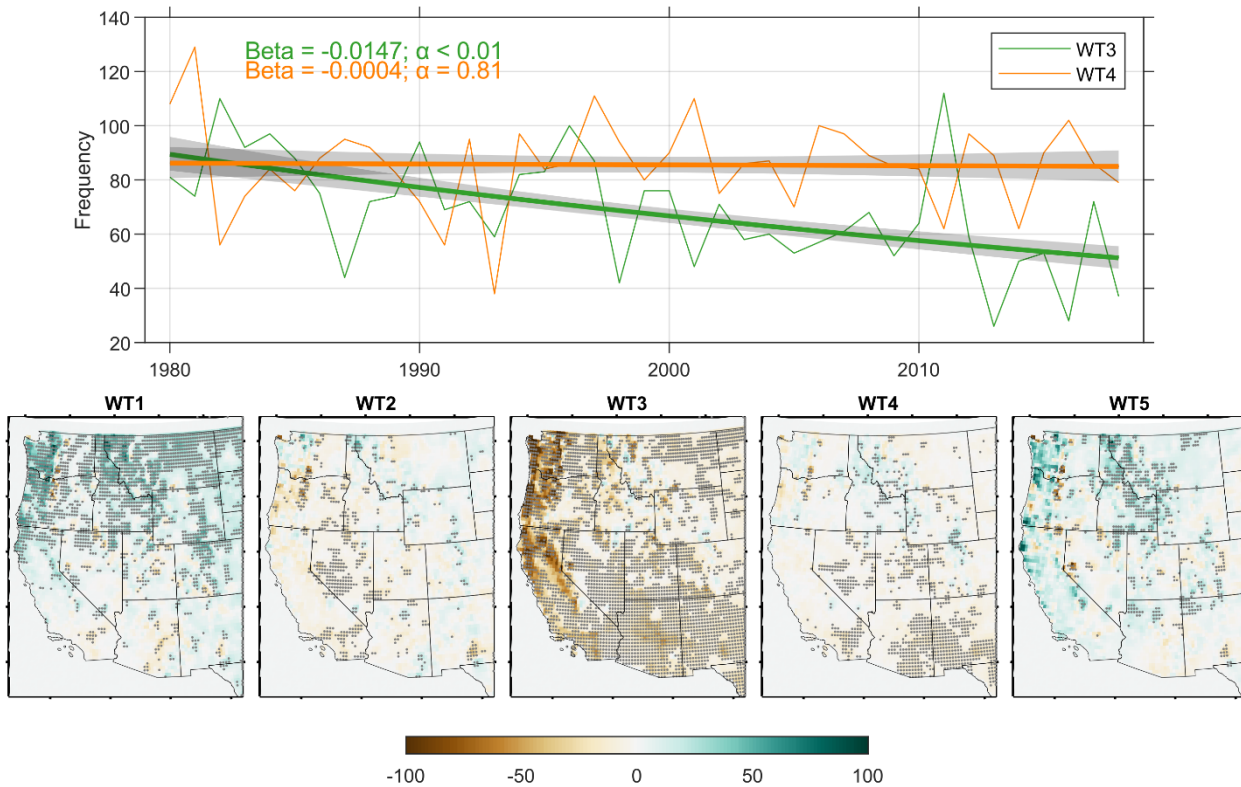


Fig. 2 Top: Frequencies (unit: events) and fitted trends (unit: events/yr) for weather types 3 and 4 during 1980-2018. Bottom: precipitation trend (unit: mm/decade) in the western US associated with WT1-5. The shading in the top panel represents the 95% confidence interval of the fitted trend using Poisson regression. The hatched regions are statistically significant at the 0.05 level.

weather types in shaping the change of annual total precipitation, we see the decrease in WT3-associated precipitation has a magnitude and pattern (Fig. 2), similar to that of the annual total precipitation declines from 1980–2018 (Fig. 1a) while the change in WT4 exhibits a much weaker and insignificant trend.

4. Concluding remarks

While an amplified seasonal-mean ridge may play a partial role in explaining recent California droughts and warming, our analysis indicates that it is the decreasing trend in transient western troughs that has contributed the most to the downward trend in the annual and seasonal precipitation totals across the western U.S. These findings highlight the importance of improved understanding of time-mean features (*i.e.* stationary waves over North America) and transient weather patterns (*e.g.*, cold fronts), their respective trends, and possible external forcings. Our study also points to the need to improve the underlying atmospheric and oceanic drivers of western U.S. drought to aid drought prediction.

References

- Kobayashi, S., Y. Ota, Y. Harada, A. Ebita, M. Moriya, H. Onoda, K. Onogi, H. Kamahori, C. Kobayashi, H. Endo, K. Miyaoka, and K. Takahashi, 2015: The JRA-55 Reanalysis: General specifications and basic characteristics. *J. Meteor. Soc. Japan*, **93**, 5–48, <https://doi.org/10.2151/jmsj.2015-001>.
- Zhang, W., and G. Villarini, 2019: On the weather types that shape the precipitation patterns across the U.S. Midwest. *Clim. Dyn.*, **53**, 4217–4232, doi:10.1007/s00382-019-04783-4.
- , —, 2021: Greenhouse gases drove the increasing trends in spring precipitation across the central USA. *Phil. Trans. Roy. Soc. A*, **379**, 20190553, doi:10.1098/rsta.2019.0553.
- , V. Hari, S. S.-Y. Wang, M. D. LaPlante, G. Garfin, G. Affram, and R. Kumar, 2022: Fewer troughs, not more ridges, have led to a drying trend in the western United States. *Geophys. Res. Lett.*, **49**, e2021GL097089, doi:10.1029/2021GL097089.

Predictability of Summer Monsoon Extreme Rainfall Events over Taiwan Using NCEP GEFSv12 Reforecast

M. M. Nageswararao,^{1,2} Yuejian Zhu,² and Vijay Tallapragada²

¹CPAESS, UCAR at NOAA/NWS/NCEP/EMC, College Park, MD

²Environmental Modeling Center, NOAA/NWS/NCEP, College Park, MD

1. Background and motivation

The skillful prediction of monthly scale rainfall at the regional/local level is one of the challenges of the meteorological scientific community. The present existing forecast systems across the world can simulate the year-to-year variation. However, the variability on month-to-month during a year is tricky due to the considerable ambiguity associated with aberrant internal low-frequency fluctuations. The monthly scale is the bridge between medium-range and seasonal/long-range weather forecasts. It is a difficult time range for weather forecasting because much of the memory of the initial atmospheric conditions on this time scale is lost, affecting the forecast prediction skill.

Taiwan is one of the sub-tropical islands in Asia. It experiences rainfall extremes regularly, leading to landslides and flash floods in/near the mountains and flooding over low-lying plains, particularly during the summer monsoon season (June through September; JJAS). Although Taiwan has a lot of water scarcity due to steep topography that hardly holds rainwater, it is the 18th rank of the water-scarcity countries in the world. Therefore, there is high demand for accurate prediction of monthly rainfall and associated extreme rainfall events over Taiwan.

In September 2020, NOAA NCEP implemented Global Ensemble Forecast System version 12 (GEFSv12) to support stakeholders for sub-seasonal forecasts and hydrological applications. The NOAA NCEP generated consistent GEFSv12 reforecast data for 2000-2019 and the reforecasts were initialized at 00 UTC once per day out to 16 days with 5 members ensemble, except on Wednesdays when the integrations were extended to 35 days with 11 members. In this study, the performance of GEFSv12 monthly scale rainfall and associated extreme events during JJAS over Taiwan for the period (2000-2019) has been evaluated.

2. Data and methodology

In this study, the NCEP GEFSv12 rainfall products over Taiwan based on every Wednesday 00 UTC initial conditions up to 35-day forecast lead times with 11 members during JJAS for the reforecast period 2000-2019 have been considered. These products are available at Amazon Web Services (AWS, <https://registry.opendata.aws/noaa-gefs/>) in grib2 format at 3 (6) hour intervals at 0.25° (0.5°) resolution for the first 10 days (beyond 10 days) of the forecast. For uniformity, day-1 to 10 forecasts are also considered on the same grid points of day-11 to 35 forecasts. The reforecast products are based on the current operational Global Forecast System version 15.1 (GFSv15.1). It uses the Geophysical Fluid Dynamics Laboratory (GFDL) FV3 Cubed-Sphere dynamical core. The GEFSv12 model horizontal resolution is ~25 km (C384 grid) with 64 vertical hybrid levels. The top layer is centered around 0.27 hPa (~55 km). The scale-aware parameterization convection scheme is used instead of Simplified Arakawa-Schubert (SAS) shallow and deep convection scheme. The scale-aware parameterization convection scheme was further modified to reduce excessive cloud-top cooling for the model stabilization. The hybrid Eddy-Diffusivity Mass-flux (EDMF) scheme is used for the vertical mixing process of the planetary boundary. The cloud microphysics scheme used from GFDL includes five predicted cloud species (cloud water, cloud ice, rain, snow, and graupel). The shortwave and longwave radiative fluxes are estimated from the rapid radiative transfer model (RRTM) developed at Atmospheric and Environmental Research (Clough *et al.* 2005). The convective gravity wave drag estimation uses the schemes

developed by Chun and Baik (1998). The GFS orographic gravity wave drag and mountain blocking schemes follow that developed by Alpert (1988). A two-tiered Sea Surface Temperature (SST) and Near Sea Surface Temperature (NSST) approach (Zhu *et al.* 2017, 2018) is used for estimating the SST boundary condition, which accounts for the day-to-day variability and diurnal variation of SST, respectively. The stochastic kinetic energy backscatter (SKEB) and stochastically perturbed parameterization tendencies (SPPTs) were used to improve the model's uncertainty. A brief description of the GEFSv12 forecast system at NOAA NCEP can be found in Zhou *et al.* (2019; 2021). CMORPH multi-satellite-based precipitation data for the same period (2000–2019) were acquired from the official FTP server of the Climate Prediction Center of the National Oceanic and Atmospheric Administration and used as a reference data (<https://www.ncei.noaa.gov/data/cmorph-high-resolution-global-precipitation-estimates/access/daily/0.25deg/>). The GEFS-SubX reforecast at TL574L64 (day 0–8; ~34 km horizontal resolution) and TL382L64 (day 8–35; ~52 km horizontal resolution) is considered a benchmark dataset to compare the ability of the GEFSv12 to predict the summer monsoon (JJAS) daily rainfall over Taiwan with different forecast lead times Day-1 to 35 based on every Wednesday initial conditions. More details about the GEFS-SubX system and the configurations can be found in Zhu *et al.* (2018).

The numerical weather models' raw products are not skillful at extended/monthly/sub-seasonal scale, and suitable statistical post-processing is highly required for skillful forecast guidance and increase of its usability. For further precipitation prediction skill improvement, in the previous studies, various ensemble-based statistical post-processing techniques have been used, *e.g.*, Frequency Match Method (FMM, Zhu and Luo 2015), Quantile-Quantile mapping Method (Nageswararao *et al.* 2021; Guan *et al.* 2021), "poor man's ensemble" (Ebert 2001), and analog method (Hamill and Whitaker 2006). This study uses the quantile-quantile mapping post-processing technique to calibrate GEFSv12 rainfall reforecast data to improve prediction skills. The main advantage of this calibration method is to transform rainfall simulated by GEFSv12 to bias-corrected data statistically and make it applicable for use in the impact assessment of the GEFSv12 model. The technique is also called 'histogram equalization and/or 'rank matching' (Piani *et al.* 2010). In this study, the statistics of daily rainfall for CMORPH and GEFSv12 reforecasts were determined independently for each lead time (Day-1 to 30 forecast lead times) and grid point over Taiwan during JJAS. This method is applied to an ensemble of 11-members and to each member separately. For example, in the June analysis, the GEFSv12 reforecast data is based on 16th May to 15th June 00 UTC weekly once (every Wednesday) initial conditions. The corresponding sample size is 89 at each grid point for each lead time and each member. The July, August, and September analyses have been practiced like June analysis for implementing the calibration method. The rainfall intensity distributions for both CMORPH and GEFSv12 reforecasts are well approximated by the gamma distribution. The empirical probability distributions of CMORPH and GEFSv12

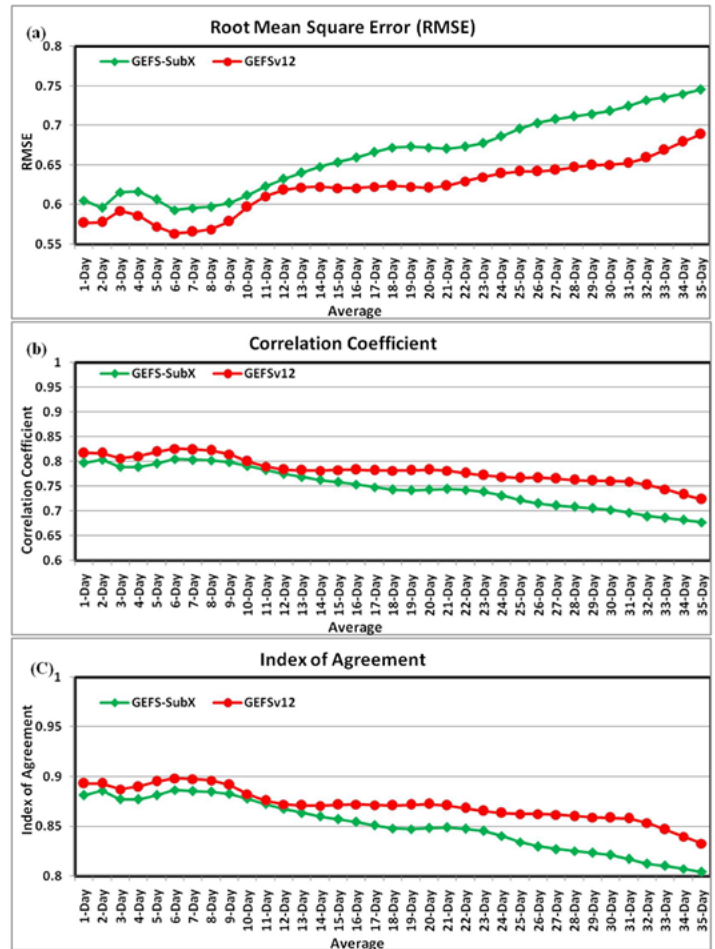


Fig. 1 (a) RMSE (b) Correlation Coefficient and (c) Index of Agreement of GEFS-SubX and GEFSv12 in depicting East-Asian-Summer-Monsoon-Index (EASMI) against 20th CR NCEP Reanalysis based on every Wednesday initial conditions for forecast lead time Day-1 to 35 for the period 2000–2019.

rainfall values have been used in this technique. The calibrated output is the inverse of the cumulative distribution function (CDF) of CMORPH values at the probability corresponding to the GEFSv12 model output CDF at the particular value (F_t). The bias is not calculated explicitly in this method. Suppose CDFs, F_{CMORPH} for CMORPH, and $F_{GEFSv12}$ for an ensemble member rainfall forecast of the GEFSv12 model. For F_t , the bias-corrected value Q will then be as follows:

$$Q = F_{CMORPH}^{-1}(F_{GEFSv12}(F_t))$$

Here, F^{-1} is an inverse of CDF. Thus, the technique of the quantile mapping is a transformation between two CDFs of the CMORPH and GEFSv12 model. The leave-one-out cross-validation procedure has been practiced in the entire process. Hereafter, the Raw and calibrated outputs of GEFSv12 are mentioned as Raw-GEFSv12 and QQ-GEFSv12, respectively.

3. Preliminary results

The East Asian Summer monsoon Index (EASMI) can capture the interannual and interdecadal variations in EASM-related climate anomalies and it is good at describing precipitation and air temperature variations over East Asia (Zhao *et al.* 2015). The EASMI is tightly associated with the East Asian–Pacific or the Pacific–Japan teleconnection and there is a possible role of internal dynamics in the EASM variability. It is also significantly linked to El Niño–Southern Oscillation (ENSO) and tropical Indian Ocean Sea surface temperature anomalies. There is a need to know the predictability of EASMI from GEFS-SubX and GEFSv12 to understand the predictability of EASM rainfall over Taiwan. In this study, the performance evaluation of GEFS-SubX and GEFSv12 Day-1 to 35 forecast lead times for EASMI against 20th CR NCEP (<https://data.tpc.ac.cn/en/download/1c0c4197-5e5d-4f0d-bd38-03dae3658a06/>) for the period 2000–2018 has been done by using standard skill metrics. Both models are good in capturing the EASMI with all forecast lead times against 20th CR NCEP reanalysis (Fig. 1). However, the RMSE of both the models in depicting EASMI is increasing with lead time (Fig. 1a). It is interesting to notice that the RMSE of GEFSv12 for EASMI is relatively lesser for all forecast lead times than the GEFS-SubX. The correlation coefficient (CC) and Index of agreement (IOA) of both models in representing the EASMI is significantly high for all lead time forecasts ($CC > 0.65$ and $IOA > 0.8$) and it is particularly more up to Day-23 forecast lead time ($CC > 0.75$ and $IOA > 0.85$). However, the CC and IOA of both models for EASMI decreases with lead time. The prediction skill of EASMI from GEFSv12 is relatively higher for all forecast lead times than the GEFS-SubX (Fig. 1b & c). The remarkable

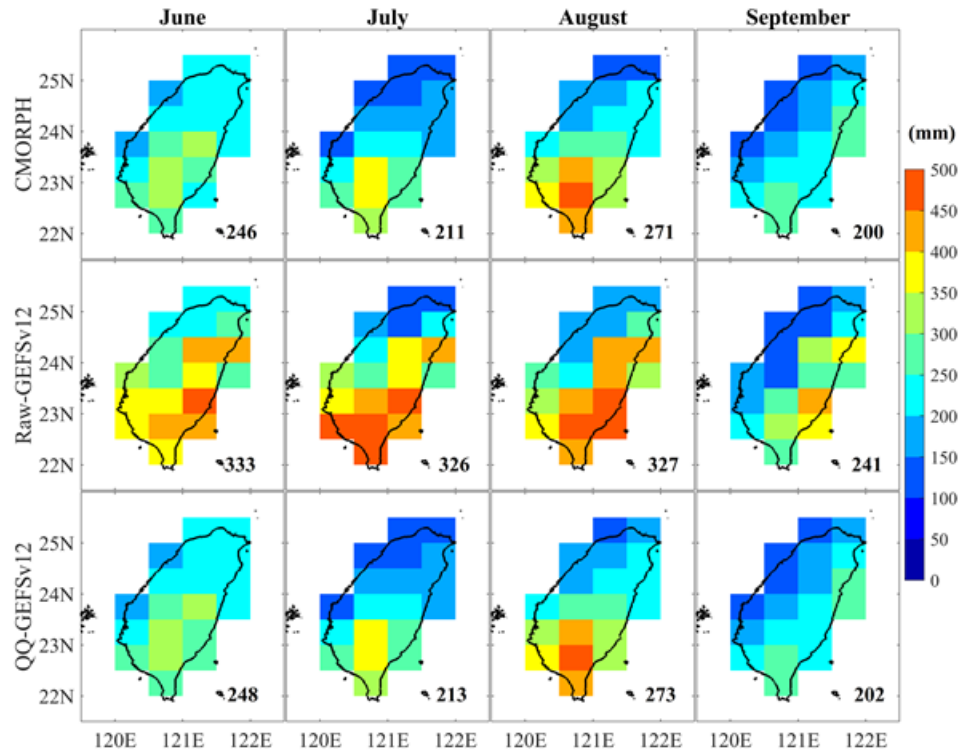


Fig. 2 Spatial distribution of monthly rainfall (mm) during summer monsoon from CMORPH, Raw, and QQ-GEFSv12 (Ensemble mean of 11 ensembles) based on weekly once initial conditions for the period 2000–2019. The value at the bottom right corner of each panel indicates the average climatological mean of monthly rainfall in Taiwan.

improvement in prediction skill of GEFSv12 in representing the ESM circulation dynamics and its influence on ESM rainfall over Taiwan as compared to GEFS-SubX. The improvements in prediction skill of GEFSv12 may be attributed to the combined influence of better initial conditions, more advanced microphysics schemes, updated stochastic schemes, finer resolution and a new FV3 dynamic core.

The monthly summer monsoon rainfall analysis over Taiwan reveals that the Raw-GEFSv12 can capture summer monsoon monthly rainfall patterns over Taiwan like CMORPH (Fig. 2). The monthly rainfall is more prominent over the western windward sloping areas in central and southern Taiwan (300-500 mm), decreasing north and west. During June, July, and August, the maximum rainfall is mainly due to the convective systems embedded within the southwesterly monsoon flow. In addition, the Mei-yu frontal systems from southern China also frequently bring heavy precipitation towards Taiwan. It is also noticed that the rainfall is quite significant along with the Central Mountain Range (CMR) as the orographic impact. During June, the pronounced monsoon rainfall covers most parts of Taiwan and reduces west to east as the season progresses. Among the months, the maximum observed monthly rainfall occurred during August (271 mm), followed by June (246 mm), July (211 mm), and September (200 mm). After August, the maximum rainfall zone shifted towards the north with season progress, and similar patterns can be seen from Raw-GEFSv12 (Fig. 2). However, the Raw-GEFSv12 has a large wet bias in most parts of Taiwan during all months, and it is notably more during July (115 mm) followed by June (87 mm), August (56 mm) and September (41 mm). After calibration, the monthly rainfall patterns over Taiwan are very similar to CMORPH, the wet bias is significantly reduced for all the months, and the magnitude of the monthly rainfall from QQ-GEFSv12 is relatively closer to CMORPH than Raw-GEFSv12.

The correlation analysis reveals that the Raw-GEFSv12 is good in capturing the year-to-year variations of the monthly summer monsoon rainfall over Taiwan for all the months against CMORPH (Fig. 3a). The Raw and QQ-GEFSv12 have significant correlation coefficient values in most parts of Taiwan ($CC > 0.4$ at 90% confidence level) in predicting monthly rainfall over Taiwan in most of the months except July. The correlation coefficient of Raw-GEFSv12 in depicting monthly rainfall is particularly larger for September (0.43) followed by August (0.41), June (0.39) and July (0.19). After calibration, a similar correlation coefficient pattern over Taiwan has been noticed from QQ-GEFSv12 for all the months. However, a slight decrease in the correlation coefficient has been seen. The reason is that the quantile-quantile mapping method is matching the forecast probability distributions to observations, but the temporal structure may be lost by using this method. It might not correct timing errors also. Further, the Index of Agreement (IOA) analysis indicates the Raw-GEFSv12 has well agreed with CMORPH in predicting monthly rainfall over Taiwan for all the months (Fig. 3b). The IOA values from Raw-GEFSv12 in most parts of Taiwan are

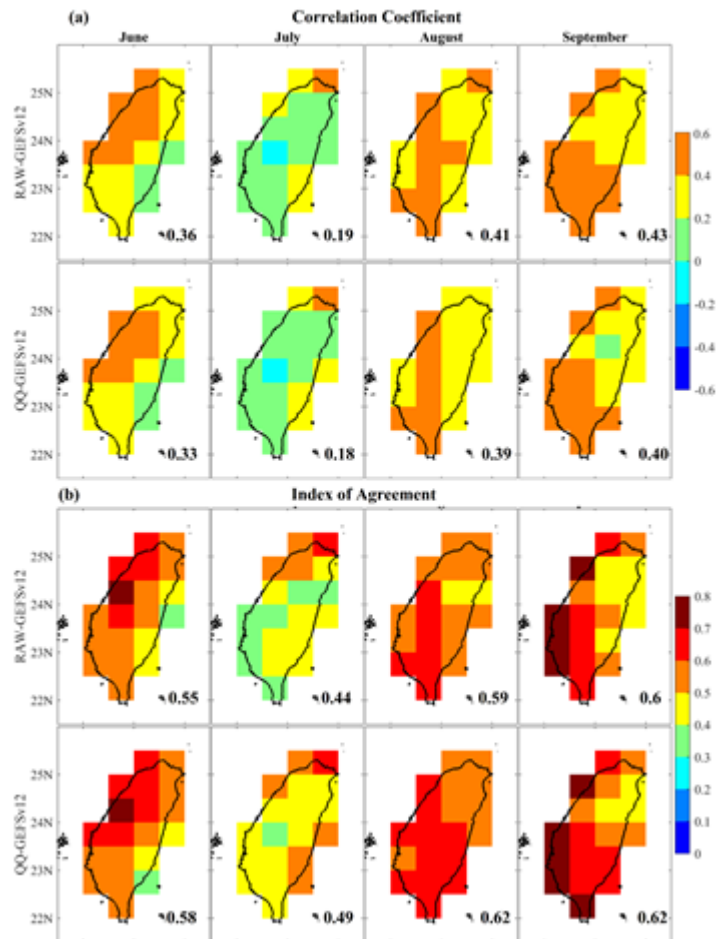


Fig. 3 (a) Correlation Coefficient and (b) Index of Agreement of Raw and QQ-GFSv12 against CMORPH in depicting Monthly rainfall during Summer Monsoon for the period 2000-2019. The value at the bottom right corner of each panel indicates the average CC/IOA of monthly rainfall in Taiwan.

remarkably high (IOA>0.5) for all the months, and it is higher for September (0.6) followed by August (0.59), June (0.55) and July (0.44). After calibration, there has been significant improvement in the IOA of QQ-GEFSv12 in most parts of Taiwan. The IOA values for Taiwan as a whole for June, July, August, and September are 0.58, 0.49, 0.62, and 0.62, respectively. The analysis concluded that the calibration method improves the prediction skill of GEFSv12 in depicting the monthly rainfall over Taiwan for all the monsoon months.

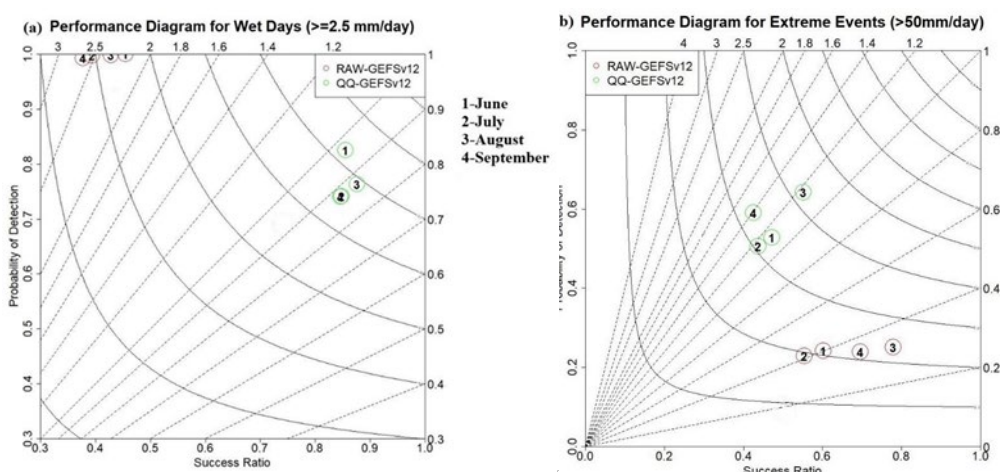


Fig. 4 Performance diagram summarizing the SR, POD, Frequency Bias, and TS statistical categorical skill scores of Raw, and QQ-GEFSv12 against CMORPH for (a) Wet days, and (b) Extreme rainfall events on monthly scale over Taiwan during June, July, August, and September for the period 2000-2018. The solid and dashed lines represent TS and Frequency Bias scores, respectively.

The performance diagram is a suitable method for summarizing the several categorical skill scores such as Probability of Detection (POD), Frequency Bias, Treat Score (TS), and Success Ratio (SR) (1-FAR) in a single graph (Huang and Luo 2017). The solid contour lines in the performance diagram (Fig. 4) show the TS, while the dash lines indicate the Frequency Bias with extended labels on X and Y's upper (2nd) axes. From Fig. 4a, the Raw-GEFSv12 has a considerable overestimation (Frequency Bias is >2) of wet days (>2.5 mm/day) over Taiwan for all months, whereas the POD is remarkably high (POD >0.9). After calibration, a substantial reduction of overestimated wet days over Taiwan from QQ-GEFSv12 has been detected for all months. In contrast, the POD of wet days over Taiwan decreased remarkably for all months. After calibration, the TS and SR skill scores for wet days over Taiwan from QQ-GEFSv12 (TS> 0.5 and SR >0.8) are extraordinarily higher than Raw-GEFSv12 (TS< 0.5 and SR <0.5) for all the months (Fig. 4a). The Raw-GEFSv12 has a considerable underestimation of extreme rainfall (ER) events over Taiwan for all the months (Frequency bias <0.4), and the POD in most of the months also is low (POD <0.3) (Fig. 4b). After calibration, the POD and TS of ER events over Taiwan from QQ-GEFSv12 increased remarkably for all the months. The frequency of ER events over Taiwan from QQ-GEFSv12 has risen notably for all the months. It is mainly due to adjusting the probability distribution of various intensity rainfall events from GEFSv12 to the CMORPH.

4. Conclusions

- There is a remarkable improvement in prediction skill of GEFSv12 in representing the East Asian summer monsoon circulation dynamics and its influence on summer monsoon rainfall over Taiwan as compared to GEFS-SubX, and these improvements may be attributed to the combined influence of better initial conditions, more advanced microphysics schemes, updated stochastic schemes, finer resolution and a new FV3 dynamic core.
- The GEFSv12 is good at representing the spatial patterns of monthly rainfall over Taiwan during the summer monsoon season. The monthly rainfall is more during August (271 mm), followed by June (246 mm), July (211 mm), and September (200 mm) respectively. After August, the maximum rainfall zone shifts towards the north with season progress and similar patterns are found from GEFSv12. However, GEFSv12 has a large overestimation of monthly rainfall in Taiwan and it is more for July (115 mm) followed by June (87 mm), August (56 mm) and September (41 mm).

- The RMSE and mean bias errors of monthly rainfall from Raw-GEFSv12 are high in the south and eastern part of Taiwan, whereas prominent monthly rain and its IAV are significantly high. The errors of RMSE and mean bias decrease from south to north and east to west during all the months. After calibration, both errors notably decreased in most parts of the country.
- The prediction skill (Correlation coefficient and Index of Agreement) of GEFSv12 in depicting the summer monsoon monthly rainfall over Taiwan for all the months is significantly high (CC and IOA >0.4) in most parts of Taiwan and particularly more during peak monsoon months August, June and also for September. It is interesting to notice that after calibration, the prediction skill remarkably increased (>0.5) from QQ-GEFSv12 for all months.
- A considerable overestimation of wet days (>2.5 mm/day) in most parts of Taiwan from Raw-GEFSv12 has been found during all months, whereas an underestimate of ER events is there. After calibration, the probability distribution of various intensity rainfall events from QQ-GEFSv12 is well adjusted to that from the CMORPH in most parts of the country during all the months. The QQ-GEFSv12 can depict ER events (> 50 mm/days), in which rainfall events lead to floods and landslides over Taiwan.
- The calibration method significantly improved the most statistical categorical skill scores of QQ-GEFSv12 for wet and ER events, while the POD is significantly improved for ER events in most parts of the country for all the months.

Acknowledgments. This work is carried out with generous funding support from the NCEP Visiting Scientist Program managed by the University Corporation for Atmospheric Research (UCAR) Cooperative Programs for the Advancement of Earth System Science (CPAESS). The authors are grateful to the Ensemble Team members at the NCEP Environmental Modeling Center (EMC) for providing access to the datasets used in this study.

References

- Alpert, J. C., M. Kanamitsu, P. M. Caplan, J. G. Sela, G. H. White, and E. Kalnay, 1988: Mountain induced gravity wave drag parameterization in the NMC medium-range forecast model. *Proc. Eighth Conf. on Numerical Weather Prediction*, Baltimore, MD, **8**, *Amer. Meteor. Soc.*, 726-733.
- Chun, H. Y., and J. J. Baik, 1998: Momentum flux by thermally induced internal gravity waves and its approximation for large-scale models. *J. Atmos. Sci.*, **55**, 3299-3310, doi:10.1175/1520-0469(1998)055<3299:MFBTII>2.0.CO;2.
- Clough, S. A., M. W. Shephard, E. J. Mlawer, J. S. Delamere, M. J. Iacono, K. Cady-Pereira, S. Boukabara, and P. D. Brown, 2005: Atmospheric radiative transfer modeling: A summary of the AER codes. *J. Quant. Spectrosc. Radiat. Transfer*, **91**, 233-244, doi:10.1016/j.jqsrt.2004.05.058.
- Ebert, E. E., 2001: Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Mon. Wea. Rev.*, **129**, 2461-2480, doi:10.1175/1520-0493(2001)129%3C2461:AOAPMS%3E2.0.CO;2.
- Guan Hong, and Coauthors, 2021: GEFSv12 reforecast dataset for supporting sub-seasonal and hydrometeorological applications. *Wea. Forecasting* (in review).
- Hamill, T. M., and J. S. Whitaker, 2006: Probabilistic quantitative precipitation forecasts based on reforecast analogs: Theory and application. *Mon. Wea. Rev.*, **134**, 3209-3229, doi:10.1175/MWR3237.1.
- Huang, L., and Y. L. Luo, 2017: Evaluation of quantitative precipitation forecasts by TIGGE ensembles for South China during the pre-summer rainy season. *J. Geophys. Res. Atmos.*, **122**, 8494-8516, doi:10.1002/2017JD026512.
- Nageswararao, M. M., Yuejian Zhu, Vijay Tallapragada, 2021: Prediction skill of GEFSv12 for southwest summer monsoon rainfall and associated extreme rainfall events on extended range scale over India. *Wea. Forecasting* (submitted).
- Piani, C., J. Haerter, and E. Coppola, 2010: Statistical bias correction for daily precipitation in regional climate models over Europe. *Theor. Appl. Climatol.*, **99**, 187-192, doi:10.1007/s00704-009-0134-9.
- Zhao, G. J., G. Huang, R. G. Wu, W. C. Tao, H. N. Gong, X. Qu, and K.M. Hu, 2015: A new upper-level circulation index for the East Asian summer monsoon variability. *J. Climate*, **28**, 9977-9996, doi:10.1175/JCLI-D-15-0272.1.

- Zhou, X., Y. Zhu, B. Fu, D. Hou, J. Peng, Y. Luo, and W. Li, 2019: The development of next NCEP Global Ensemble Forecast System. *Extended Summary, Climate Prediction S&T Digest, 43rd NOAA Climate Diagnostics and Prediction Workshop*, Santa Barbara, CA, DOC/NOAA, 159-163, doi:10.25923/ae2c-v522.
- Zhou, X., Y. Zhu, D. Hou, B. Fu, W. Li, H. Guan, E. Sinsky, W. Kolczynski, X. Xue, Y. Luo, J. Peng, B. Yang, V. Tallapragada, and P. Pegion, 2021: The introduction of the NCEP Global Ensemble Forecast System version 12. *Wea. Forecasting* (in review).
- Zhu, Y., and Y. Luo, 2015: Precipitation calibration based on Frequency Matching Method (FMM). *Wea. Forecasting*, **30**, 1109-1124, doi:10.1175/WAF-D-13-00049.1.
- Zhu, Y., X. Zhou, M. Pena, W. Li, C. Melhauser, and D. Hou, 2017: Impact of sea surface temperature forcing on weeks 3 and 4 forecast skill in the NCEP Global Ensemble Forecast System. *Wea. Forecasting*, **32**, 2159-2173, doi:10.1175/WAF-D-17-0093.1.
- Zhu, Y., X. Zhou, Wei Li, H. Dingchen, M. Christopher, Eric Sinsky, P. Malaquias, Bing Fu, G. Hong, K. Walter, W. Richard, and V. Tallapragada, 2018: Toward the improvement of sub-seasonal prediction in the NCEP Global Ensemble Forecast System (GEFS). *J. Geophys. Res.: Atmos.*, **123**, 6732-6745, doi:10.1029/2018JD028506.

Sea-Surface Temperatures and Vertical Wind Shear as Precursors to Tropical Cyclone Activity in the Caribbean and an Expanding Main Developing Region

Keneshia Hibbert,^{1,2} Thomas M. Smith,³ Equisha Glenn,^{1,4} and Jorge E. Gonzalez^{1,5}

¹NOAA Center for Earth System Science and Remote Sensing Technologies

²Earth & Atmospheric Science Department, The City College of New York, New York, NY

³NOAA/STAR/SCSB and CISESS/ESSIC, University of Maryland, College Park MD

⁴Civil Engineering Department, The City College of New York, New York, NY

⁵Department of Mechanical Engineering, The City College of New York, NY

Sea surface temperatures (SSTs) and vertical wind shear (VWS) are essential to tropical cyclone (TC) formation. TCs need warm SSTs and low shear for genesis. Increasing SSTs and VWS changes influence storm development. This work analyzes SST and VWS trends for the Caribbean, the surrounding region, and the Atlantic hurricane main developing region (MDR) from 1982 to 2020. Tropical storm intensity increases significantly during this period. Annual and seasonal trends show regional SSTs in the MDR are warming annually ($0.0219^{\circ}\text{C yr}^{-1}$) and per season ($0.0280^{\circ}\text{C yr}^{-1}$). Simultaneously, VWS decreases during the late rainfall season (LRS) at 0.0556m/s yr^{-1} in the MDR and 0.0167m/s yr^{-1} in the Caribbean and surrounding area, while the Atlantic Warm Pool (AWP) is expanding at 0.51km^2 per decade. Increased upper atmospheric winds are driving VWS changes. Correlations of large-area averages do not show significant relationships between TC intensity and frequency and SSTs or VWS during the LRS. The observed changes appear to be associated with regional warming SSTs impacting TC changes.

1. Introduction

Sea surface temperatures (SSTs) in the mid-Atlantic and the Caribbean have steadily increased for several decades (Antuna *et al.* 2015; Glenn *et al.* 2015 and 2021). Mid Atlantic SSTs are warming faster during the late rainfall season (LRS – August to November) than any other season (dry – December to March) and early – April to July) with temperatures reaching 26.5°C and higher, the threshold considered for deep convection (Fig.1) (Gadgil *et al.* 1984; Graham and Barnett 1987). Upward trends in the Atlantic SSTs may lead to tropical cyclonic (TC) activity intensification, posing more significant risks to coastal communities. Research shows that storms in the North Atlantic have increased in frequency and intensity since the 1980s (Melillo *et al.*, 2014). Additionally, the Caribbean and the U.S. have experienced increasingly intense hurricanes in recent years, potentially linked to increasing SSTs (IPCC, 2014). More than 80% of Atlantic TC systems that turn into major hurricanes are formed within the main developing region (MDR, 10°N - 20°N , 20°W - 80°W), where TCs primarily form from easterly waves originating in Western Africa (Goldenberg and Shapiro, 1996). Thus, understanding the connection between a changing climate and TCs frequency and intensification trends in this region is particularly important.

Previous studies have analyzed the role of various environmental factors in major hurricane activity in the MDR. Mann and Emanuel (2006) showed growing SST trends strongly correlated to TC counts in the MDR, unrelated to the Atlantic Multidecadal Oscillation (AMO). While Vecchi and Soden's 2007 work spoke about how wind shear is supposed to increase due to climate change, making it more difficult for hurricanes to form. Although previous studies of the correlations between SSTs and TC frequency and SST and deep convection have been conducted elsewhere, we have not found research on these topics within the MDR. Thus, the

objective of this study is to investigate the connections between recent SSTs (1982-2020), vertical wind shear (VWS) trends, and the possible linkages to recent observations in TC activity in the Caribbean, the surrounding region, and the MDR.

2. Materials and methods

Climatological analyses of the study region were analyzed over 71 years. The first SST dataset used in this study is the NOAA's Optimum Interpolated Sea-Surface Temperature (OISST). This is a 0.25° High-Resolution Optimum Interpolation (OI) Sea Surface Temperature v2.1 dataset with a daily temporal resolution combined with the infrared satellite SST estimates in-situ observation data from buoys and ships (Huang *et al.*, 2021). For the earlier period Pathfinder, infrared SST is used, and operational satellite SST estimates are used for daily updates. The second SST dataset used in this study is the ERSST v5, derived from the International Comprehensive Ocean-Atmosphere Dataset (ICOADS). It is a monthly $2^\circ \times 2^\circ$ horizontally gridded dataset that uses the NOAA Global Surface Temperature product to integrate ERSST data with land surface temperature from the Global Historical Climatology Network Monthly dataset combined surface temperature analyses (Huang *et al.*, 2021). The NCEP/NCAR Reanalysis 1 dataset provided vertical wind shear data with monthly temporal and $2.5^\circ \times 2.5^\circ$ spatial resolution. Atlantic hurricane data used for this project are taken from the NOAA HURDAT2 Reanalysis project, last updated in November 2019.

3. Results and discussions

There have been warming SSTs in the Caribbean, the surrounding region, and the MDR over the past 39 years (1982-2020), with the most significant changes occurring over the past 25-27 years. These changes have implications for increased hurricane activity and intensity. SSTs are warmer at the end of the season, October

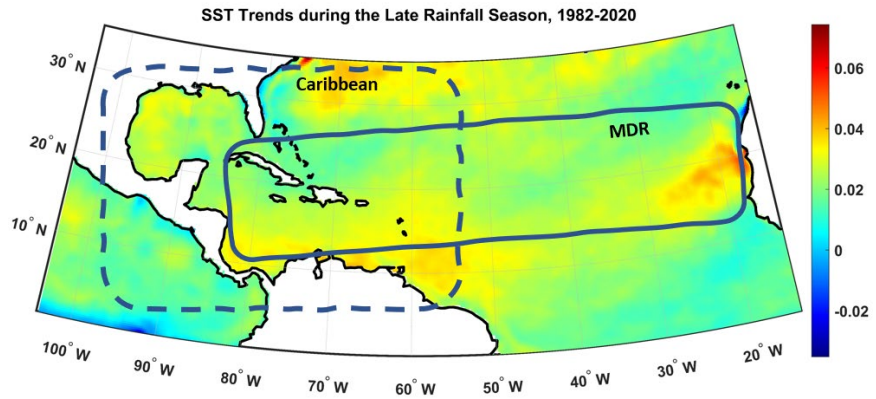


Fig. 1 Spatial Map depicting trends from 1982 to 2020 during the Late Rainfall Season (LRS), indicating warming trends up to and exceeding 0.4°C per decade in the Main Developing Region (MDR) (10°N - 20°N , 20°W - 80°W), and the Caribbean and the Surrounding Region (5°N - 31°N , 100°W - 55°W).

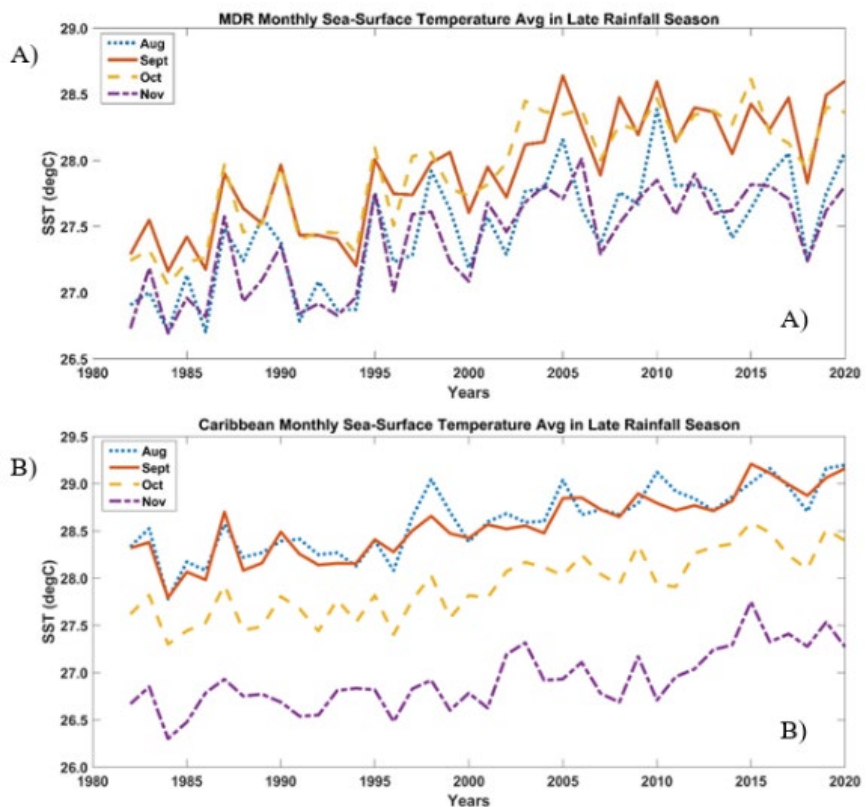


Fig. 2 Monthly regional SST trends from 1982 to 2020 using NOAA OISST product: (a) SST temperatures during the LRS in the MDR, (b) SST in the LRS during the Caribbean and the surrounding region.

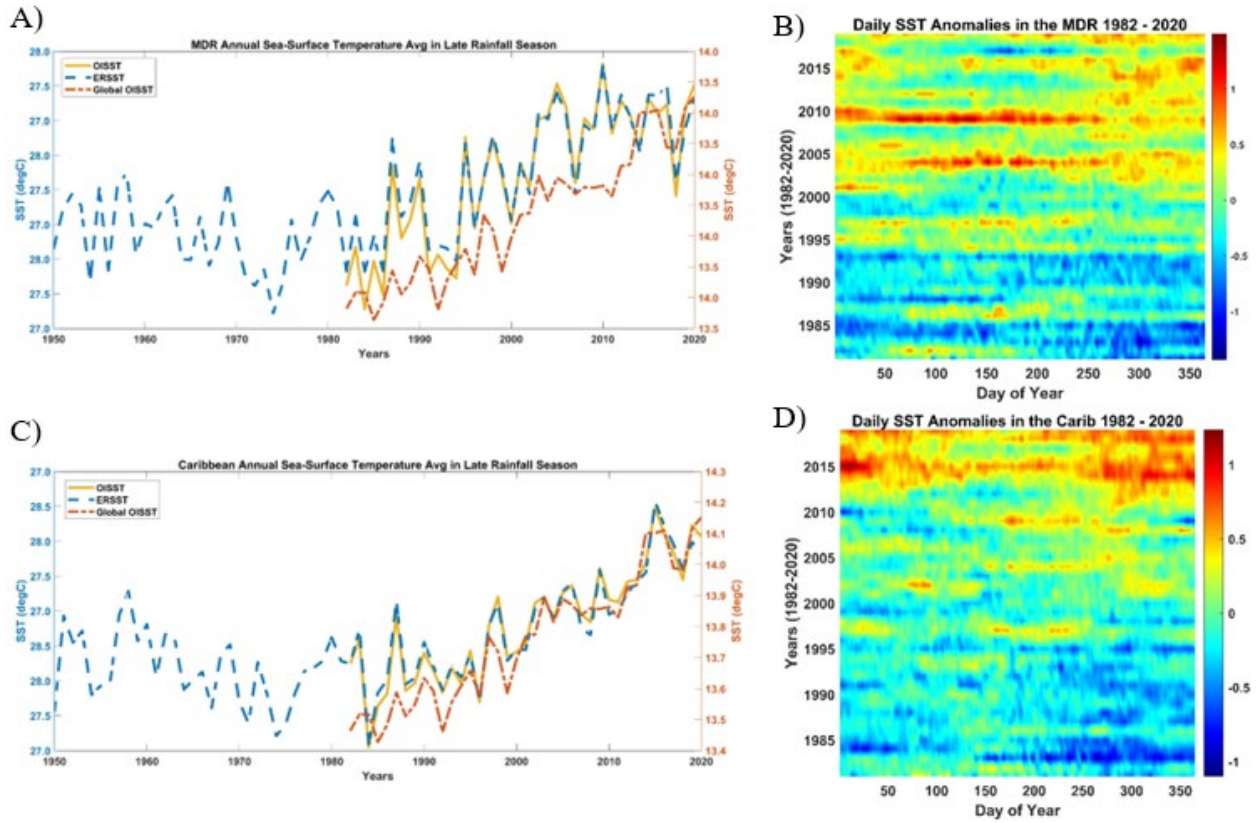


Fig. 3 Annual SST trends and daily anomalies. (a,c) annual average SSTs during the LRS for the MDR and the Caribbean and the surrounding region, depicting current and global SST trends from 1982-2020 using the NOAA OISST and 1950-2020 ERSST dataset. (b,d) regional daily SST anomalies during the LRS for the MDR and the Caribbean and the surrounding region. The color bar is SST in degrees Celsius.

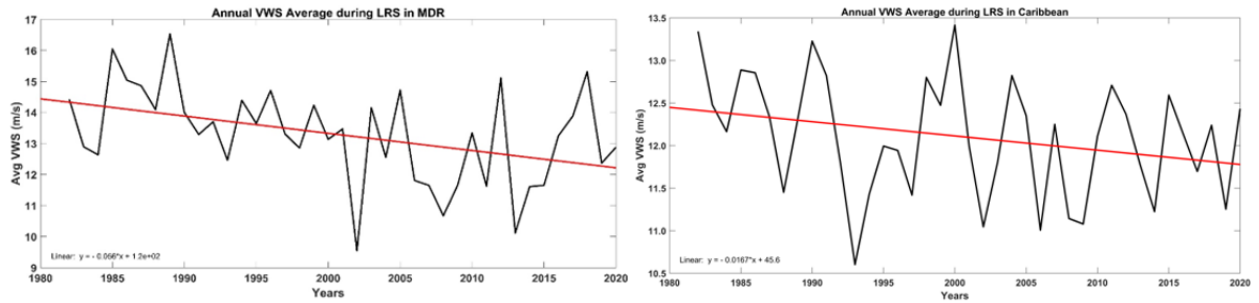


Fig. 4 Left: Annual average of VWS during the LRS from 1982 to 2020 in the MDR, showing a decreasing trend at 0.56ms^{-1} per decade; Right: Annual average of VWS during the LRS from 1982 to 2020 in the Caribbean and the surrounding region, showing a decreasing trend at 0.16ms^{-1} per decade.

SSTs are warming over time (Figs. 2 and 3). The AWP increases in magnitude and intensity during the LRS, most prominently in the latter half of the study period at 0.51km^2 per decade. This is important because of the oceanic-atmospheric interactions caused by changes to the AWP, where Atlantic TCs typically form and grow.

VWS results show that annual averages are stable. There are slight decreasing trends (Fig. 4), but the subtle change does not seem to be a significant cause for changes in Atlantic TC activity. Analysis of long-term yearly and seasonal trends for SSTs and VWS suggests that warming SSTs is the primary cause of increased hurricane frequency and intensity. Hurricane intensity is highly correlated with sea-surface temperature, implying that future warming will lead to more frequent or intense storms, increasing the chances of TCs making landfall with increased destructive potential.

While August and September are peak hurricane season months, increasing frequency and intensity in October suggest we may need to consider October as a part of the peak hurricane season (Fig.5). Despite recently reported higher shears, the warmer SSTs could be a factor in the increasing number of October. Additionally, with the expansion of the AWP, SST warming trends in the Caribbean, the surrounding region, and the MDR and increasingly frequent and intense storms in the area, we may need to consider expanding the definition of the MDR.

Future research is needed and will continue to analyze the impact of additional potential variables that influence TC activity, specifically ocean heat content, sea-level height, and mixed-layer depth.

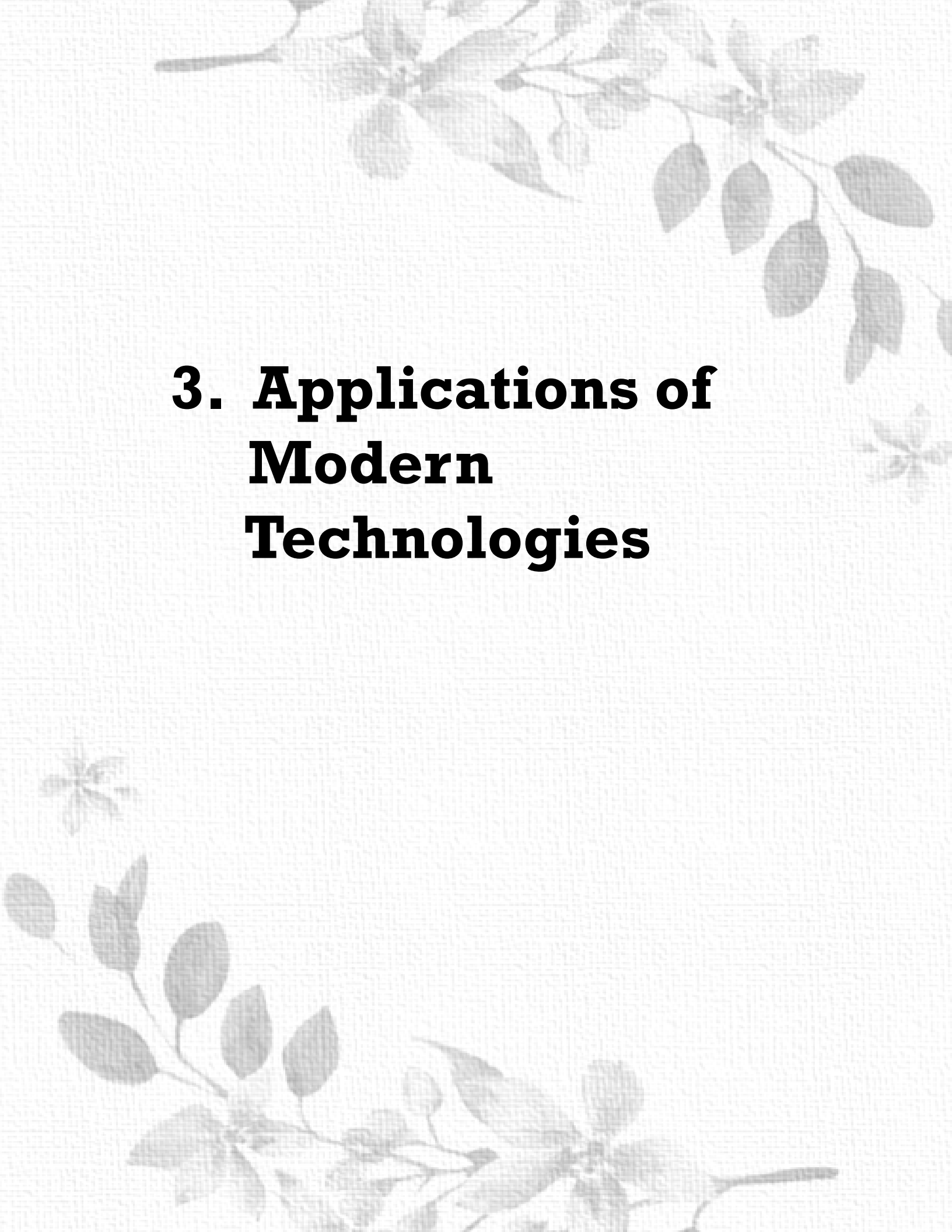


Fig. 5 Categories 1&2 Atlantic Hurricane (HU) and Categories 3-5 Atlantic major hurricane (MH) counts over the LRS months, 1851-2020. The blue bars (HU) indicate low-intensity storm counts, and the orange bars (MH) indicate high-intensity storm counts. Tables show the overall storm counts vs. average storm occurrence per year for that study period.

References

Antuña-Marrero, J. C., O. H. Otterå, A. Robock, and M. D. S. Mesquita, 2015: Modelled and observed sea surface temperature trends for the Caribbean and Antilles. *Int. J. Climatol.*, doi:10.1002/joc.4466.

- Gadgil, S., and Coauthors, 1984: Ocean-atmosphere coupling over monsoon regions. *Nature*, **312**, 141–143, doi:10.1038/312141a0.
- Glenn, E., and Coauthors, 2021: Tropical convection in the Caribbean and surrounding region during a regional, warming sea-surface temperature period, 1982–2020. *Hydrology*, **8**, 56, <https://doi.org/10.3390/hydrology8020056>.
- , D. Comarazamy, J. E. González, and T. Smith, 2015: Detection of recent regional sea surface temperature warming in the Caribbean and surrounding region, *Geophys. Res. Lett.*, **42**, doi:10.1002/2015GL065002.
- Graham, N. E., and T. P. Barnett, 1987: Sea surface temperature, surface wind divergence, and convection over tropical oceans. *Science*, **238**, 657–659, doi:10.1126/science.238.4827.657.
- Huang, B., C. Liu, V. Banzon, E. Freeman, G. Graham, B. Hankins, T. Smith, and H.-M. Zhang, 2021: Improvements of the Daily Optimum Interpolation Sea Surface Temperature (DOISST) version 2.1. *J. Climate*, **34**, 2923–2939, doi:10.1175/JCLI-D-20-0166.1.
- IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (Eds.)]. IPCC, Geneva, Switzerland, 151 pp.
- Mann, M. E., and K. A. Emanuel, 2006: Atlantic hurricane trends linked to climate change. *Eos, Trans. Amer. Geophys. Union*, **87**, 233, 238, 241, doi:10.1029/2006eo240001.
- Melillo, J. M., T. C. Richmond, and G. W. Yohe, Eds., 2014: Climate change impacts in the United States: The third national climate assessment. *U.S. Global Change Research Program*, <https://doi.org/10.7930/j0z31wj2>.
- Vecchi, G. A., and B. J. Soden, 2007: Increased tropical Atlantic wind shear in model projections of global warming. *Geophys. Res. Lett.*, **34**, L08702, doi:10.1029/2006GL028905.



3. Applications of Modern Technologies

Deep Learning for Subseasonal Precipitation and Temperature Errors

Maria J. Molina, Jadwiga Richter, Judith Berner, Anne A. Glanville, Katie Dagon,
 Abby Jaye, Aixue Hu, and Gerald Meehl

National Center for Atmospheric Research, Boulder, CO

1. Introduction

Every year, extreme precipitation and drought disrupt life, destroy infrastructure, and result in fatalities across the United States and the world. Skillful precipitation forecasts with a lead time of several weeks (*i.e.*, subseasonal) can help stakeholders of societally relevant public sectors (*e.g.*, water management, agriculture, and health) understand imminent threats and take protective actions to mitigate harm (Vitart *et al.* 2017, Pigeon *et al.* 2019). However, prediction skill of subseasonal precipitation from Earth system models remains poor (Fig. 1). Subseasonal prediction remains particularly challenging because the sources of predictability at such timescales are limited. Predictability stemming from atmospheric initial conditions is substantially reduced beyond approximately two weeks and the ocean generally does not offer added predictability until a trajectory reaches the seasonal timescale (Meehl *et al.* 2021).

Recent studies have shown that deep learning models can produce subseasonal to multiyear forecasts with skill that exceeds current dynamical forecasting systems (Ham *et al.* 2019, He *et al.* 2020, Kim *et al.* 2021, Weyn *et al.* 2021). Deep learning is well suited for geoscience prediction problems because of its ability to extract patterns from large amounts of data and its ability to learn multivariate relationships. Our objective is to leverage deep learning approaches with observational and reanalysis products to improve already existing subseasonal reforecasts (Richter *et al.* 2022) created using the Community Earth System Model v2 (CESM2; Danabasoglu *et al.* 2020). Subseasonal forecasts that are further bias-corrected and skillful could be of significant societal value for agricultural productivity, water management, and transportation systems.

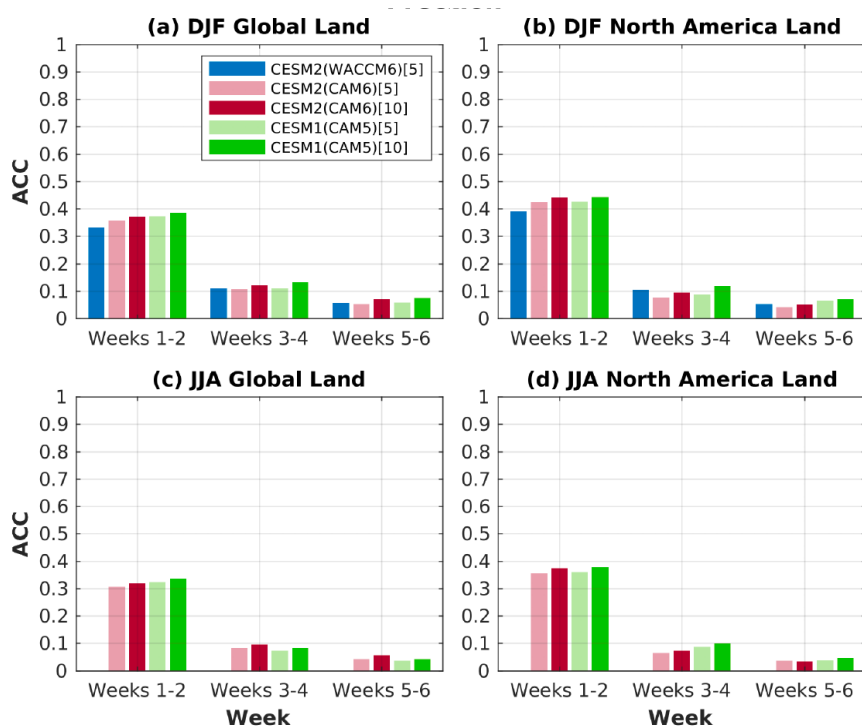


Fig. 1 Adapted from Richter *et al.* (2022, under review). DJF and JJA anomaly correlation coefficient (ACC) for CESM2 precipitation averaged over land (1999-2015).

2. Data and methodology/experimental design

The 11-member ensemble CESM2 subseasonal reforecasts (Richter *et al.* 2022) were initialized weekly every Monday (1999-2019) and carried out following SubX protocol (Pegion *et al.* 2019). Weekly real time forecasts created using CESM2 (Richter *et al.* 2022) are being contributed to the multi-model mean ensemble

used to issue the experimental NOAA weeks 3-4 outlooks as part of the SubX experiment. We focused on bias correction of global subseasonal temperature and precipitation forecasts produced using CESM2. Observational and reanalysis products used include the NOAA Climate Prediction Center (CPC) global daily gridded surface air temperature and global unified gauge-based analysis of daily precipitation, the NOAA Global Precipitation Climatology Project data (GPCP; Adler *et al.* 2003), and the ECMWF Reanalysis 5th Generation data (ERA5; Hersbach *et al.* 2020).

Preliminary tests conducted involved the training of numerous deep learning model architectures, including densely connected neural networks, long short-term memory networks, and convolutional neural networks. The deep learning model architecture, named U-Net (Ronneberger *et al.* 2015), resulted in comparatively higher skill when predicting subseasonal precipitation prediction errors. The U-Net involves the use of numerous convolutional and max pooling layers to downsample and upsample the input data as it is propagated through the network's layers and weights (Fig. 2). The U-Net also contains numerous cross connections across network layers, which helps reduce vanishing gradient issues during training, where the measured loss may not be propagated fully throughout the network's deep layers. Evaluation of model skill in bias correcting precipitation forecasts was assessed using anomaly correlation coefficient, root mean squared error, and ranked probability skill score.

The open-source Python-language software used to develop and train the deep learning models is PyTorch (<https://pytorch.org/>) and ongoing work is contained in an accessible Github repository (<https://github.com/mariajmolina/ML-for-S2S>). PyTorch can be run on both CPU and GPU resources, but preference is for running on GPU resources given the substantial speed-up in training. The PyTorch Application Programming Interface (API) contains the capability to scale deep learning model training across numerous GPUs, based on the number of GPUs allocated to a job when launched.

The numerical approach of this project involves the use of already existing CESM2 subseasonal reforecasts (11-ensemble members; 1999-2019) and observational products. The data products were preprocessed into lead time bias corrected anomalies for deep learning model training and were made available in NetCDF format. The files were used to train the U-Net to predict CESM2 subseasonal reforecast errors for temperature and precipitation (weeks 3, 4, 5, and 6), which were then used to bias correct CESM2 subseasonal forecasts of precipitation and temperature. CESM2 subseasonal forecasts and deep learning bias corrected fields were output using SubX protocol (Pegion *et al.* 2019) on a nominal 1-degree global grid and evaluated with daily temporal resolution. Years 1999 through 2015 were used for training and validation, and the years 2016 through 2019 were used for final evaluation (*i.e.*, test set).

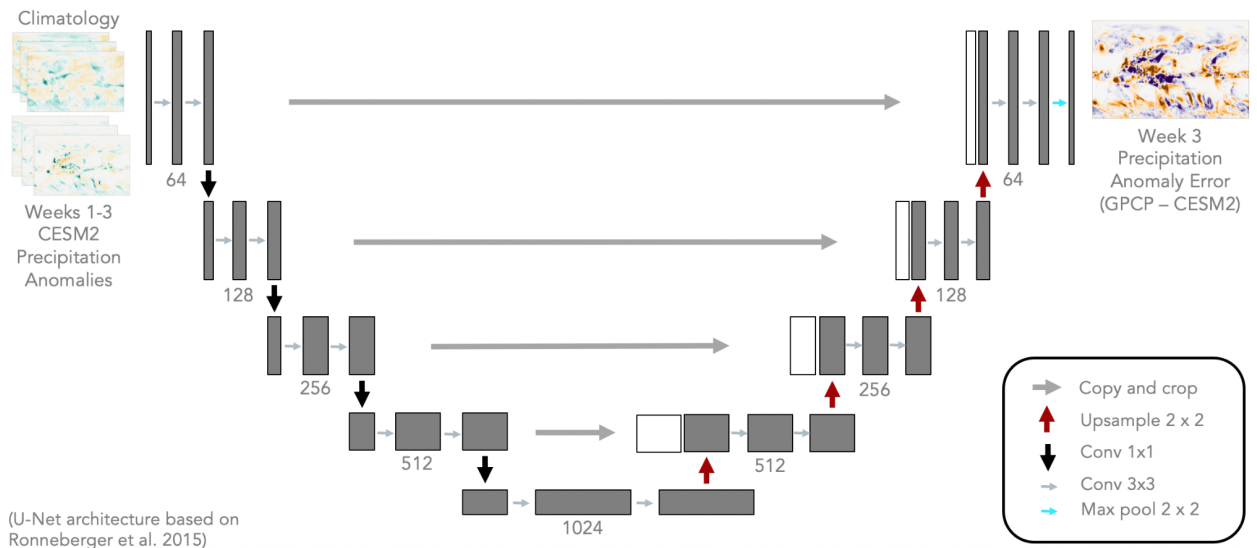


Fig. 2 U-Net architecture trained to learn CESM2 subseasonal precipitation prediction errors.

3. Results

Using the U-Net architecture above (Fig. 2), tests were performed using various fields as input (*i.e.*, predictors), including CESM2 precipitation anomaly subseasonal predictions (varying the number of preceding weeks as input) and the respective weekly climatology. The U-Net was trained to output (*i.e.*, predictand) the errors associated with the CESM2 predicted precipitation and temperature global anomalies, with the error computed relative to an observational field that contained data over land and ocean points (e.g., NOAA GPCP for precipitation and ERA5 for 2-meter temperature). Training was conducted using all the individual CESM2 ensemble members (not the ensemble mean) to take advantage of the larger sample size that such an approach provides. This project is ongoing, but here we provide some very preliminary results.

We found that the U-Net was able to predict error globally with skill across numerous lead times (weeks 3-6). For week 3 (years 2016 through 2019), the anomaly correlation coefficient (ACC) for prediction of temperature anomaly errors over land were approximately 0.4 (cosine/area weighted), whether stratifying by season (DJF, MAM, JJA, and SON) or considering annual skill. Similarly, for week 3 prediction of precipitation anomaly errors over land, ACC skill was approximately 0.3, whether stratifying by season (DJF (Fig. 3), MAM, JJA, and SON) or considering annual skill. While the U-Net was able to predict anomaly errors with skill across specific localized areas, adding the U-Net predicted errors back to the original CESM2 prediction (*i.e.*, bias correction) did not result in substantial improvement in precipitation or temperature prediction skill. We hypothesize that multiple U-Net models should be trained and used to bias correct global subseasonal prediction errors (for temperature and precipitation separately), as such an approach would enable the U-Net to learn finer-scale regional patterns that could provide more skill, rather than the current global approach used.

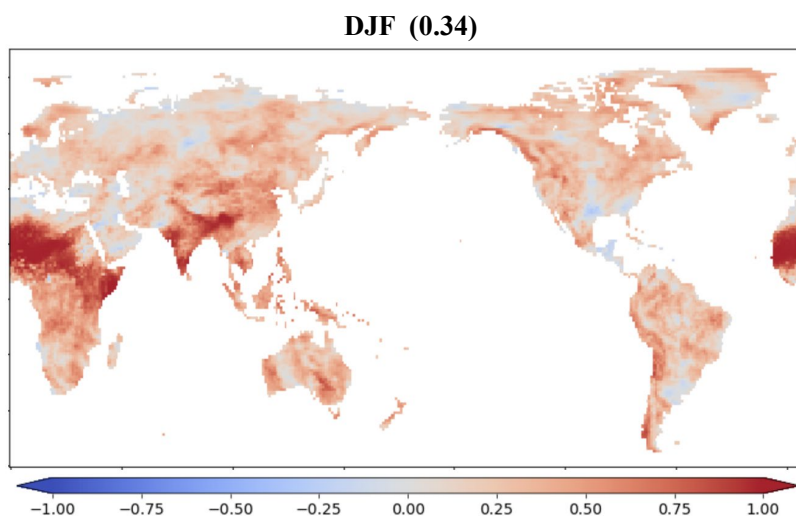


Fig. 3 DJF ACC (blue is less skill, red is more skill) for U-Net test set of week 3 precipitation prediction errors.

4. Conclusions

Future work will involve an extensive hyperparameter grid search to search for more optimal machine learning hyperparameters that may help the network converge and learn a more optimal solution. We will also employ the use of transfer learning, which proved very useful in Ham *et al.* (2019) for multi-year deep learning ENSO prediction, where we will first train the U-Net using all the individual ensemble members and upon loss plateauing during training (when loss is minimized), we will subsequently transfer learn with just one ensemble member, so that the neural network's weights can adjust to learn the single ensemble member's prediction errors. Future work will also involve the development of a large ensemble of deep learning models for an ensemble of bias-corrected forecasts, which will also allow us to explore sensitivity to variations in network hyperparameter settings, such as weight initializations. Finally, and importantly, we will also train using specific seasons and smaller spatial regions as opposed to the full yearly and global scale, allowing the network to learn localized patterns that are specific to smaller climate regions and seasons. These approaches will entail considerable computational usage (on GPUs), which will take place during an Accelerated Scientific Discovery period on NCAR's new supercomputer named Derecho (this coming Fall 2022).

References

- Adler, R. F., G. J. Huffman, A. Chang, R. Ferraro, P. Xie, J. Janowiak, B. Rudolf, U. Schneider, S. Curtis, D. Bolvin, A. Gruber, J. Susskind, and P. Arkin, 2003: The Version 2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-Present). *J. Hydrometeor.*, **4**, 1147-1167.

- Danabasoglu, G., J. F. Lamarque, J. Bacmeister, D. A. Bailey, A. K. DuVivier, J. Edwards, L. K. Emmons, J. Fasullo, R. Garcia, A. Gettelman, and C. Hannay, 2020: The Community Earth System Model Version 2 (CESM2). *J. Adv. Model. Earth Syst.*, **12**, e2019MS001916. doi:10.1029/2019MS001916.
- Ham, Y. G., J. H. Kim, and J. J. Luo, 2019: Deep learning for multi-year ENSO forecasts. *Nature*, **573**, 568-572, doi:10.1038/s41586-019-1559-7.
- He, S., X. Li, T. DelSole, P. Ravikumar, and A. Banerjee, 2020: Sub-seasonal climate forecasting via machine learning: Challenges, analysis, and advances. arXiv:2006.07972 [cs.LG], doi:10.48550/arXiv.2006.07972.
- Hersbach, H., B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, and A. Simmons, and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999-2049, doi:10.1002/qj.3803.
- Kim, H., Y. G. Ham, Y. S. Joo, and S. W. Son, 2021: Deep learning for bias correction of MJO prediction. *Nat. Commun.*, **12**, 1-7.
- Meehl, G. A., J. H. Richter, H. Teng, A. Capotondi, K. Cobb, F. Doblas-Reyes, M. G. Donat, M. H. England, J. C. Fyfe, W. Han, and H. Kim, 2021: Initialized Earth System prediction from subseasonal to decadal timescales. *Nat. Rev. Earth Environ.*, **2**, 340-357.
- Pegion, K., B. P. Kirtman, E. Becker, D. C. Collins, E. LaJoie, R. Burgman, R. Bell, T. DelSole, D. Min, Y. Zhu, and W. Li, 2019: The Subseasonal Experiment (SubX): A multimodel subseasonal prediction experiment. *Bull. Amer. Meteor. Soc.*, **100**, 2043-2060.
- Richter, J. H., A. A. Glanville, J. Edwards, B. Kauffman, N. A. Davis, A. Jaye, H. Kim, N. M. Pedatella, L. Sun, J. Berner, W. M. Kim, S. G. Yeager, G. Danabasoglu, J. M. Caron, and K. W. Oleson, 2022: Subseasonal Earth system prediction with CESM2. *Weather Forecast.*, doi:10.1175/WAF-D-21-0163.1.
- Ronneberger, O., P. Fischer, and T. Brox, 2015: U-net: Convolutional networks for biomedical image segmentation. *Medical image computing and computer-assisted intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. Wells, and A. Frangi, Eds., Springer, Cham., 234-241, doi:10.1007/978-3-319-24574-4_28.
- Vitart, F., C. Ardilouze, A. Bonet, A. Brookshaw, M. Chen, C. Codorean, M. Déqué, L. Ferranti, E. Fucile, M. Fuentes, and H. Hendon, 2017: The subseasonal to seasonal (S2S) prediction project database. *Bull. Amer. Meteor. Soc.*, **98**, 163-173.
- Weyn, J. A., D. R. Durran, R. Caruana, and N. Cresswell-Clay, 2021: Sub-seasonal forecasting with a large ensemble of deep-learning weather prediction models. *J. Adv. Model. Earth Syst.*, **13**, e2021MS002502, doi:10.1029/2021MS002502.

Ensemble Predictability of Week 3/4 Precipitation and Temperature over the United States via Cluster Analysis of the Large-Scale Circulation

Greg Jennrich,^{1,3} David Straus,² Muthuvel Chelliah,³ and Cory Baggett^{1,3}

¹*Earth Resources Technology, Inc.*

²*Department of Atmospheric, Oceanic, and Earth Sciences, George Mason University, Fairfax, VA*

³*Climate Prediction Center, NOAA/NWS/NCEP, College Park, MD*

1. Introduction

Forecasting the Week 3/4 period presents many challenges, resulting in a need for improvements to forecast skill. At this time spanned from initial conditions, numerical models struggle to present skillful forecasts of temperature, precipitation, and associated extremes. Nor does this period fully extend into the boundary-dependent climate time scale. One approach to improve Week 3-4 forecasts is to utilize the better predicted, large-scale circulation to make forecasts of temperature and precipitation anomalies, using the association between the preferred patterns of geopotential height (hereafter regimes) and surface weather obtained from reanalysis products. The functionality of regime classification has been well documented (Amini and Straus 2018, Riddle *et al.* 2013, Dawson and Palmer 2015). This study explores the utility of k-means cluster analysis of geopotential heights to identify regimes and using the forecasted regimes to make skillful predictions of temperature and precipitation in the Week 3/4 period.

2. Data and methods

ERA-Interim provides data for 500-hPa geopotential heights (z500) for 1979/80 to 2018/19 winters. To match the 2-week period of Week 3/4 forecasts, we use the 14-day running mean anomalies with end dates from November 23rd to March 23rd. We consider a domain over North America spanning the central Pacific to the western Atlantic (150°-330°E, 20°-80°N), which allows an analysis of common patterns of height anomalies across the region (teleconnections). The dimensionality of the data is reduced by retaining the leading 12 principal components (PCs) from empirical orthogonal function (EOF) analysis of this data, which explains about 85% of the variance in z500 anomalies. These PCs are used as inputs to the k-means clustering algorithm.

The goal of this k-means clustering method is to separate the clusters such that the ratio of the variance, an Eulerian distance metric, between clusters is maximized and the intracluster variance is minimized (Straus *et al.* 2007). Using Monte-Carlo PC data sets in which each synthetic PC retains its auto-covariance structure but is statistically independent of all other PCs, we determine that any number of clusters greater than 3 is statistically significant. Through sensitivity testing with hindcast skill, 6 clusters or k=6 were chosen as the number of clusters to consider.

With each 14-day period from each forecast assigned a cluster, we composite several weather variables, such as temperature, precipitation totals, and storm tracks, for each of the six clusters, where the association of each cluster with a weather variable is based on reanalyses. The composites we create are frequency composites. We first categorize a given variable's observed, 14-day anomaly as below, near, or above normal if it falls within the bottom, middle, or top tercile of its climatological distribution, respectively. Here, the tercile thresholds are defined by the 33rd and 67th percentiles and have been harmonically smoothed. Next, we find the percentage of occurrence of each of the three categories at each grid point for each cluster, resulting in three, cluster specific, frequency maps. Data from CPC's Global Unified Gauge-Base Precipitation reanalysis (Chen *et al.* 2008), CPC's 2-m Daily reanalysis (Fan and van den Dool 2008), and ERA-Interim reanalysis (Dee *et al.* 2011), provide the 14-day running periods synchronous with the cluster periods.

Hindcasts for z500 from the GEFSv12 and ECMWF extended forecasts allow for an investigation into the skill of a cluster-based forecast. We consider 252 winter (Nov 15- Feb 15) weekly initialization dates from 2000/01 to 2019/20 and verify the Week 3/4 forecast period (days +15 to +28). Forecasts are constructed by first assigning individual ensemble members to a cluster. The GEFSv12 and ECMWF provide 11 ensemble members each to assign. A final forecast is constructed from weighting the cluster-based composites by the percentage of ensemble members assigned to each cluster. Skill is measured by a two-category Heidke skill score (HSS) for the CONUS and Alaska, where the forecast corresponds to the category with the maximum, cluster-weighted probability. Thus, the above and below normal forecast points are compared with the corresponding two-category observations. Near normal forecast grid boxes are ignored for scoring due to the relatively low occurrence in the forecasts. The resulting three-category HSS would be significantly reduced. The aforementioned CPC datasets are used to verify the forecasts.

3. Results

The k-means clustering provides the 6 most common height patterns (clusters) seen in the North American cold-season. Figure 1 reveals the height anomalies associated with each cluster. While the clusters are numbered from 1 to 6, there is no significance in their order. Each of the clusters is comprised of various trough/ridge patterns, some related to common teleconnections. For example, a strong positive North Atlantic Oscillation (+NAO) is evident in Cluster 2, while the opposite signal (-NAO) stands out in cluster 4. The Pacific North

DJF 500 hPa Cluster Patterns 1979-2018

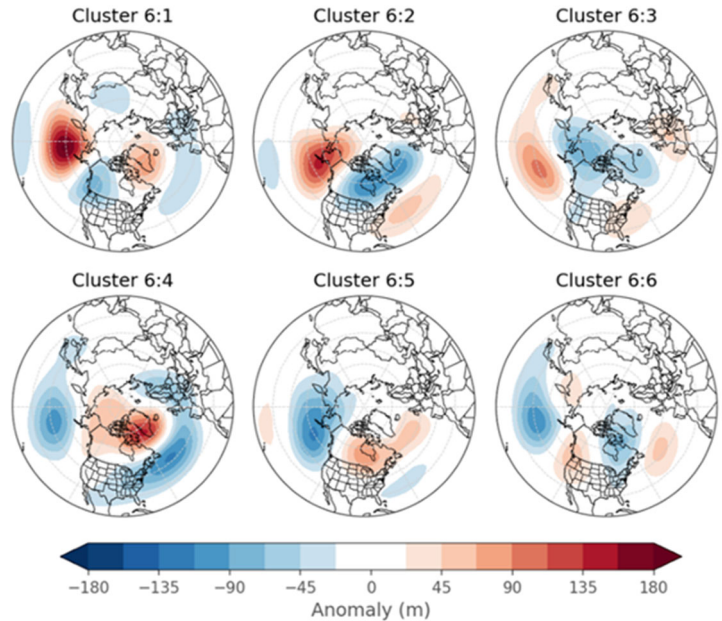


Fig. 1 500 hPa geopotential height anomalies (in meters) for the 6 North American Clusters. K-means cluster analysis provides the cluster assignments used to calculate these composites. There is no significance in the ordering of the clusters.

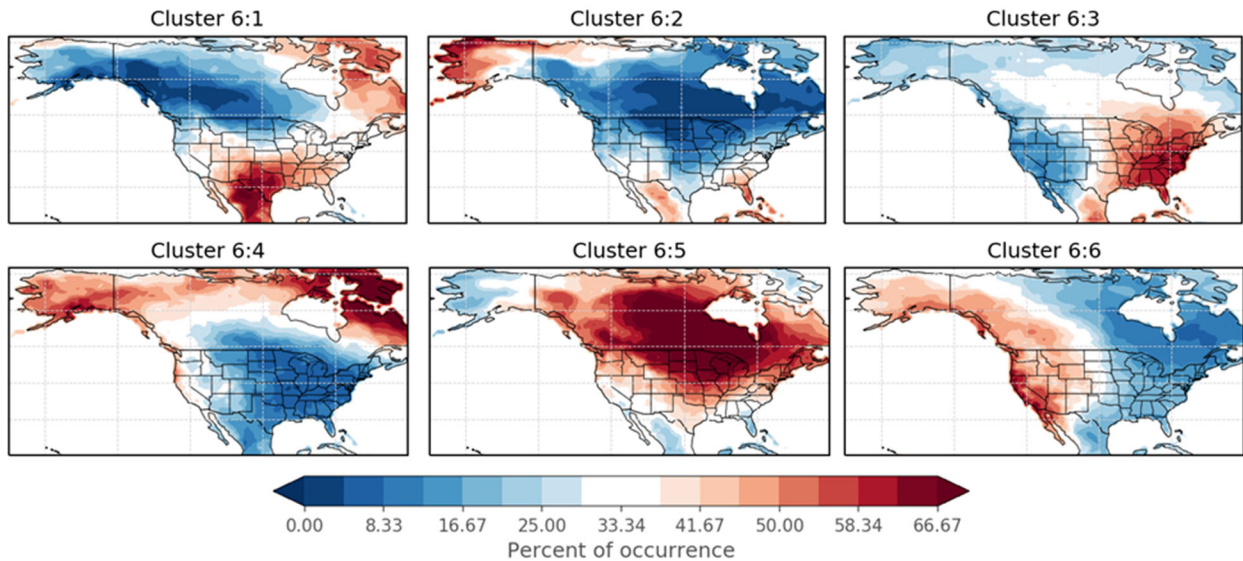


Fig. 2 Cluster composites of above normal temperature frequencies. For each 14-day period assigned a specific cluster, it is determined how frequently each grid point was observed to fall into the above normal temperature tercile. Climatological normal would be near 33.3 %.

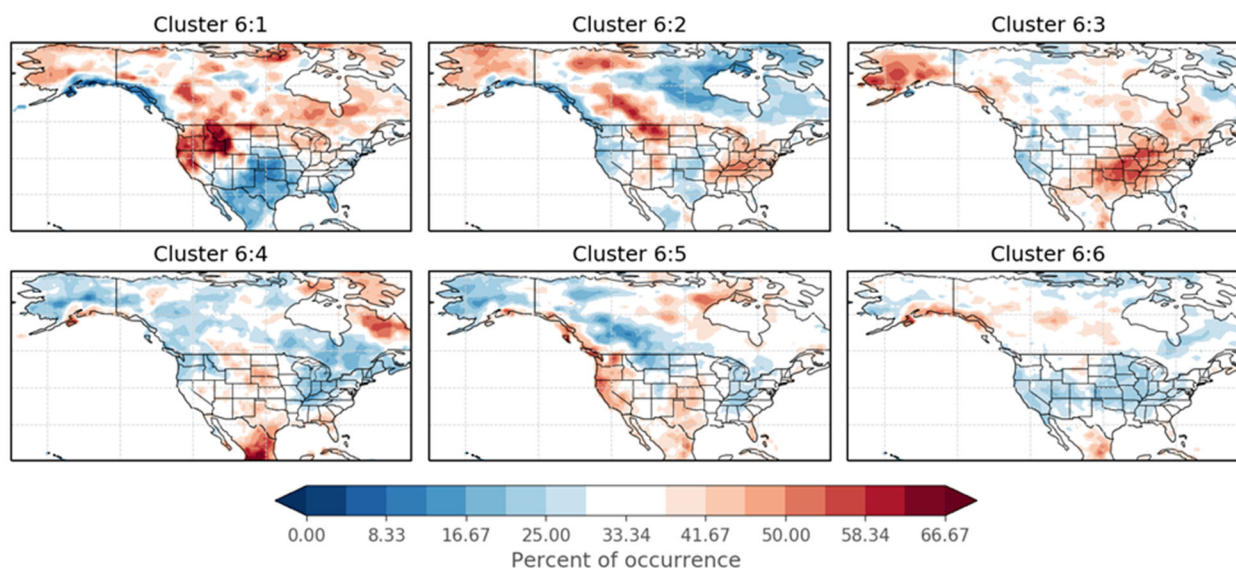


Fig. 3 Same as in Fig. 2, but for above normal precipitation.

American (PNA) pattern is also connected to anomalous wintertime weather (Wallace and Gutzler 1981; Baxter and Nigam 2015) and is loosely found in all clusters, but most apparent as a negative PNA signal in cluster 3. Additionally, each of the clusters has an opposing cluster, with near opposite height patterns. These sets include: Clusters 1 and 6, Clusters 2 and 5, and Clusters 3 and 4.

With the clusters defined, the next step is to investigate anomalous 2-m temperature and precipitation for each of the clusters. Figure 2 shows the frequency of occurrence of above normal temperatures for each of the clusters. We would expect a random sample of periods to have each tercile category occur around 33% of the time. Where frequencies are greater than 33%, there are increased probabilities of above normal temperatures. While percentages less than 33% could result in greater probabilities for the below or near normal terciles, we can assume through a similar figure for below normal temperatures (not shown) that its often an indicator for an increased frequency of below normal temperatures. Each of the 6 clusters have distinct temperature tercile occurrences. For example, cluster 2 signals a strong cold air outbreak over much of the CONUS. In contrast, Clusters 3 and 6 have split, east west terciles.

The same analysis is performed for precipitation in Figure 3. Although the signals are noisier, there are several areas of increased frequencies of above normal precipitation. For example, Cluster 3 features a swath of increased frequencies from the southern Great Plains to the Great Lakes, while the Pacific Northwest sees increased frequencies during Clusters 1 and 5. Cluster 6 suggests suppressed precipitation across much of the Lower 48 of the United States. Each of the clusters have well-defined temperature and precipitation anomalies that give insight into their associated surface conditions.

Hindcasts from the GEFS and ECMWF allow us to compare skill between forecasts based on the cluster framework and raw forecasts from the dynamical models. Table 1 displays the temperature and precipitation HSS for the cluster and model forecasts (rows 1 and 2), as well as possible forecasts of opportunity (rows 3 and 4). For all of the hindcasts, cluster forecasts are skillful, but not quite to the level of the dynamical models. We acknowledge that this framework may not be applicable to every week 3/4 forecast, thus we explore possible forecasts of opportunity in the bottom two rows. We define forecasts of opportunity as instances when 75% or more of the ensemble members are assigned to two clusters or less, which would suggest that there is more confidence in the utility of the cluster framework. This occurred in just under 25% of the hindcasts. These forecasts were ~8 and 10 HSS points better than the entire hindcast dataset for temperature and precipitation, respectively. As a fair comparison, the same initialization dates that were identified as forecasts of opportunity are scored for the dynamical models, where skill scores increased as well. Since we know the cluster assignment distribution beforehand, this method to identify forecasts of opportunity shows some promise.

Table 1 Heidke Skill Scores based on 2-category observations for GEFSv12/ECMWF winter hindcasts from 2000-2019, scored over CONUS/AK. For cluster forecasts, near normal forecast points are ignored and not scored.

Heidke Skill Scores

Forecast Type	Forecast Category	Forecast Count	Temperature	Precipitation
Cluster: Maximum Category	Above/Near/Below, Near=Ignore	252	18.9	7.9
GEFS/ECMWF: Raw Anomaly	Above/Below	252	22.9	10.3
Cluster: Top 2 Clusters Assignments sum $\geq 75\%$	Above/Near/Below, Near=Ignore	50	27.0	17.6
GEFS/ECMWF: Raw Anomaly (Same Forecasts as Row 3)	Above/Below	50	28.2	15.6

4. Discussion

This investigation sheds light on the most frequent circulation regimes during the North American winter. These regimes are often similar to combinations of teleconnections such as the NAO and PNA. The 6 clusters and their associated composites present a useful tool for forecasters by providing detailed probabilistic maps of above, near, or below normal temperature and precipitation anomalies based on the Week 3/4 height forecast. Assigned clusters from each ensemble member can also be used to weigh and formulate cluster-based temperature and precipitation forecasts.

While not scoring as high as the dynamical model forecasts during the hindcast period, the cluster forecasts provide skillful insight (see Table 1) to the temperature and precipitation Week 3/4 forecast from just the z500 field. There seems to be a few things to consider from this hindcast assessment. First, temperature forecasts rarely have near normal classification (not shown). This is in part due to the robust anomalies seen in Figures 1-3. Above and below normal terciles are much more common across North America than the near normal terciles. The lean towards end terciles is likely attributed to the anomalous nature of the cluster analysis. Clusters, by nature, are the most common anomaly regimes. Since surface variables are correlated, we should expect our cluster composites to be quite anomalous themselves.

Second, there is no established way to utilize the ensemble members to make a final forecast. The weighting of clusters from cluster assignments is a basic, yet useful, way to calculate a forecast, but there are other ways to achieve this. If some ensemble members are not closely correlated with any of the clusters, it may be best to simply throw them out, so that we do not force a cluster classification. Also, we treat all members as equal contributions. Perhaps, weighting members by the correlation with clusters would result in a better forecast.

Finally, there are a few more items that we plan to investigate. While we have a skill measurement based on hindcasts, a realtime measurement will be required. This will need to include more models to increase the ensemble members to consider. Additionally, a comparison of different methods for making forecasts, described above, would need a skill assessment in realtime. While the focus has been on temperature and precipitation, other variables should be considered. For example, an investigation into storm tracks has begun and generally corresponds with the clustering by height fields. Likewise, extremes for any variable can undergo the same analysis. Extreme cold and precipitation (15th and 85th percentiles, respectively) forecasts result in similar maps as the tercile temperature and precipitation forecasts. Lastly, the expansion outside of DJF to account for all seasons should be performed. In all, this work represents an important first step in addressing the utility of regime analysis for Week 3/4 forecasts.

References

- Amini, S., and D. M. Straus, 2018: Control of storminess over the Pacific and North America by circulation regimes. *Clim. Dyn.* doi:10.1007/s00382-018-4409-7.
- Baxter, S., and S. Nigam, 2015: Key role of the North Pacific Oscillation–west Pacific pattern in generating the extreme 2013/14 North American winter. *J. Climate*, **28**, 8109–8117.

- Chen, M., and Coauthors, 2008: Assessing objective techniques for gauge-based analyses of global daily precipitation. *J. Geophys. Res.*, **113**, D04110. doi:10.1029/2007JD009132
- Dawson, A., and T. N. Palmer, 2015: Simulating weather regimes: impact of model resolution and stochastic parameterization, *Clim. Dyn.*, **44**, 2177–2193. doi:10.1007/s00382-014-2238-x.
- Dee, D. P., and Coauthors 2011, The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q.J.R. Meteorol. Soc.*, **137**, 553–597. doi:10.1002/qj.828
- Fan, Y., and H. van den Dool, 2008: A global monthly land surface air temperature analysis for 1948-present. *J. Geophys. Res.*, **113**, D01103. doi:10.1029/2007JD008470
- Riddle, E. E., M. B. Stone, N. C. Johnson, M. L. L’Heureux, D. C. Collins, and S. B. Feldstein, 2013: The impact of the MJO on clusters of wintertime circulation anomalies over the North American region. *Clim. Dyn.*, **40**, 1741–1766.
- Straus, D. M., S. Corti and F. Molteni, 2007: Circulation Regimes: Chaotic Variability versus SST-Forced Predictability. *J. Climate*, **20**, 2251–2272
- Wallace, J. M., and D. S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Mon. Wea. Rev.*, **109**, 784–812

Meta-heuristic Ant Colony Optimization Technique to Forecast the Amount of Summer Monsoon Rainfall: Skill Comparison with Markov Chain Model

Sayantika Mukherjee

Amity University, Kolkata, India

Forecasting summer monsoon rainfall with precision becomes crucial for the farmers to plan for harvesting in a country like India where the national economy is mostly based on regional agriculture. In the present study the meta-heuristic method of ant colony optimization (ACO) technique is used to forecast the amount of summer monsoon rainfall over an urban station Kolkata (Chaudhuri *et al.* 2014). ACO technique takes inspiration from the foraging behavior of some ant species. The ants deposit pheromone on the ground in order to mark a favorable path that should be followed by other members of the colony. A range of rainfall amount replicating the pheromone concentration is evaluated during the summer monsoon season. The maximum amount of rainfall during summer monsoon season (June – September) is observed to be within the range of 7.5 to 35 mm, the Range – 4 category set by the India Meteorological Department (IMD) during 1998 to 2007. The transitional probabilities of rainfall for consecutive two days during the summer monsoon season are computed using ACO technique and compared with Markov Chain Model (MCM) (Fig. 1). The result reveals that the accuracy in forecasting the amount of rainfall for two successive days using ACO technique and MCM are 95% and 83% respectively. The accuracy of the forecast is validated with the IMD observations from 2008 and 2012.

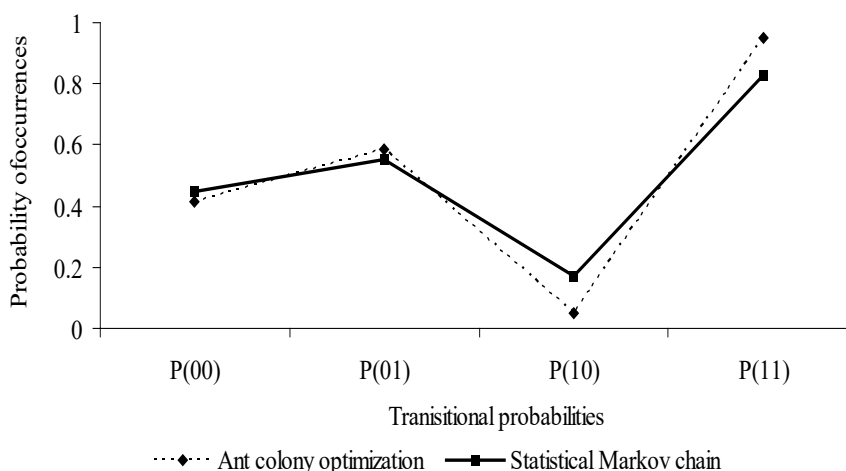


Fig. 1 Transitional probabilities for occurrences of rainfall for successive two days during monsoon season using ant colony optimization technique and statistical Markov chain model for the period from 1998 to 2007 over Kolkata.

References

Chaudhuri, S., S. Goswami, D. Das, and A. Middey, 2014: Meta-heuristic ant colony optimization technique to forecast the amount of summer monsoon rainfall: skill comparison with Markov chain model. *Theor. Appl. Climatol.*, **116**, 585–595.

Enhancing Subseasonal Temperature Prediction by Bridging a Statistical Model with Dynamical Arctic Oscillation Forecasting

Changhyun Yoo,¹ Minju Kim,¹ and Jung Choi²

¹Department of Climate and Energy Systems Engineering, Ewha Womans University, Seoul, Korea

²School of Earth and Environmental Sciences, Seoul National University, Seoul, Korea

ABSTRACT

This study proposes a hybrid approach to improving subseasonal prediction skills by bridging a conventional statistical model and a dynamical ensemble forecast system. Based on the perfect prognosis method, the phase of the Arctic Oscillation (AO) from the European Centre for Medium-range Weather Forecasts ensemble forecast system is used as a predictor in a composite based statistical model to predict the wintertime surface air temperature in the Northern Hemisphere. The hybrid model, which employs AO phases predicted by the dynamical model for weeks 1–4, generally outperforms the conventional statistical model for lead times of weeks 2–6 (Fig. 1). The improved skill score is due to the high accuracy of the AO forecast from the dynamical model and the strong lagged connection between the AO and temperature. This study thus lays the groundwork for the potential use of combining climate variability, statistical relation, and dynamical forecasting.

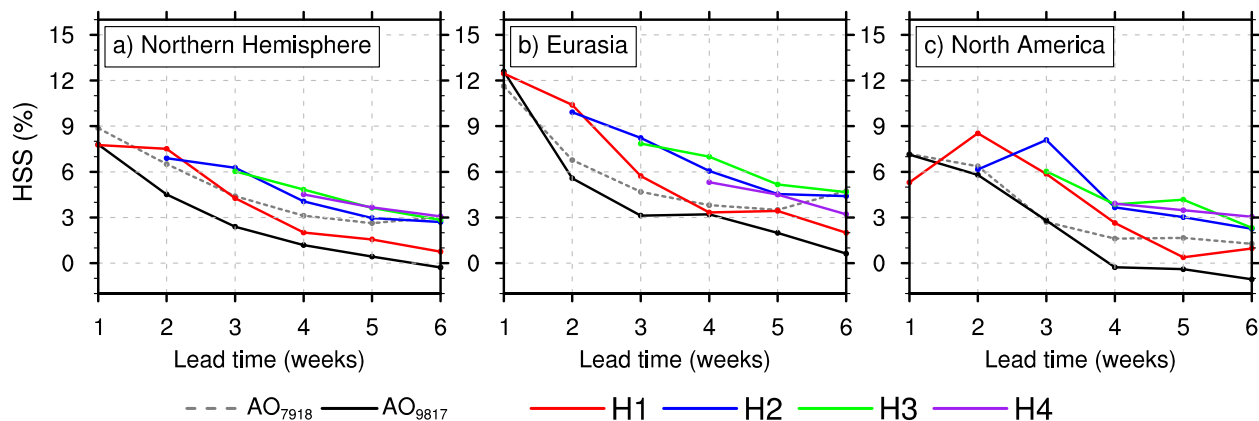


Fig. 1 Area-averaged DJF Heidke Skill Score (HSS) over (a) Northern Hemisphere, (b) Eurasia, and (c) North America from lead-time weeks 1-6. The dashed and solid black lines represent the two conventional models AO7918 and AO9817, respectively. The colored lines represent the hybrid models: H1 (red), H2 (blue), H3 (green), and H4 (purple).

* AO7918 and AO9817 are conventional statistical phase models constructed using the entire DJF between 1979/80 and 2017/18 and between 1998/99 and 2016/17, respectively. H1-H4 are different versions of hybrid dynamical-statistical model built with the dynamically forecasted AO phase for forecast weeks 1-4.

This study has been published in *Geophysical Research Letters* in 2021.

Reference

Kim, M., C. Yoo, and J. Choi, 2021: Enhancing subseasonal temperature prediction by bridging a statistical model with dynamical Arctic Oscillation forecasting. *Geophys. Res. Lett.*, **48**, e2021GL093447, doi:10.1029/2021GL093447.



4. Development and Use of Climate Data Records



U.S. Climatological Standard Normals: A Utilitarian Workhorse

Michael A. Palecki

NOAA National Centers for Environmental Information, Asheville, NC

1. Overview

The concept of the normal as a climatological reference has been around since the 19th Century, and was adopted into the terminology of the World Meteorological Organization (WMO) from its very beginning in 1950, and even before by its predecessor the International Meteorological Organization (IMO). The first normal period adopted worldwide was 1901-1930 at a meeting of the IMO in 1935. A normal dataset consists of averages and other statistics derived from a particular range of years, so as to provide a uniform context for comparing conditions across time and geography. The United States started updating normals every ten years in the mid-1950s, but the WMO required it of its member nations only every thirty years, most recently for 1931-1960, 1961-1990, and 1991-2020. In 2015, the WMO changed its requirements to recommend that nations update normals every ten years in light of the speed with which the underlying global climate was changing.

Besides being for a specific period, normals are different from averages in how they are calculated (Arguez *et al.* 2012). Specific procedures have been recommended in the WMO Guidelines on the Calculation of Climate Normals, WMO-No. 1203 (WMO 2017). In the United States, official climatological standard normals for observation stations are calculated by the NOAA National Centers for Environmental Information (NCEI) and provided to the National Weather Service (NWS) and all other users in the public and private sector so they are the same for all. NCEI procedures correct temperature values for non-climatic influences using a process of pairwise homogenization (Menne and Williams 2009), and daily temperature normals are smoothed by applying a constrained harmonic fit to the values (Arguez and Applequist 2013). Precipitation values are improved by filling missing values from surrounding stations (Durre *et al.* 2013). At the end of these processes, the accumulated daily normals match precisely to monthly normals.

Normals are designed to perform two main tasks: 1) act as a uniform baseline for understanding how today's weather compares to the weather of a uniform reference period, and 2) provide a set of averages and statistics that can be used to make decisions for actions/activities that are impacted by weather but are occurring at times beyond the reach of reliable dynamical weather forecasts. The first is always done with the standard 30-year period, while the second can utilize the conventional normals or other baseline periods, such as 2006-2020 15-year normals that are also provided by NCEI. In total more than 500 variables are calculated over annual, seasonal, monthly, daily, and hourly time intervals, including averages, maxima and minima, percentiles, threshold exceedance counts, growing season statistics, and others.

2. Highlights

The main release of the 2020 U.S. Climate Normals took place at the beginning of May 2021, and all U.S. normals can be found through a single newly designed web site located at: <https://www.ncei.noaa.gov/products/land-based-station/us-climate-normals>. A variety of access methods are available, from downloading the entire dataset to extracting temperature and precipitation normals for a single station of interest (Fig. 1).

Climate normals are available for stations at more than 15000 locations for precipitation and more than 7000 locations for temperature. Most of these data are collected by citizen scientists in the NWS Cooperative Observer Program (COOP) Network or the Community Collaborative Rain, Hail and Snow (CoCoRaHS) Network. Despite this coverage, users require normals for every location in the contiguous U.S. While this is not possible with observation stations, sets of gridded normals have been constructed at a 4 km grid scale (1/24

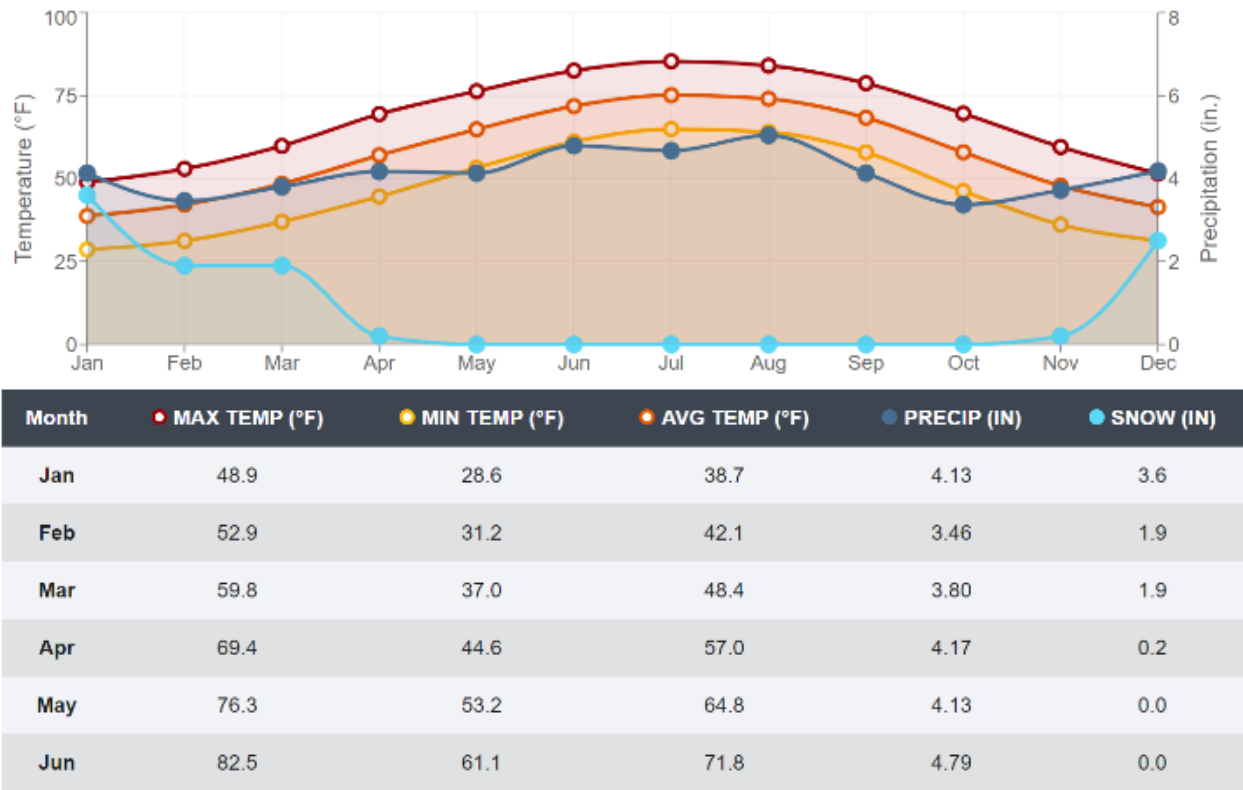


Fig. 1 Conventional monthly normals for Asheville Regional Airport., NC.

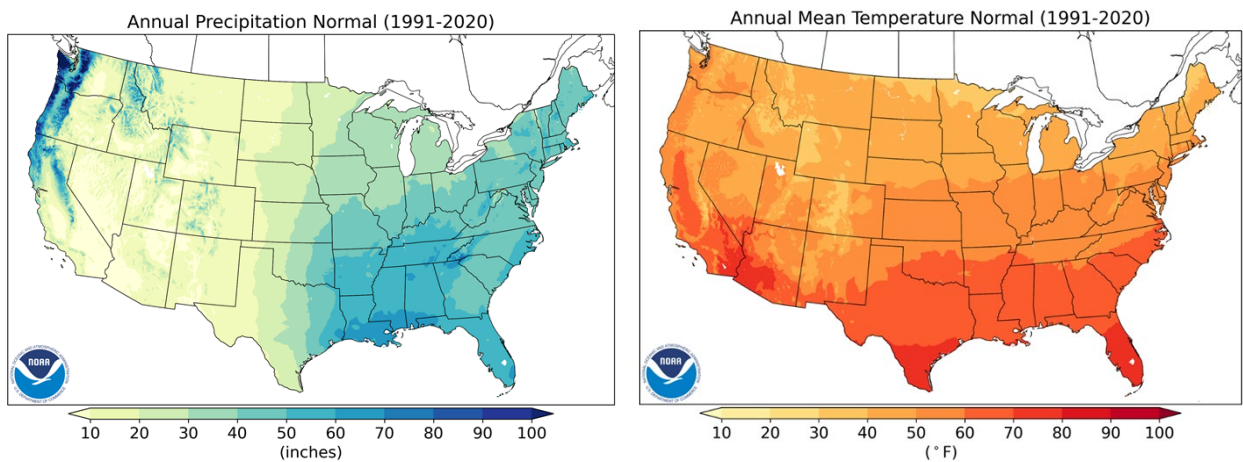


Fig. 2 Annual gridded climate normals for the contiguous U.S.

by 1/24 degree latitude / longitude) due to the development at NCEI of monthly grids calculated using climatologically guided interpolation techniques (Vose *et al.* 2014). Maps are available for users (Fig. 2), as well as the data values for each monthly climate normal grid.

One of the highlights of the availability of the new normals was the opportunity to look at how conventional normals have changed since the last cycle. Subtracting 1981-2010 normals from 1991-2020 yields patterns of substantial changes that are not uniform (Fig. 3).

Conditions have become generally wetter in the central and east U.S., and drier in the west and southwest U.S., while the U.S. is warmer everywhere except in the north-central U.S. However, changes do vary by month, with patterns shifting geographically through the months. April temperature normals are cooler in a wide swath

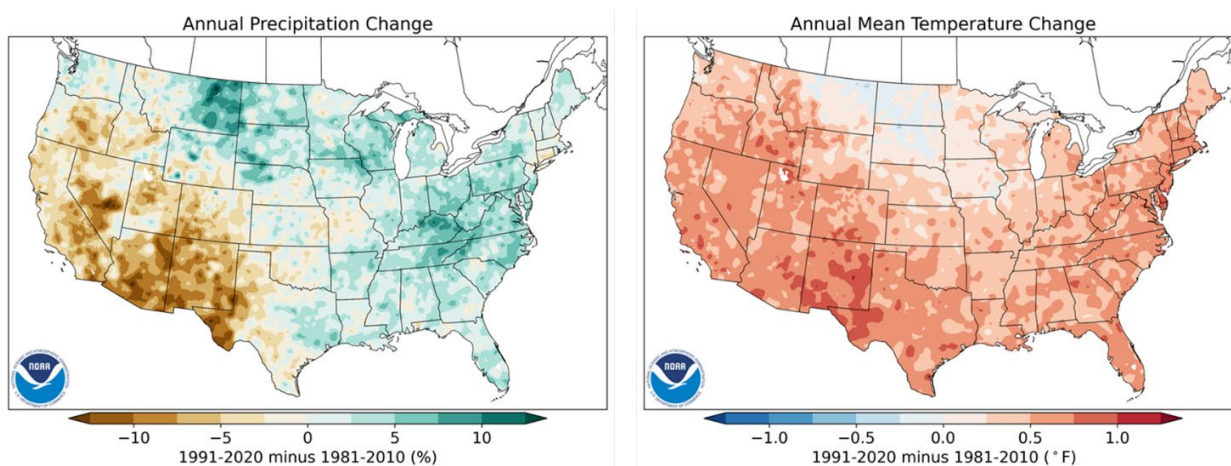
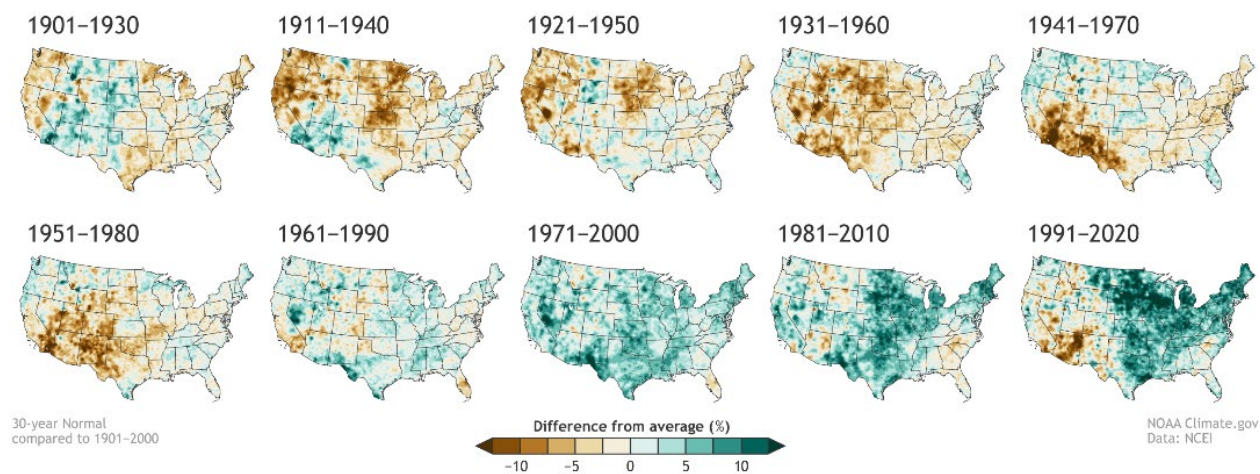


Fig. 3 Change in climate normals from 1981-2010 to 1991-2020.

U.S. ANNUAL PRECIPITATION COMPARED TO 20th-CENTURY AVERAGE



U.S. ANNUAL TEMPERATURE COMPARED TO 20th-CENTURY AVERAGE

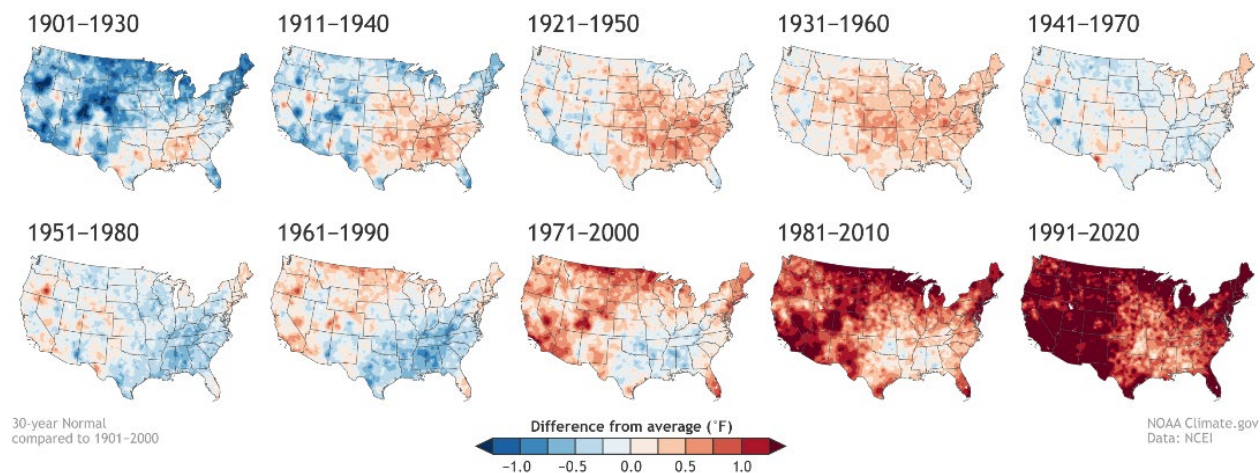


Fig. 4 The fingerprints of human-induced climate change have emerged since the 1970s.

of the north-central U.S., indicating the impacts of atmospheric circulation and snow cover differences between the 1980s and 2010s, since both normals have the years 1991-2010 in common. In the southeast U.S., October temperatures have warmed, November temperatures cooled, and December temperatures warmed, indicating a change in the shape of the seasonal cycle in that region.

3. Discussion

While the use of 2020 U.S. Climate Normals is required by NOAA NWS and many other government and private sector entities, the normals concept should be adopted more broadly. It is very important to use the same period-of-record and calculation approaches when trying to compare anomalies from different datasets or models. Models are especially flexible in the selection of a baseline or normals period, but if comparing between models and observations, or two or more different models, it is best practice to base those comparisons on a time period drawn from when they overlap. This time period must also link to the application. If trying to understand responses to anomalies or extremes in the current climate, that period should be as close to the present as is feasible. However, there has been consideration given to the use of older or longer baseline periods for examining climate change over time, to avoid the possibility that the scale of change will be obscured by using a shifting baseline over the decades (Hulme 2020). WMO recognizes 1961-1990 as a best period to use when examining long term climate change (WMO 2017), while NCEI climate monitoring products use a centennial period between 1901-2000, and still other periods have been used by the Intergovernmental Panel on Climate Change in their reports over time. However, the power of using both by comparing 30-year normals to a 1901-2000 baseline can be seen in Figure 4, which uses a series of gridded normals to illustrate that climate change formed coherent trajectories after the mid-1900s that continue into the present. Normals are a utilitarian workhorse as defined, and can also be a useful concept applied in a variety of circumstances beyond their basic origins.

References

- Arguez, A., I. Durre, S. Applequist, R. S. Vose, M. F. Squires, X. Yin, R. R. Heim Jr., and T. W. Owen, 2012: NOAA's 1981-2010 U.S. climate normals: An overview. *Bull. Amer. Meteor. Soc.*, **93**, 1687-1697, doi:10.1175/BAMS-D-11-00197.1.
- , and S. Applequist, 2013: A harmonic approach for calculating daily temperature normals constrained by homogenized monthly temperature normals. *J. Atmos. Oceanic Technol.*, **30**, 1259-1265, doi:10.1175/JTECH-D-12-00195.1.
- Durre, I., M. F. Squires, R. S. Vose, X. Yin, A. Arguez, and S. Applequist, 2013: NOAA's 1981-2010 U.S. climate normals: Monthly precipitation, snowfall, and snow depth. *J. Appl. Meteorol. Climatol.*, **52**, 2377-2395, doi:10.1175/JAMC-D-13-051.1.
- Hulme, M., 2020: Climates multiple: Three baselines, two tolerances, one normal. *Academia Lett.*, Article 102, doi:10.20935/AL102.
- Menne, M. J., and C. N. Williams Jr., 2009: Homogenization of temperature series via pairwise comparisons. *J. Climate*, **22**, 1700-1717, doi:10.1175/2008JCLI2263.1.
- Vose, R. S., S. Applequist, M. Squire, I. Durre, M. Menne, C. N. Williams Jr., C. Fenimore, K. Gleason, and D. Arndt, 2014: Improved historical temperature and precipitation time series for U.S. climate divisions. *J. Appl. Meteor. Climatol.*, **53**, 1232-1251, doi:10.1175/JAMC-D-13-0248.1.
- World Meteorological Organization (WMO), 2017: *WMO Guidelines on the Calculation of Climate Normals* (WMO-No. 1203). Geneva.

Communicating Uncertainty in SST Analysis

Caihong Wen,¹ Arun Kumar,¹ Wanqiu Wang,¹ Michelle L'Heureux,¹
Pingping Xie,¹ Zeng-Zhen Hu,¹ and Bert Katz^{1,2}

¹Climate Prediction Center, NOAA/NWS/NCEP, College Park, MD

²Earth Resources Technology, Inc., Laurel, MD

1. Introduction

Sea surface temperature (SST) is one of most important indicators of Earth's climate because of its influence on the heat, momentum, and freshwater exchanges at the ocean-atmosphere interface on weather and climate time scales. It is an essential variable for weather and climate forecasting, forecast verification, and monitoring. SST analysis, which provides smooth SST data with global coverage, is often used at operational agencies to monitor the state of climate variability.

Several operational centers and research groups have developed and released real-time SST products. The SST analyses are created by applying interpolation and statistical techniques (*e.g.*, observational/background error correlation scales, input data bias correction) on SST observations (*i.e.*, satellite based, in situ platforms, ship, autonomous vehicles *etc.*) (O'Carroll *et al.* 2019). Uncertainty among SST analyses exists because of different observation data sources, sampling, and analysis techniques. The Climate Prediction Center (CPC) currently utilizes multiple SST data sets to monitor climate variability and outlooks on time scales ranging from weeks to seasons. On occasion, there are large differences among SST products in areas where SST variations have important influence on climate variability, such as the ENSO region of the central-eastern equatorial Pacific Ocean, hurricane main development area, *etc.* This creates challenges for interpreting climate monitoring and prediction products that depend on the sign and amplitude of SST anomalies. Our objectives are to help users understand consistency and discrepancy among SST products used at CPC, and to inform users about the characteristics of these data sets, which informs their use in applications.

In addition to the importance to climate monitoring, SST analysis is widely used as boundary conditions in the weather and climate forecast systems and provides a constraint on SST evolution in ocean data assimilation systems. For ocean reanalysis, the ocean models are forced by surface heat fluxes from atmospheric reanalysis, which often have biases, leading to drifts in model SSTs. A strong relaxation of model SST to an SST analysis is applied as a flux correction, so that the evolution of model SST stays close to the evolution of specified SST analysis. This is also referred as SST nudging, which is widely used in operation ocean reanalysis systems, such as NCEP Global Ocean Data Assimilation System (GODAS) (Behringer and Xue 2004), Climate Forecast System Reanalysis (CFSR) (Saha *et al.* 2010), ECMWF Ocean Reanalysis System 4 (ORAS4) (Balmaseda *et al.* 2013) and ORAS5 (Zuo *et al.* 2019). Previous studies reported that the fidelity of ECMWF ocean reanalysis was sensitive to the SST nudging sources in the data assimilation systems (Balmaseda *et al.* 2013; Zuo *et al.* 2019). NCEP CFSR ocean reanalysis switched its SST nudging source from NOAA daily OISSTv2 to a new NOAA/EMC analysis called "Near Surface Sea Surface Temperature (NSST)" in early 2020. Moreover, the CFSR ocean reanalysis is used to provide oceanic initial conditions for NCEP Climate Forecast System version 2 (CFSv2) (Saha *et al.* 2014). How the CFSR and CFSv2 respond to the change in the SST nudging source has not been investigated yet. The other objective of this study is to explore the potential impacts of SST nudging source replacement on CFSR and CFSv2 monthly forecasts.

2. Data sets and methods

CPC currently employs five SST analyses for climate monitoring and forecast verification. It includes: NOAA weekly Optimum Interpolation (OI) v2.0 SST or NCEP OI SST (hereafter referred to as OISST, Reynolds *et al.* 2002).

- NOAA Extended Reconstructed SST version 5 (ERSSTv5, Huang *et al.* 2017).
- NOAA $\frac{1}{4}$ Daily OISSTv2.1 (hereafter referred to as OISSTv2.1, Reynolds *et al.* 2007; Huang *et al.* 2020).
- NCEP Near Surface SST (NSST) (NWS, 2020).
- U.K. Met Office Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system (Donlon *et al.* 2012 a, b).

These SST analyses depend on different observational input and reanalysis techniques, and therefore, meet different requirements. OISST, OISSTv2.1, NSST and OSTIA use both in situ and satellite observations. These data sets rely on different satellite streams and have different capabilities of resolving temporal and spatial variations. OISST provides global weekly data on $1^\circ \times 1^\circ$ grid. OISSTv2.1, NSST and OSTIA were specifically designed to provide accurate high spatial and temporal resolution SST estimates, so that they can be utilized in operational applications such as real-time analysis, and/or boundary conditions in numerical weather/climate forecasting models. It is noteworthy that OISSTv2.1 shares the same data value of daily OISSTv2.0 prior January 2016. ERSSTv5 only uses in-situ observations, which is relatively sparse compared to the satellite coverage, but this trade-off allows for a much longer record than SST datasets that use satellite observations. This dataset is specifically designed to be suitable for long-term monitoring of global and basin-wide climate variability. For example, NOAA uses ERSSTv5 to define ENSO conditions going back to 1950 because this data set is more homogeneous over time and allows more consistent comparisons among a larger number of historical ENSO events.

SST analyses can also be classified based on the depth, for which the SST analysis representatives of OISST, OISSTv2.1 and ERSSTv5 represent “Bulk SST” (roughly 0.5m in depth), while NSST and OSTIA are representatives of the foundation temperature (*i.e.*, the temperature is free or nearly free of diurnal cycle at roughly 10 m in depth). The characteristics of the five SST analyses are summarized in Table 1.

Except for NSST, all the other SST anomalies are computed as departures from their own climatology reference period 1991-2020. NSST is only available after 2015, which is too short to define a 30-year climatology, so its anomalies are defined as departures from the OISSTv2.1 climatology. Maps of CPC real-time daily, weekly, and monthly SST analyses are available at https://origin.cpc.ncep.noaa.gov/products/GODAS/multiSST_body.html.

In this study, all the SST products were gridded on a common spatial resolution as OISST (regular $1^\circ \times 1^\circ$ grid). The uncertainty/consistency among the SST analyses is quantified by taking the difference, root-mean-square difference (RMSD), and the correlation between an individual SST product and a “reference product”. In this study, OISSTv2.1 serves as the benchmark analysis. Hereafter, bias refers to the difference between the

Table 1 SST product comparison summary

Dataset	Spatial resolution	Temporal resolution	Time range	Type of SST	Usage examples
NCEP OISST, (OISST)	Global $1^\circ \times 1^\circ$	Weekly	Nov 1982 - present	Bulk SST	CPC operational products (weekly ENSO update, Ocean briefing, <i>etc.</i>)
NOAA ERSSTv5 (ERSST)	Global $2^\circ \times 2^\circ$	Monthly	1854 - present	Bulk SST	ENSO Diagnostic Discussion
NOAA Daily OISSTv2.1 (OISSTv2.1)	Global $0.25^\circ \times 0.25^\circ$	Daily	Sep 1981-present (1981 - 2015 is identical with Daily OISSTv2)	Bulk SST	CFSv2 forecast validation; SST nudging source for CFSR ocean reanalysis prior February 2020
NCEP NSST (NSST)	$0.083^\circ \times 0.083^\circ$	Daily	July 2015 - present	Foundation SST	SST nudging source for CFSR ocean reanalysis since February 2020
Met Office OSTIA (OSTIA)	$0.05^\circ \times 0.05^\circ$	Daily	1985 - present	Foundation SST	EMC model validation

various individual SST analyses and OISSTv2.1. For the SST data set with daily updates (OISSTv2.1, NSST and OSTIA), weekly data is derived from the weekly average centered on Wednesday, the same timestamp as the weekly OISST. Monthly data is the average of daily data over a month. For OISST, which has a native resolution of weekly averages, the monthly data is derived by a linear interpolation of the weekly output to the daily field and then averaging the daily values over the month.

We also assessed the fidelity of SST products by validating against the in-situ TAO moored buoy data from January 2016 to September 2021 (McPhaden *et al.* 1998). ERSST was excluded from this comparison of weekly data because it is only available in monthly resolution. All the SST analyses are interpolated onto TAO's grid and were sampled identically in time as the buoy data. For example, when the buoy data were not available at a specific period, SST analysis data for the same period was also set as missing values.

We also assess the impact of the recent SST nudging source replacement on CFSR ocean reanalysis and CFSv2 forecasts. For CFSR, SST anomalies are departures from its 1991-2020 monthly average. For CFSv2 forecasts, SST ensemble anomaly forecasts are downloaded from IRI Data Library (Kirtman *et al.* 2014, <https://iridl.ldeo.columbia.edu/SOURCES/.Models/.NMME/>).

3. Results

3.1 Uncertainty among SST analyses

ENSO is the most dominant mode of interannual coupled atmosphere-ocean variability. At NOAA, the Niño3.4 anomaly (SST anomaly average in [170°-120°W, 5°S-5°N]) is one of the most important indicators to monitor the occurrence and evolution of ENSO. Figure 1a compares the time series of weekly Niño3.4 from various SST analyses. It shows that uncertainty among the four SST analyses varies with time. The most striking feature is that OISST (red line) exhibits pronounced high-frequency fluctuations superimposed on ENSO timescales. Sometimes the differences between OISST and other datasets are greater than 0.5°C. For example, Niño3.4 was near zero in OISSTv2.0 on September 29, 2021 and then dropped to -0.6°C on October 6, 2021. Although OISSTv2.1, OSTIA and NSST are not identical with each other, they were all consistently below -0.5°C (threshold value to define La Niña) in these two weeks. CPC has used the weekly OISST for the official weekly ENSO monitoring

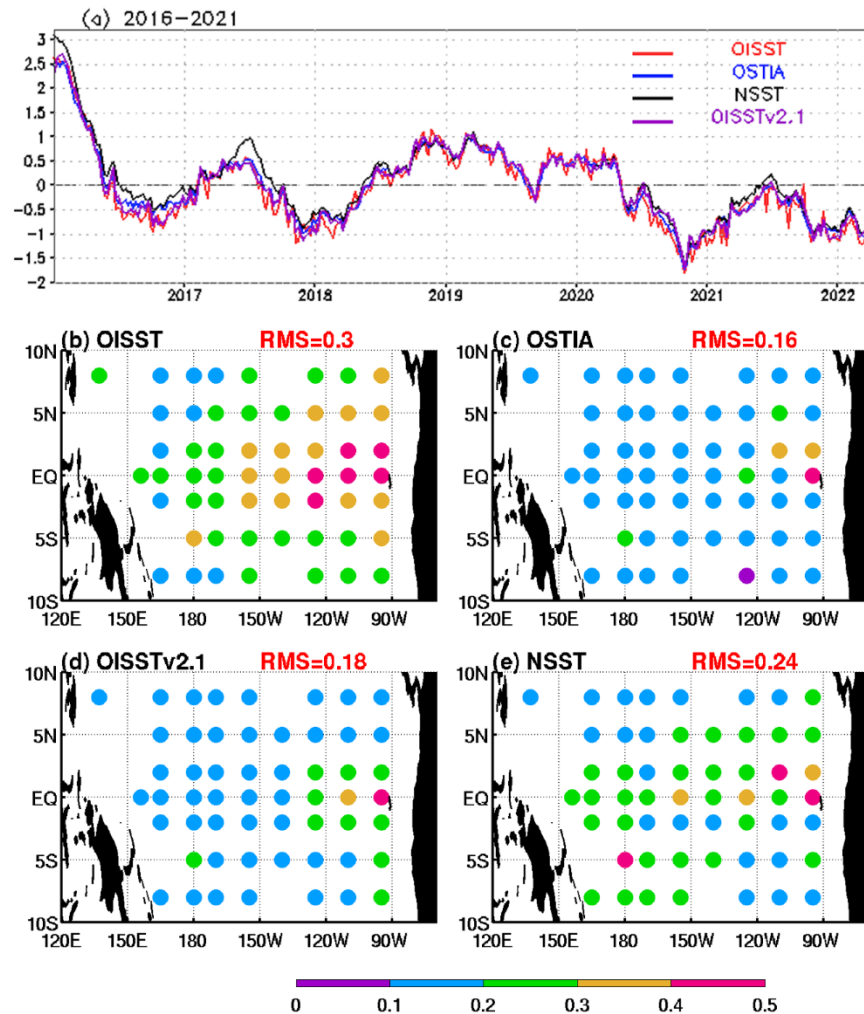


Fig. 1 (a) Time series of weekly Niño3.4 from various SST analysis products. (b)-(e) Root-mean-square (RMS) error of weekly SST for each SST analysis against the TAO/TRITON data in 2016-2021. Red texts represent basin average RMS and the color at TAO location indicates the amplitude of RMS error. Unit in °C.

products since the early 2000s. However, once the questionable fluctuations in OISST became evident the CPC ENSO team started to use OISSTv2.1 for weekly ENSO monitoring after October 11, 2021. This demonstrates the importance of knowing which SST analysis is closer to the observations. Figure 1 b-e shows the root-mean-square error (RMSE) of individual SST analysis against TAO moored buoy measurements. OISST has the largest basin average RMSE. In particular, RMSE of OISST is generally greater than 0.3°C in the eastern Pacific (east of 155°W). OSTIA and OISSTv2.1 have the smallest RMSE. The analysis suggests that weekly OISST has some issues in capturing weekly SST variation in the key regions of ENSO development.

For regions outside the tropical Pacific Ocean, the moored buoy data is very limited. To assess the uncertainty among SST products, the temporal evolution of global mean difference between individual SST analysis and OISSTv2.1 (Fig. 2a) was examined. Both ERSSTv5 and OISST global average SST are persistently warmer than OISSTv2.1 since 1982. The difference between OSTIA and OISSTv2.1 is the smallest among the SST analyses. We also compared the climatology differences. The results showed that OISST and ERSSTv5 are significantly warmer than OISSTv2.1 near the ITCZ and high latitudes of the southern hemisphere, while the difference between OSTIA and OISSTv2.1 is not statistically significant in most regions (not shown).

Spatial root-mean-square difference (RMSD) is plotted in Fig. 2b. ERSSTv5 has the largest global averaged RMSD from OISSTv2.1 ($\sim 0.5^{\circ}\text{C}$) and the difference remained stable in the last forty years. OSTIA has the smallest RMSD from OISSTv2.1 and RMSD drops from 0.3°C to 0.2°C after 2017.

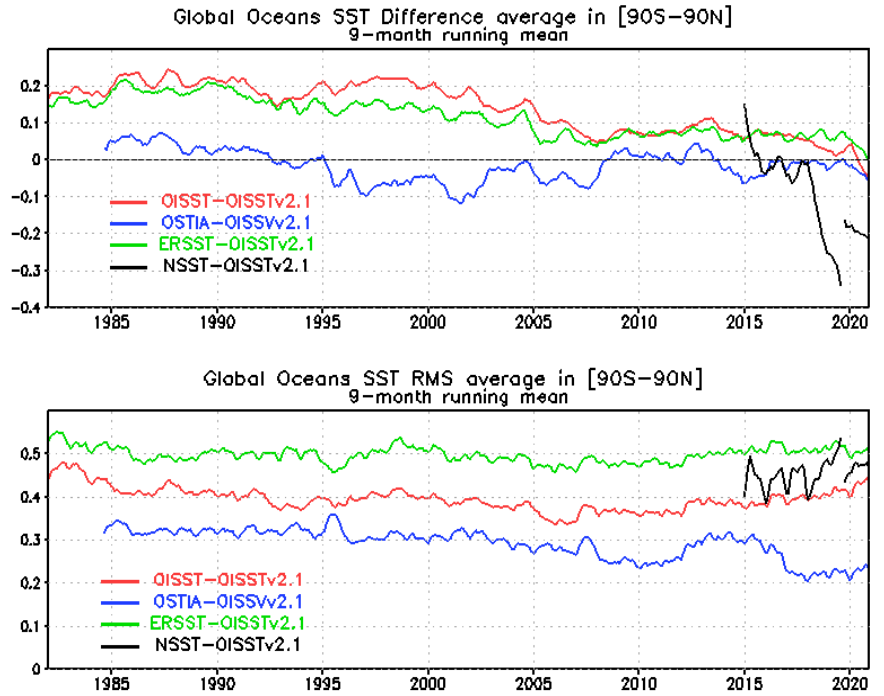


Fig. 2 Time series of (upper panel) global SST and (bottom panel) global root-mean-square difference between OISSTv2.1 and other SST data sets. Unit in $^{\circ}\text{C}$. Nine-month running mean is used to filter the time series.

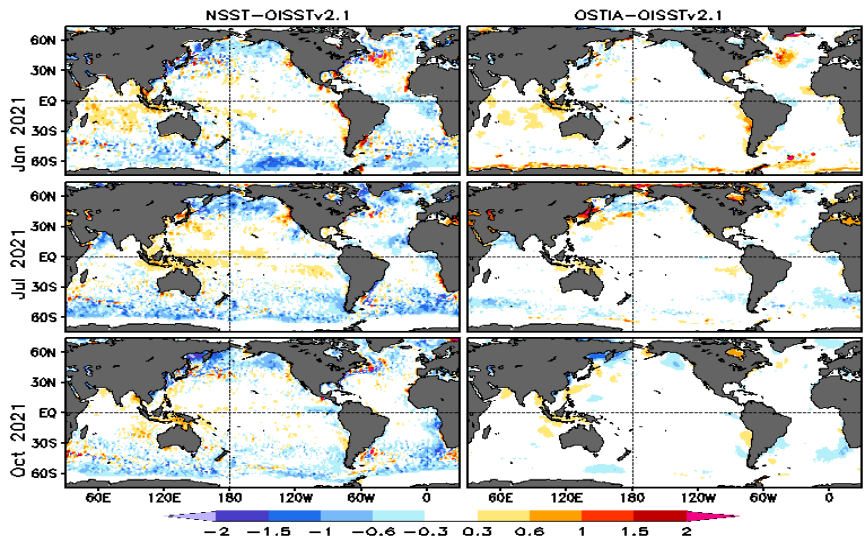


Fig. 3 Left panels: Monthly SST difference between NSST and OISST v2.1 (NSST minus OISST v2.1) at different months of 2021. Right panels are the same as left panels except for that between OSTIA and OISSTv2.1. Unit in $^{\circ}\text{C}$.

3.2 Impact of replacing SST nudging source on CFSR and CFSv2 forecasts

NSST has much shorter temporal coverage compared with other SST datasets, and it has been updated three times since the operational implementation in July 2017. These changes from the updates prohibit the use of NSST as a climate monitoring resource because such a dataset requires consistency over the historical record. Otherwise, it is uncertain whether a shift reflects a physical change or is due to a change in how the data was gathered and processed. The latest version was implemented in November 2019 and is currently used as the SST nudging source for the CFSR, NOAA's current operational climate reanalysis system. As shown in Fig. 2a, the global average NSST is significantly colder than the other three SST products. The global mean NSST bias in 2020 is about 0.2°C , which is about three times higher than the other SST products ($<0.05^{\circ}\text{C}$).

Figure 3 shows the biases of NSST and OSTIA in different months of 2020. Overall, NSST was significantly colder than OISSTv2.1 in the mid-to-high latitudes year-round, while biases in the tropics exhibit strong seasonal dependence. Large cold biases developed near the upwelling zone of eastern Pacific during the Northern Hemisphere summer and fall, while the bias is negligible during winter. Clear cold bias is also found near upwelling zones of the Atlantic Ocean in different months. Interestingly, although both NSST and OSTIA measure foundation temperature, OSTIA biases were negligible in most of the regions. It suggests that large cold biases in NSST is not purely due to the difference in SST type. One factor contributing to the cold bias in the tropical upwelling regions could be the usage of partly clear AVHRR radiance in NSST when it is not well detected by the satellite (personal communication with EMC, 2022). Further investigation is needed to clarify causes for the large difference between NSST and OISSTv2.1.

The CFSR ocean reanalysis was strongly nudged to daily OISSTv2.0 from January 1979 to January 2020 and then nudged to NSST starting from February 2020. How sensitive is the CFSR ocean reanalysis to such change in the source of SST nudging? To address this, the evolution of difference between CFSR and OISSTv2.1 is examined. Given the large difference between NSST and OISSTv2.1, one can expect a rapid shift in CFSR SST if the SST nudging source has a strong influence. Because the CFSR climatology is fixed, the bias shift in the total SST field should manifest in the SST anomaly field. To illustrate this, we selected the Pacific Ocean as an example. Figure 4a displays the temporal evolution of zonal average of difference. CFSR SST follows very closely with OISSTv2.1 prior February 2021 with very small bias, but rapidly developed a similar large cold bias to that of NSST after the nudging replacement (Fig. 4b). This demonstrates that the switch of SST analysis introduces a spurious negative anomaly in the mid-to-high latitudes. It is evident that the switch of nudging SST analysis source has a strong negative impact on temporal consistency of CFSR SST.

CFSR ocean reanalysis provides oceanic initial conditions to CFSv2. In this study, we only focus on the impact of the SST nudging source change on ENSO predictions. Both 2020-21 and 2021-22 were La Niña years. Whether CFSv2 forecasts share common features for these two events can help us understand the

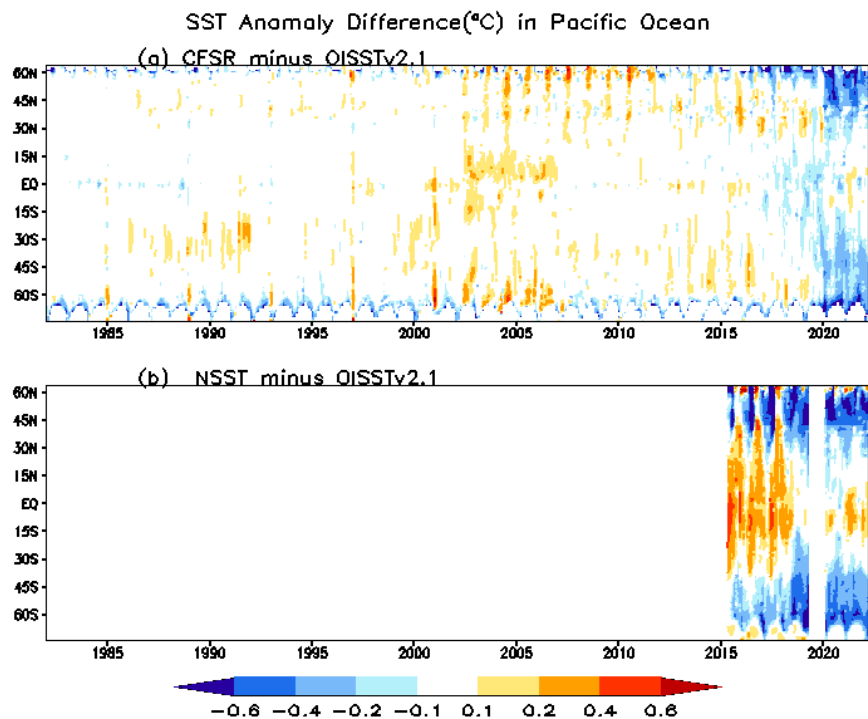


Fig. 4 Zonal average monthly SST anomaly bias of (a) CFSR, (b) NSST in the Pacific Ocean during 1982-2021. Unit in $^{\circ}\text{C}$.

impact. Figure 5 displays the Hovmöller diagram of NSST, CFSR and 0-month lead CFSv2 SST biases along the equator in the Pacific Ocean.

During 2020-2021, NSST biases exhibit similar seasonal variations where the center of the maximum negative bias ($<-0.8^{\circ}\text{C}$) peaks in September-November in the far eastern Pacific and, at other times, a warm bias ($>0.2^{\circ}\text{C}$) emerges in the western-central Pacific (Fig. 5a). CFSR bias was very small prior to the NSST nudging implementation, and then developed a similar pattern to NSST (Fig. 5b). The analysis reveals that the SST analysis replacement creates spurious negative (positive) SST signals in the eastern Pacific (warm pool) where air-sea interactions are very active for ENSO development. As expected, the CFSR bias leaves a clear footprint nearly immediately in the 0-lead CFSv2 SST forecast (Fig. 5c). However, the forecast bias evolution is slightly different from that of CFSR. Strong negative SST bias expands from the far eastern Pacific westward near the dateline from September to December. This was not observed in the CFSR bias. The warm bias in the 0-lead forecasts is also stronger than that of CFSR. It indicates that the spurious enhanced west-east SST gradient in CFSR can trigger positive feedback in the coupled system, and thereby intensify the original SST bias.

Figure 6 shows Niño3.4 predictions from CFSv2 over the course of 2020-2021. The most striking feature is that the CFSv2 forecasts of La Niña are far too cold for initializations during October to November. Forecasts overestimated the peak of La Niña by more than 1°C . This is consistent with the problems with the NSST/CFSR analyses shown in Fig. 5. Given that these months are short-lead forecasts, and are close to the mature phase of ENSO, we should expect higher prediction skill (Barnston and Tippett *et al.* 2013). That the forecasts are so strongly biased suggests that NSST error is so large that spurious strong positive feedbacks are induced in the coupled system, degrading the CFSv2 ENSO forecast skills.

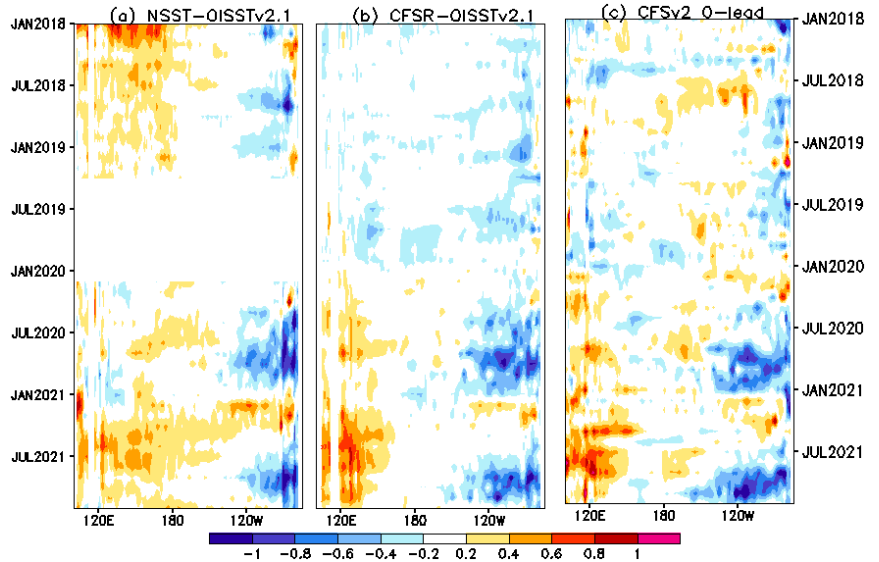


Fig. 5 Longitude-time diagrams of (a) NSST, (b) CFSR, and (c) 0-month lead CFSv2 SST anomaly bias at the equator during 2018-2021. Unit in $^{\circ}\text{C}$.

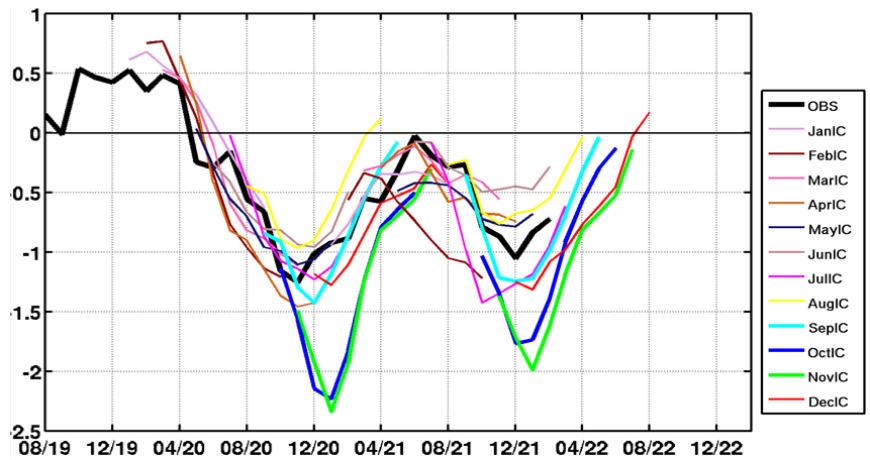


Fig. 6 Nine-month ensemble average predictions of Niño3.4 index from CFSv2 during 2020-2021. The black solid line is the observation based on OISSTv2.1. Color solid lines represent forecasts of individual forecasts at different initial months. Thick color solid lines represent months when CFSR has strong SST cold bias shown in Fig. 5.

4. Summary

SST analyses have been widely used for weather and climate forecasting and monitoring. However, on occasion, significant differences exist among SST products, which are caused by differences in the input data sources, sampling, analysis procedures, and other factors. The uncertainty creates challenges for interpreting climate monitoring and prediction products, which depend on the signs and amplitudes of SST anomalies. In this study, we assessed the fidelity of SST products currently employed in CPC operational products. We also investigated the impact of the SST nudging source change on the CFSR ocean data assimilation system and on CFSv2 SST forecasts. The main results are:

1. Uncertainty among SST data sets varies with time and location. OSTIA and OISSTv2.1 are consistent with each other, and these two data sets have smaller bias against the in situ TAO moored buoy than those of OISST and NSST.
2. NSST is significantly colder than OISSTv2.1 as well as other SST products in the mid-to-high latitudes. In the tropics, NSST biases are mainly located in the upwelling regions and Indo-Pacific warm pool regions with seasonally varying amplitude.
3. The replacement of NSST as the source of SST nudging in the analysis gives rise to the systematic errors in CFSR SST and anomalies after February 2020. CFSR biases subsequently leave clear footprints in the 0-lead CFSv2 forecast and beyond.
4. Large cold biases from the NSST in the eastern equatorial Pacific during the fall season contributed to overly negative predictions in Niño3.4 for DJF 2020-21 and 2021-22.

References

- Balmaseda, M. A., K. Mogensen, and A. T. Weaver, 2013: Evaluation of the ECMWF ocean reanalysis system ORAS4. *Quart. J. Roy. Meteor. Soc.*, **139**, 1132–1161, doi:10.1002/qj.2063.
- Barnston, A.G. and Tippet, M.K., 2013. Predictions of Nino3.4 SST in CFSv1 and CFSv2: A diagnostic comparison. *Clim. Dyn.*, **41**, 1615-1633.
- Behringer, D., and Y. Xue, 2004: Evaluation of the global ocean data assimilation system at NCEP: The Pacific Ocean. Preprints, *Eighth Symp. on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface*, Seattle, WA, Amer. Meteor. Soc., 6 pp. [Available online at <http://ams.confex.com/ams/pdfpapers/70720.pdf>]
- Donlon, C. J., M. Martin, J. Stark, J. Roberts-Jones, E. Fiedler, and W. Wimmer, 2012: The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system. *Remote Sens. Environ.*, **116**, 140–158, doi:10.1016/j.rse.2010.10.017.
- Huang, B., and Coauthors, 2017: Extended Reconstructed Sea Surface Temperature, version 5 (ERSSTv5): Upgrades, validations, and intercomparisons. *J. Climate*, **30**, 8179–8205, doi:10.1175/JCLI-D-16-0836.1.
- , C. Liu, V. Banzon, E. Freeman, G. Graham, B. Hankins, T. Smith, and H.-M. Zhang, 2020: Improvements of the Daily Optimum Interpolation Sea Surface Temperature (DOISST) version 2.1. *J. Climate*, **34**, 2923–2939, doi:10.1175/JCLI-D-20-0166.1.
- McPhaden, M. J., and Coauthors, 1998: The tropical ocean–global atmosphere (TOGA) observing system: A decade of progress. *J. Geophys. Res.*, **103**(C7), 14169–14240.
- NWS, 2020: EMC's Near Surface Sea Surface Temperature (NSST) [Available online at https://www.emc.ncep.noaa.gov/emc/pages/numerical_forecast_systems/sst.php]
- O'Carroll, A. G., and Coauthors, 2019: Observational needs of sea surface temperature. *Front. Mar. Sci.*, **6**, 420, doi:10.3389/fmars.2019.00420.
- Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang, 2002: An improved in situ and satellite SST analysis. *J. Climate*, **15**, 1609–1625.
- , T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, 2007: Daily high-resolution-blended analyses for sea surface temperature. *J. Climate*, **20**, 5473–5496, doi:10.1175/2007JCLI1824.1.
- Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057.

-
- , and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185–2208.
- Stark, J. D., C. J. Donlon, M. J. Martin, and M. E. McCulloch, 2007: OSTIA: An operational, high-resolution, real time, global sea surface temperature analysis system. *Oceans 2007 – Europe*, Aberdeen, UK, IEEE, doi: 10.1109/OCEANSE.2007.4302251.
- Zuo, H., M. Balmaseda, S. Tietsche, K. Mogensen, and M. Mayer, 2019: The ECMWF operational ensemble reanalysis-analysis system for ocean and sea-ice: A description of the system and assessment. *Ocean Sci. Discuss.*, doi:10.5194/os-2018-154.

Updating the CPC T2M Observational Verification Dataset and Impact on the Seasonal T2M GPRA

Mike Halpert, David DeWitt, and Jon Gottschalck
Climate Prediction Center, NOAA/NWS/NCEP, College Park, MD

The Climate Prediction Center (CPC) Government Performance and Results Act (GPRA) metric is the 48 month running mean Heidke Skill Score (HSS) of the favored category for the first lead seasonal 2-meter temperature (T2M) forecast. The forecasts are given as the probability of tercile categories: below, near, and above normal with the GPRA metric increasing (decreasing) if the score from 48 months ago was lower (higher) than the current seasonal score that replaces it in the running mean.

Since 2004, CPC has verified its seasonal temperature forecasts using real time data from River Forecast Centers (including from the Hydrometeorological Automated Data System – HADS) with a climatology based on the Cooperative Observer Program (COOP) network. Over the past 5-10 years, HADS has greatly increased its number of reporting stations, many of them out West at elevation, which has resulted in a cold bias in the CPC verification data set. Figure 1 shows the station density for this dataset on 6JAN2014 (left) and again 7 years later on 6JAN2021 (right). In particular, note the greater number of stations (generally at elevation) in California, Utah, Colorado, Montana, and Wyoming. Since some of the skill of seasonal temperature forecasts is linked to trends, this cold bias has often resulted in poorer scores than should have been obtained from unbiased verification of the forecasts.

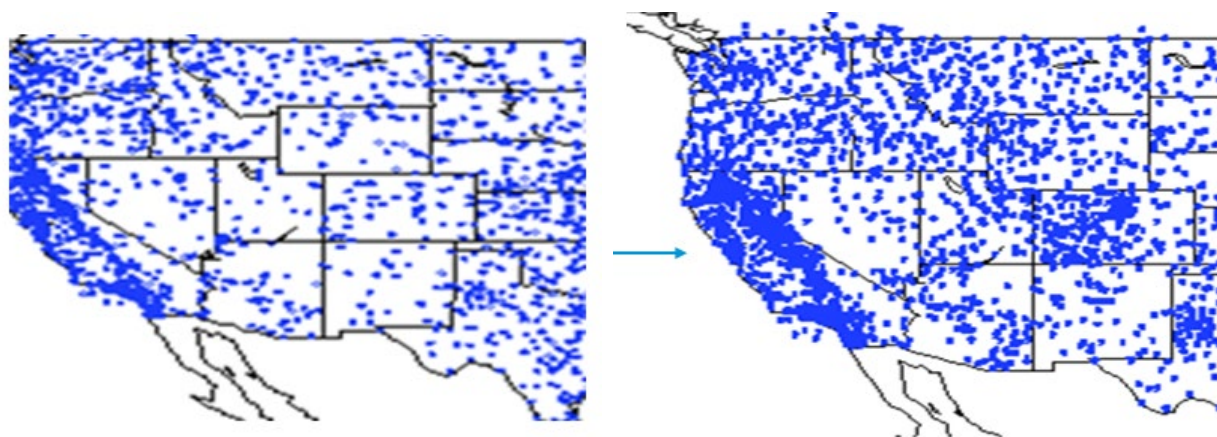


Fig. 1 Station locations for the historical database used to verify CPC seasonal outlooks on January 6, 2014 (left) and January 6, 2021 (right).

In order to more properly characterize a season into tercile categories, CPC has developed a new dataset which consists of Synoptic and METeorological Aerodrome Report (Metar) stations only, and which is consistent (same stations) throughout the historical and real time periods. Although containing fewer stations across the country (somewhere around 1500 stations), the use of a database with the same set of stations used in both the climatology and in the real time is highly desirable and ensures that the observed departures from normal (both above and below) will be the result of differences in the seasonal climate and not the result of the changing locations of stations used to verify the forecast.

A comparison of the individual HSS for the period 2011-20 (not shown) shows that scores are fairly consistent during the first half of the period. However, beginning in about 2016, scores obtained when verifying

against the new database often (but not always) exceed the HSS when verifying against the older database. Figure 2 shows the 48 month running mean of the first lead seasonal temperature outlook HSS calculated using the older, biased database (blue line) and the newer, consistent database (green line). As noted earlier, scores for the first half of the period are fairly similar, with the significant separation between the two lines developing fairly rapidly around 2017. Once established, the difference remains fairly consistent, indicating that the bias in HSS over the past 5 years has remained at about 5-7 points.

The change of datasets used for the verification of CPC's seasonal outlooks is expected to be implemented at the beginning of FY2023.

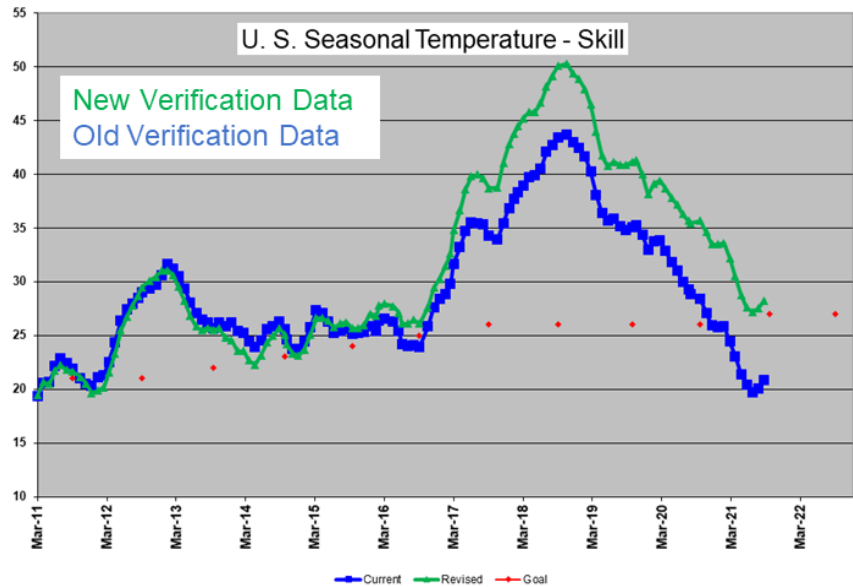


Fig. 2 Time series of the 48 month running mean of the Heidke Skill Score (HSS) of first lead seasonal temperature outlook scored using CPC's older verification dataset (blue line) and newer, consistent verification dataset (green line).

Understanding US Drought in Past 120+ Years

Yun Fan

Climate Prediction Center, NOAA/NWS/NCEP, College Park, MD

1. Introduction

Drought is one of major natural disasters which accounts for about one-third of all-natural disaster impacts. Those severe mega-droughts that cover large spatial areas and last long term periods can bring major disruption to earth's ecosystems and devastating disasters to the human society. Therefore, to better understand the attributions of these major droughts and to accurately predict them are important for science and protecting human society. In this study, a long-term (1895 to present) land surface hydrological dataset over the US was used to study long-term variations of the US droughts over the past 120+ years. Detailed studies were conducted to investigate if any drifting of the frequencies, magnitudes and locations of the land surface hydrological extremes happened over the past 12 decades, and how the land surface hydrological extremes responded to land surface hydrological forcing. Several land surface hydrological modelling experiments have been conducted to

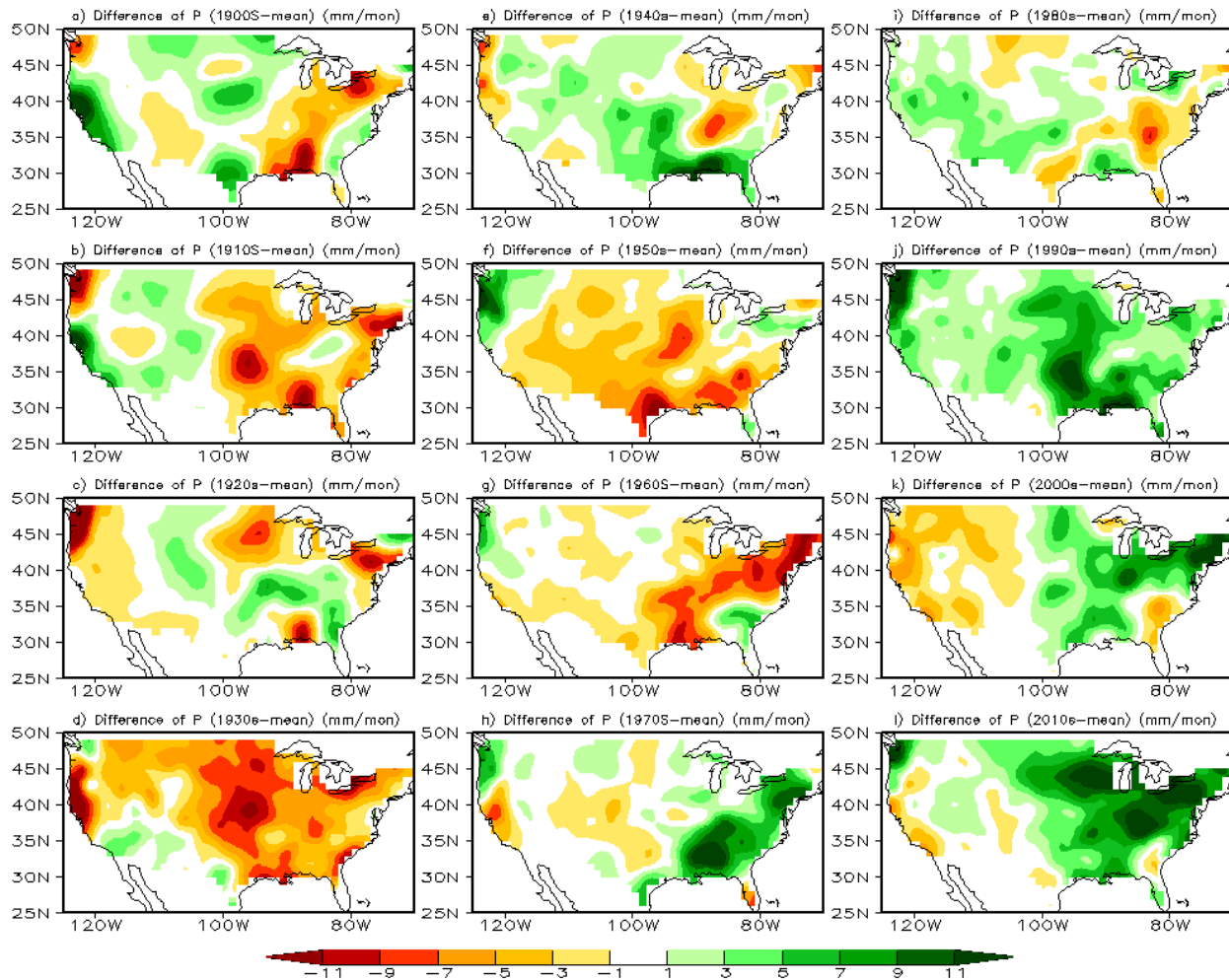


Fig. 1 Decadal variations of the observed precipitation for period 1901 to 2020 (120 years).

quantify the impacts of climate variations, such as long-term rainfall deficit and anomalous warm up temperatures, on the US drought. Further studies were zoomed into the Southwest, where frequently being hit by server droughts during the recent two decades, such as the Colorado River Basin (CRB).

2. Data and methodology

Land surface soil moisture responds to incoming precipitation and other atmospheric forcing like a low-pass filter, and it is an ideal and important physical variable for drought monitoring and studies. A long-term land surface hydrological dataset (from 1895 to present) was derived from the CPC leaky bucket land surface hydrological model (Huang *et al.* 1996), forced with the improved National Centers for Environmental Information (NCEI) monthly observed precipitation (P) and 2m surface air temperature (T2m) (well quality controlled) over the 344 US climate divisions (Vose *et al.* 2014).

There were some concerns (McRoberts *et al.* 2011) that variance of the observed data in the early period (*i.e.* 1895 ~ 1930) may be too low and would be “different” from the data collected in late period. Our analysis (Fan *et al.* 2017) shows that variabilities of the observed data in the early period look quite “normal” and there is no clear discontinuity or disruption when compared to those in the late period.

3. Decadal variations for past 120 years

The significant decadal spatial-temporal variations of the observed P and T2m for past 120 years are showed in Figs. 1 and 2, which depict evolution of the major decade-long wet (dry) and warm (cool) events across the

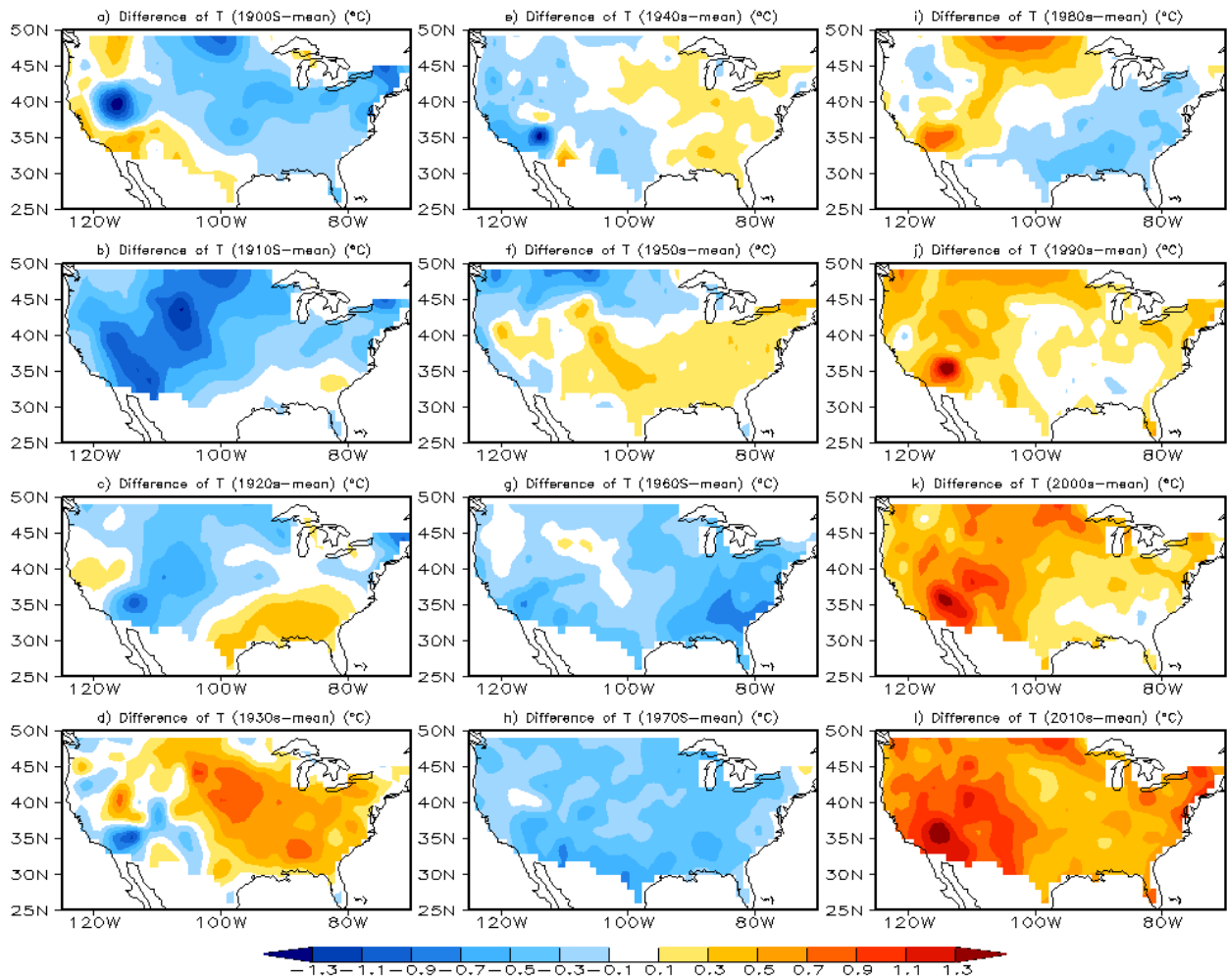


Fig. 2 Decadal variations of the observed 2m temperature for period 1901 to 2020 (120 years).

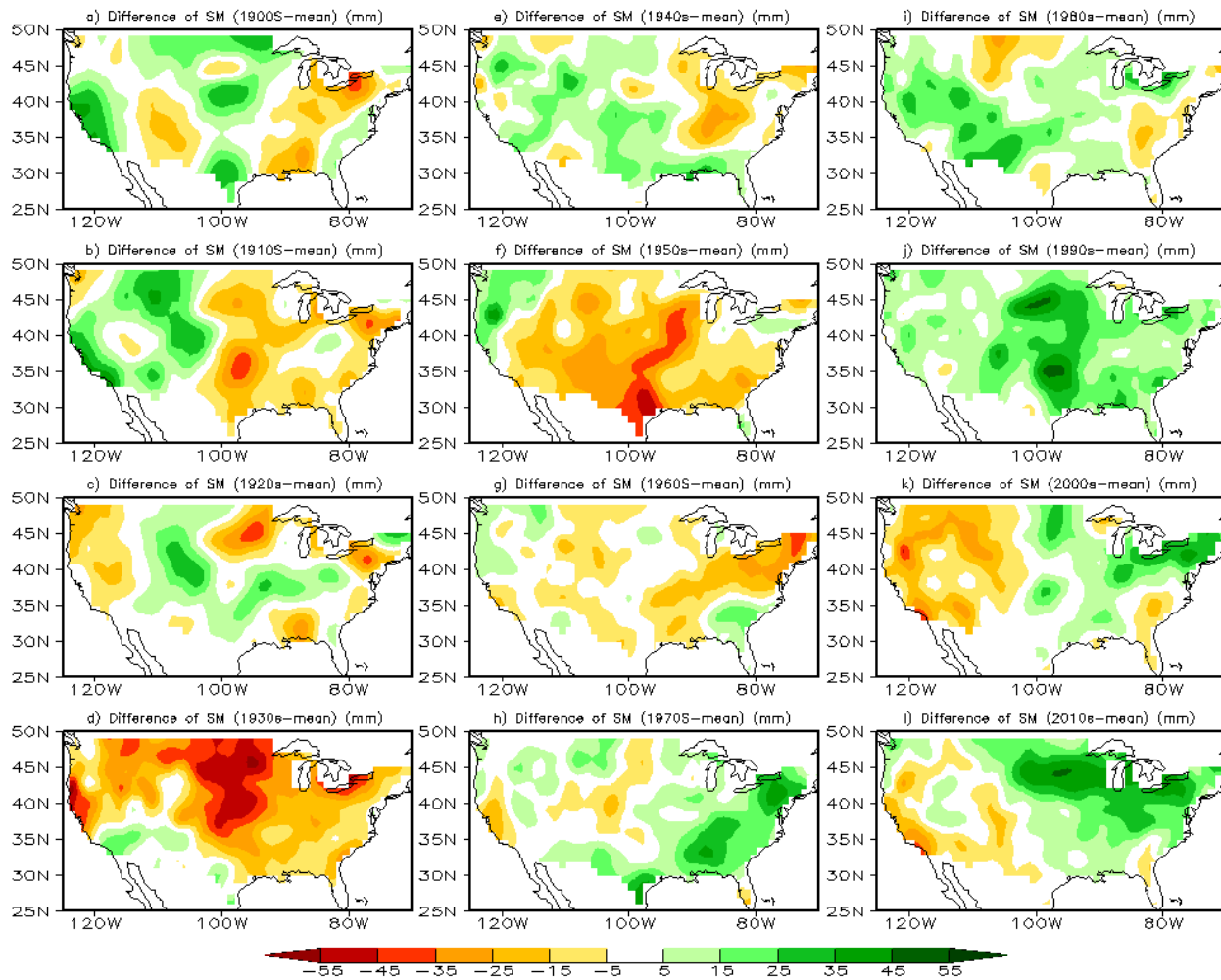


Fig. 3 Decadal variations of the simulated land surface soil moisture for period 1901 to 2020 (120 years).

CONUS. Another striking feature is the long-term trend in observations, such as much of the eastern CONUS trending to wetter while the whole CONUS trending to warmer, especially the southwestern CONUS.

The simulated decadal soil moisture variations (Fig. 3) reveal that the major land surface hydrological extreme events occurring in the past 12 decades were well captured in terms of spatial distribution, severity and duration, such as the decade-long major droughts (dust bowl era) in the 1930s in large parts of the US, multi-year severe droughts in the 1950s in the US (e.g. the southern Great Plain). The preliminary analysis shows that the decadal spatial-temporal variations of the derived soil moisture carry very similar patterns as the observed P, indicating that soil moisture variations are primarily driven by the variations of the observed P. However, the signature of the observed T2m modulation is also clearly indicated, such as anomalous warmer period in 1930s and 1950s enhancing already drier conditions, while relatively cool period in 1960s relieving the drier conditions in the eastern CONUS.

In order to further quantify the impacts of climate variations, such as long-term rainfall deficit and anomalous warm up temperatures, on the US drought, several modelling experiments have been conducted with different forcing (Control: observed daily P and T2m for 1979 to 2021; Exp1: 1991-2020 observed daily P climatology and observed daily T2m; Exp2: observed daily P and 1991-2022 observed daily T2m climatology; Exp3: 1991-2020 observed daily P and T2m climatology). The results suggested that observed P inter-annual variability may account for about 85~90% inter-annual variability for the simulated soil moisture, while observed T2m inter-annual variability may contribute another 10~15% inter-annual variability for the simulated soil moisture.

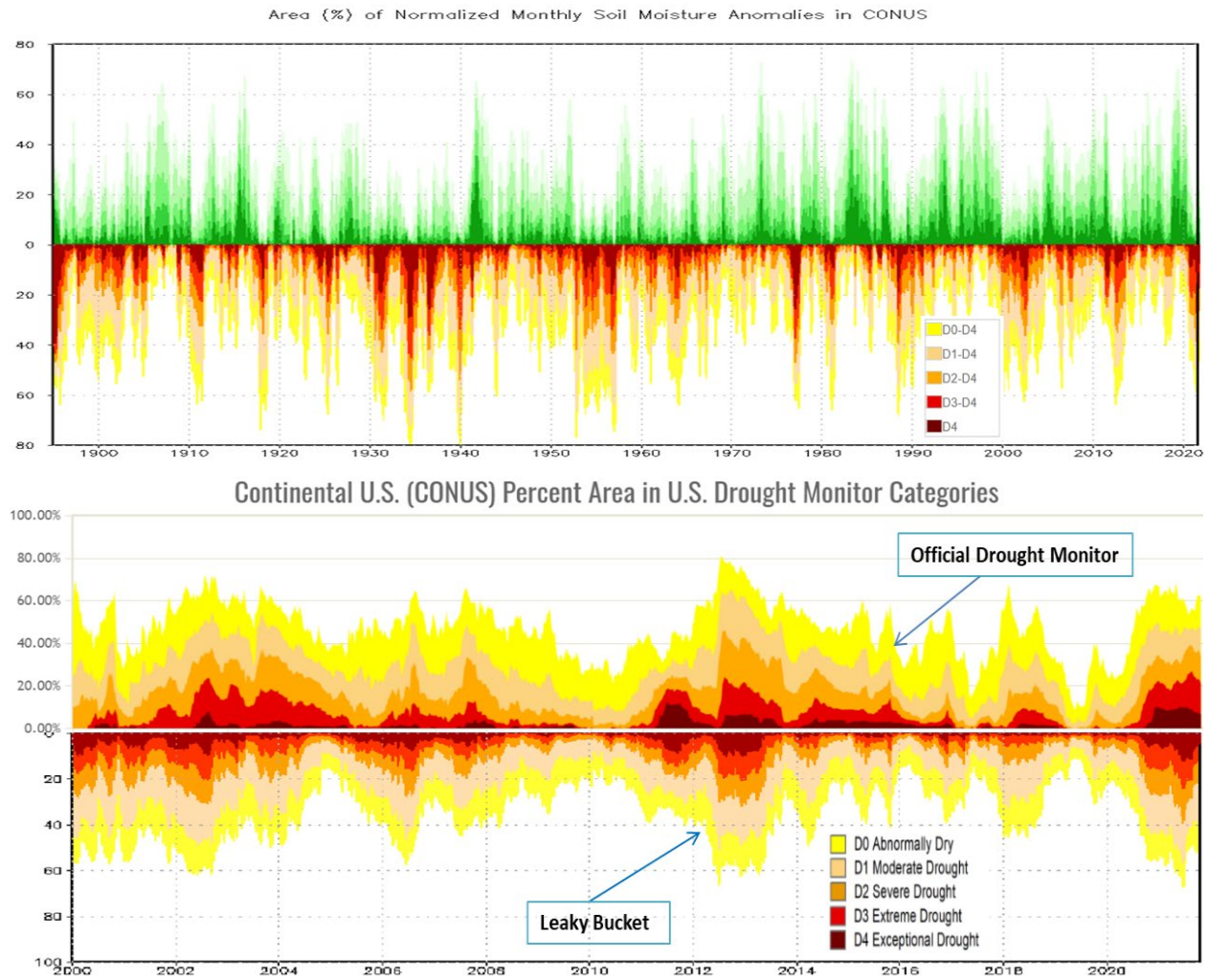


Fig. 4 Time series of areas (%) covered by classified (D0 to D4; yellow to dark red) US droughts: (1) upper panel from CPC Leaky Bucket model for period of 1895-2021 ; (2) lower panel for the official US drought monitor against CPC Leaky Bucket model for period of 2000-2021.

4. Mutual validation against US drought monitor (USDM)

Figure 4 provides a centennial long categorized drought area (%) index for the CONUS (upper panel: D0 to D4 with yellow to dark red bars) based on the simulated soil moisture data, while the counterpart (green shaded bars) can be viewed as wet extreme events (*e.g.* flood). In terms of the CONUS wide during the past 120+ years, the worst drought periods happened in 1930s (dust bowl era) with about 4 major extreme events spared across the decade and in 1950s with two to three more major droughts clustered in middle of the decade.

To verify the quality of the used simulated soil moisture data, traditional way is directly against to observation. Other indirect ways can be done by validation with other drought indicators. Here the weekly official USDM was picked up to perform a mutual validation. The lower-panel of Fig. 4 shows very good agreement (temporal mapping) between drought indices from the USDM and CPC Leaky Bucket modeled soil moisture. Figure 5 further displays a nice spatial mapping between the two drought indicators.

5. Drought in the Western CONUS and Colorado River Basin (CRB)

Here more detailed studies were conducted to focus on the US West, where widespread droughts frequently hit the region in recent decades and water supply routinely becomes a major issue for the people living in the

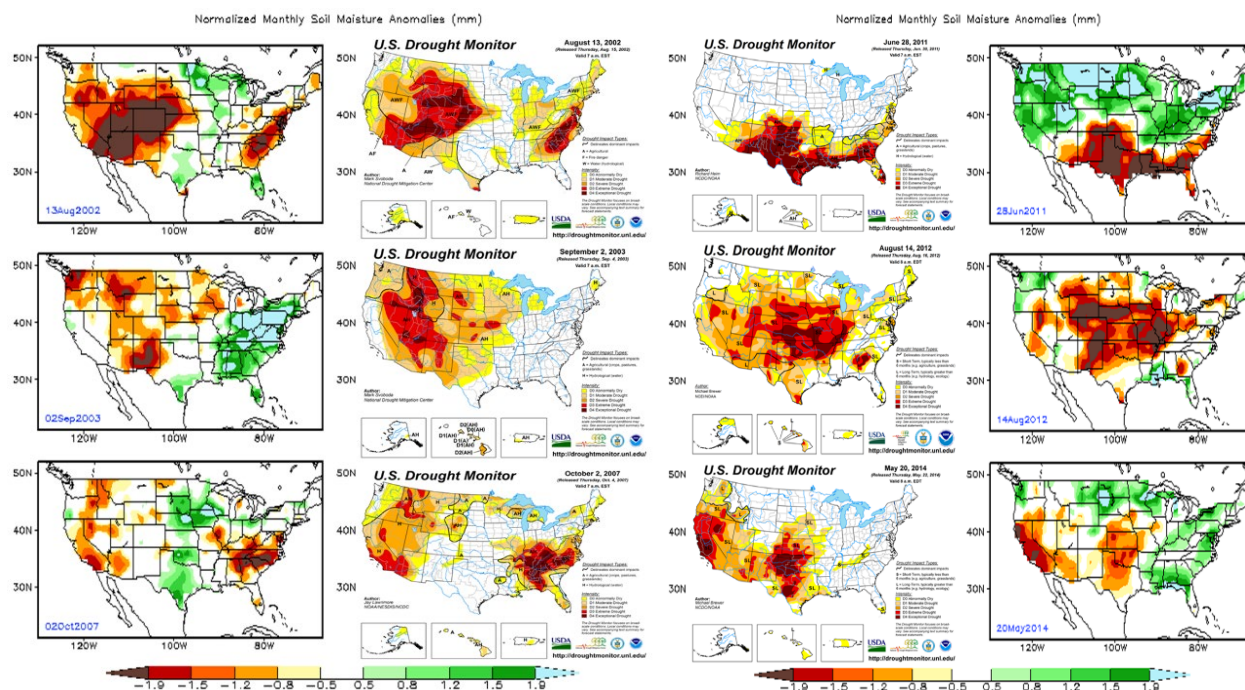


Fig. 5 Selected 6 US drought events from the official US drought monitor and classified CPC Leaky Bucket model.



Fig. 6 EOF analysis from simulated soil moisture in the western CONUS for period 1895 to 2021.

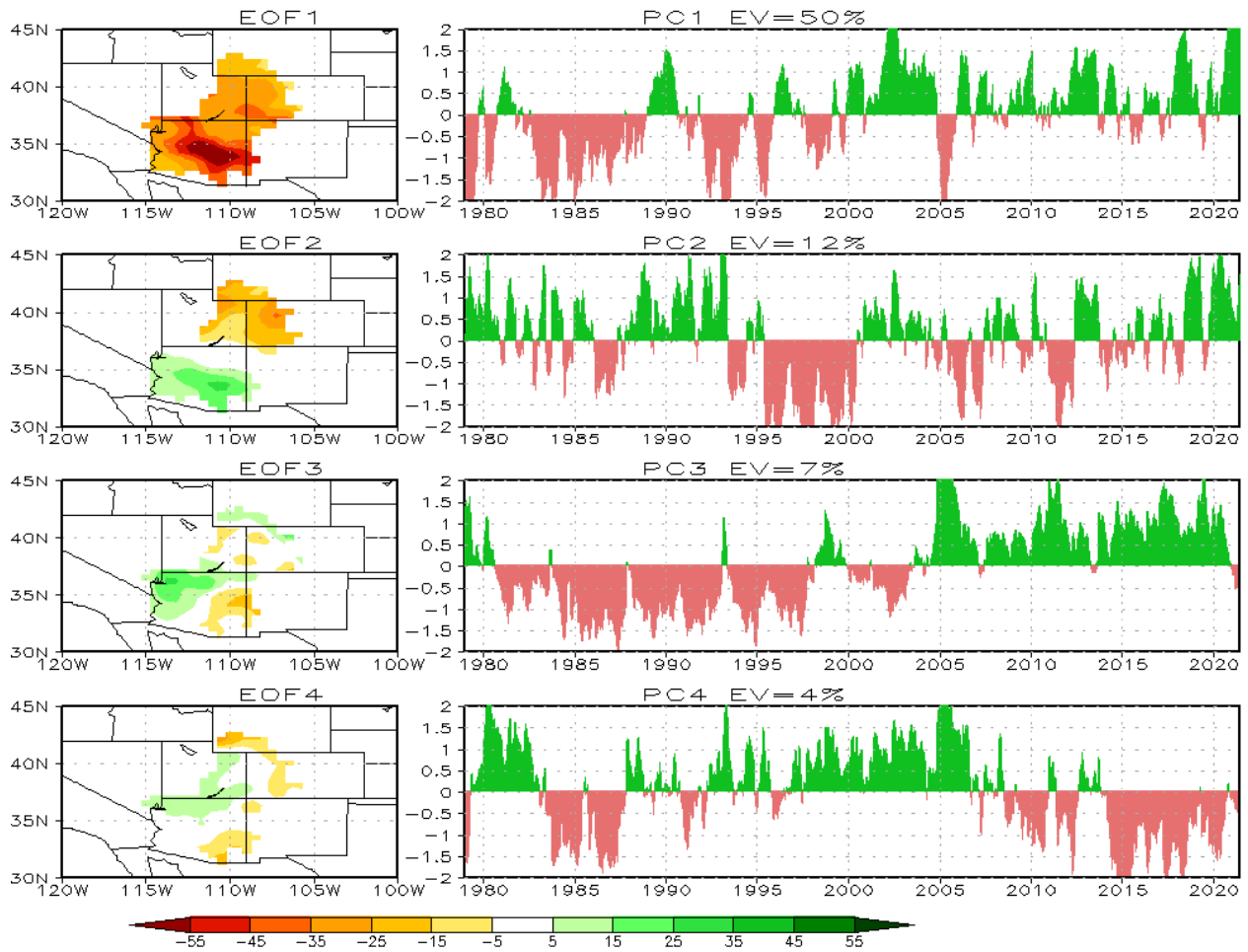


Fig. 7 EOF analysis from simulated daily soil moisture in the CRB for period 1979 to 2021.

West. The EOF analysis (Fig. 6) depicts very clear decadal and inter-annual variability. The dominant first 4 leading EOFs with relatively simple and large-scale patterns can explain nearly 69% total soil moisture variance.

The studies were further zoomed into the Southwest, where is the core region associated with warm-up trend and frequently hit by severe droughts during the recent decades, such as the Colorado River Basin (CRB). The daily EOF analysis for period of 1979 to 2021 shows that the even more dominant first 4 leading EOF modes could account for up-to 73% (94% from the monthly data) total soil moisture variance (Fig. 7). Interestingly, the second EOF mode shows the naturally divided geological Lower Basin and Upper Basin and they vary in opposite direction. The most striking low-frequency feature is the leading EOF modes display very significant inter-decadal variations and lead to a clearly long-term dry trend. Decreasing water supply and increasing water use (also partly due to population increase) will make the water demanding crises even worse for more than 40 million people who live in the CRB and rely on its water. If this trend continues, better water management becomes more important for the region's future.

6. Conclusions

The preliminary studies reveal the significant decadal spatial-temporal variations in the observed P and T2m across the CONUS for past 120+ years and depict the striking features of the long-term trend in observations: much of the eastern CONUS trending to wetter while the whole CONUS trending to warmer, especially the southwestern CONUS. The severe precipitation anomalies modulated by long-term anomalous temperatures were the immediate reasons for those major hydrological extreme events.

With the studies further zoomed into the US West, especially the Southwest such as the CRB, the more simple soil moisture structures emerged with the most striking low-frequency features showing a clearly long-

term drier trend. Decreasing water supply and increasing water use will make the water demanding crises even worse. However, from forecast point of view, the CRB soil moisture variability is strongly dominated by very few simple and low-frequency soil moisture anomaly patterns and may indicate drought in the CRB potentially more predictable.

References

- Huang, J., H. M. van den Dool, and K. P. Georgakakos, 1996: Analysis of model-calculated soil moisture over the United States (1931-1993) and applications to long-range temperature forecasts. *J. Climate*, **9**, 1350-1362.
- Vose, R. S., S. Applequist, I. Durre, M. J. Menne, C. N. Williams, C. Fenimore, K. Gleason, D. Arndt, 2014: Improved historical temperature and precipitation time series for U.S. climate divisions. *J. Appl. Meteorol. Climatol.*, **53**, 1232–1251, doi: 10.1175/JAMC-D-13-0248.1.
- McRoberts, D. B., and J. W. Nielsen-Gammon, 2011: A new homogenized climate division precipitation dataset for analysis of climate variability and climate change. *J. Appl. Meteorol. Climatol.*, **50**, 1187–1199, doi: 10.1175/2010JAMC2626.1
- Fan, Y. and H. van den Dool, 2017: Long-term variation of US land surface hydrological extremes. *Climate Prediction S&T Digest*, 42nd NOAA Climate Diagnostics and Prediction Workshop, Norman, OK, DOC/NOAA, 130-135, doi:10.7289/V5/CDPW-NWS-42nd-2018.

Developing an Experimental Week 3-4 Storm Track Outlook over North Pacific, North America, and North Atlantic

Yutong Pan,^{1,2} Wanqiu Wang,¹ Edmund K.M. Chang,³ Hui Wang,¹ David DeWitt¹

¹*Climate Prediction Center, NOAA/NWS/NCEP, College Park, MD*

²*Earth Resources Technology, Inc., Laurel, MD*

³*School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY*

1. Introduction

Extratropical storms are accompanied by heavy precipitation and damaging winds, causing high impacts on public safety, property, and economy. Reliable predictions of enhanced and suppressed storm activities over mid- and high-latitude regions, including Alaska, are in great demand and of critical importance. The goal of this study is to develop a forecast tool for week 3-4 storminess by extending the week-2 forecast system (Pan *et al.* 2021), which already exists at the NOAA Climate Prediction Center (CPC), to week 3-4.

Storms are detected and tracked using 6-hourly sea level pressure (SLP) data from the real-time GEFSv12 35-day forecasts and a storm-tracking algorithm developed by Serreze (1995). The week 3-4 outlook products include storm activities (storm tracks and track density, storm intensity and duration), SLP day-to-day variance, storm-associated precipitation, SLP and 10-m wind speed over North Pacific, North America, and North Atlantic, derived from the GEFSv12 week 3-4 forecasts for both total and anomaly fields. In addition to deterministic forecasts (ensemble mean forecast), the outlooks also provide probabilistic forecasts of precipitation and 10-m wind speed exceeding 75% and 90% percentiles, and storm intensity lower than 990, 980, 970, and 960 hPa. Verifications for the real-time week 3-4 forecasts are also conducted using the NCEP Climate Forecast System Reanalysis (CFSR). The week 3-4 storminess outlook is updated daily.

2. Data and methodology

2.1 Data

In this study, GEFSv12 6-hourly 35-day forecasts (31 ensemble members) are utilized for the week 3-4 storminess outlooks. The variables used include SLP, precipitation, and 10-m winds. We have also used the 21-year (1999-2019) GEFSv12 hindcast dataset to derive model climatology and assess the forecast skill, and the CFSR data as observations for the forecast verification and skill assessment.

2.2 Methodology

Similar to the week-2 outlook (Pan *et al.* 2021), the week 3-4 storm detecting and tracking are based on the algorithm developed by Serreze (1995), with the following criteria:

- Using 6-hourly SLP data on the $2.5^\circ \times 2.5^\circ$ grid
- Storm center SLP ≤ 1000 hPa
- Storm center SLP at least 1 hPa lower than surrounding grid points
- Maximum distance a storm can travel is 800 km/6 hour

Storm track density is defined as total number of storm centers within a 250-km radius for each grid point divided by ensemble members. Storm intensity (center SLP) denotes the mean storm center pressure within a 250-km radius for each grid point. Storm duration is the mean lifetime of storms passing through a domain of 250-km radius for each grid point.

3. Week 3-4 storm track outlook, verification and skill evaluation

The real-time week 3-4 forecast products are available on the following forecast website, with a daily update (<https://ftp.cpc.ncep.noaa.gov/hwang/YP/week2/>). Figure 1 shows an example of the week 3-4 outlook issued on January 30, 2022, for the 14 days from February 13 to 27, 2022, including storm tracks, storm track density, storm intensity and duration. The left panels in Fig. 1 are the total fields and the right panels are the anomaly fields. Forecasts for other variables, as well as the sub-regional maps can be found in the forecast webpage.

The verification of the week 3-4 forecast against the CFSR is done when the CFSR data are available for the forecast target weeks. Therefore, there is a 28-day delay for the real-time verification. Figure 2 shows both forecast and verification, in which left panels are the model forecast (same as in Fig. 1) and the right panels are the CFSR verification of storm tracks, storm track density, storm intensity and duration for the forecast shown in both Fig. 1 and Fig. 2 (left).

The outlook tool is assessed using the 21-year (1999-2019) GEFSv12 hindcast data. The forecast skill is determined by the anomaly correlation (AC) between the forecasts and the CFSR during the GEFSv12 hindcast period. Figures 3-4 display the AC skills of week 3-4 storm track density, SLP day-to-day variance, precipitation, and SLP, respectively, for January and July. The results indicate a relatively high skill for the week 3-4 storm track density (Fig. 3 left) in the climatological storm-active regions (circled by dash curves). The week 3-4 forecasts of SLP day-to-day variance (Fig. 3 right), precipitation and SLP (Fig. 4) show higher skills than the week 3-4 storm track density in both January and July. Overall, the AC skills in the winter month (January) are higher than in the summer month (July).

4. Summary

A real-time GEFS-based week 3-4 storminess outlook tool was developed at NOAA CPC, with a daily update and the CFSR verification. Anomaly correlations of week 3-4 storm track density, SLP day-to-day variance, precipitation and SLP between the GEFSv12

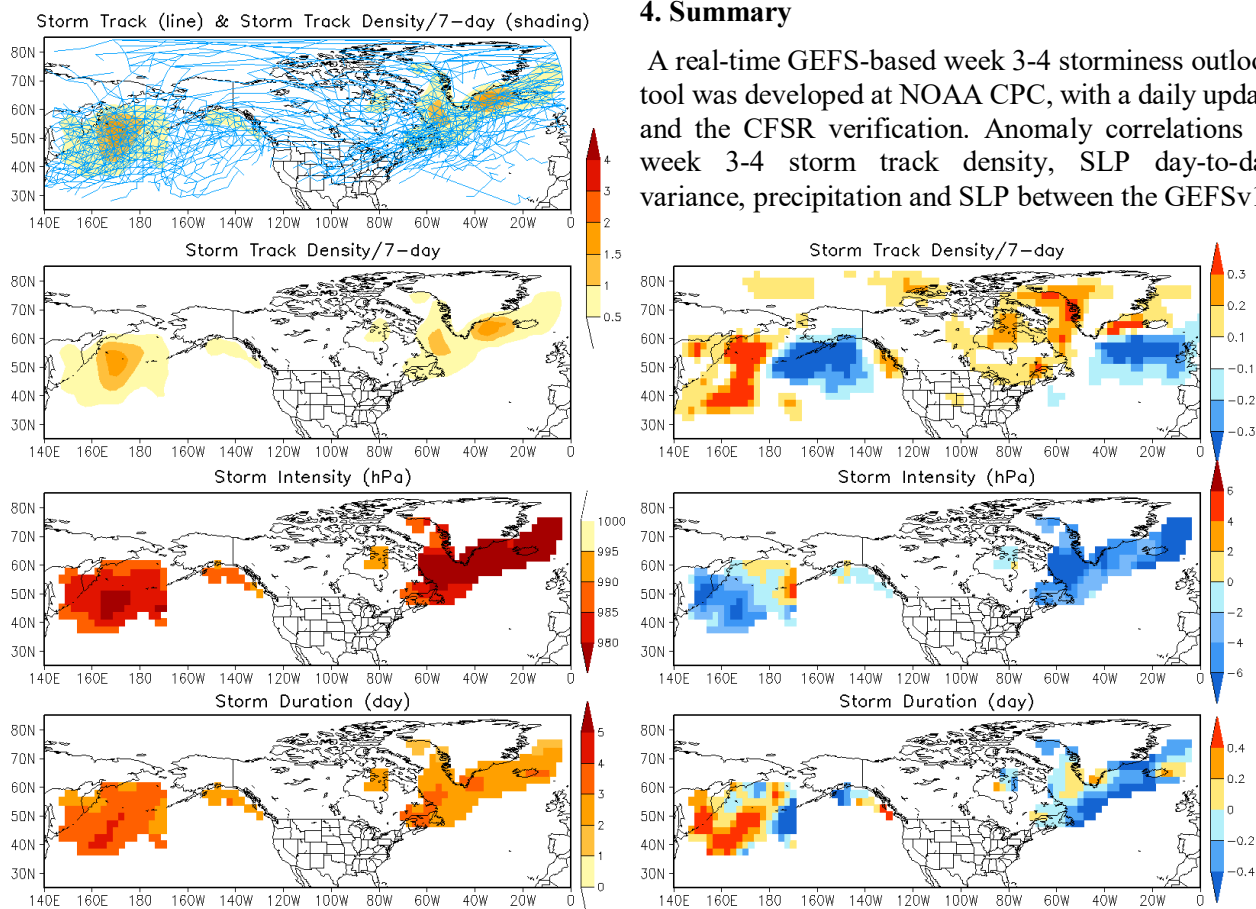


Fig. 1. GEFSv12-based week 3-4 forecasts of storm tracks, track density, storm intensity and duration for both total (left) and anomaly fields (right). The forecast date was January 30, 2022 for week 3-4 from February 13 to February 27, 2022.

21-year hindcasts and the CFSR indicate a certain level of skills for storm track density over the storm-active regions, and better skills for SLP day-to-day variance, precipitation and SLP. Overall, a winter month has a higher skill than a summer month.

To improve the forecast skill, the week 3–4 storminess outlooks using Climate Forecast System version 2 (CFSv2) 45-day operational forecast and GEFSv12+CFSv2 combined forecast have also been developed and implemented in real time, which are also available on the forecast website.

Acknowledgments. We thank Michael Staudenmaier, Eugene Petrescu, Ray Wolf, Joseph Dellicarpini, Christopher Buonanno, Eric Lau for their helpful comments and suggestions on the development of the storm-track outlook products.

References

- Pan, Y., W. Wang, H. Wang, and D. DeWitt, 2021: Developing an experimental week-2 storm track outlook over North Western Hemisphere. *Climate Prediction S&T Digest*, 45th NOAA Climate Diagnostics and Prediction Workshop, Virtual Online, DOC/NOAA, 123–126, doi: 10.25923/tpfe-4n87.
- Serreze, M. C., 1995: Climatological aspects of cyclone development and decay in the Arctic. *Atmos.–Ocean*, **33**, 1–23.

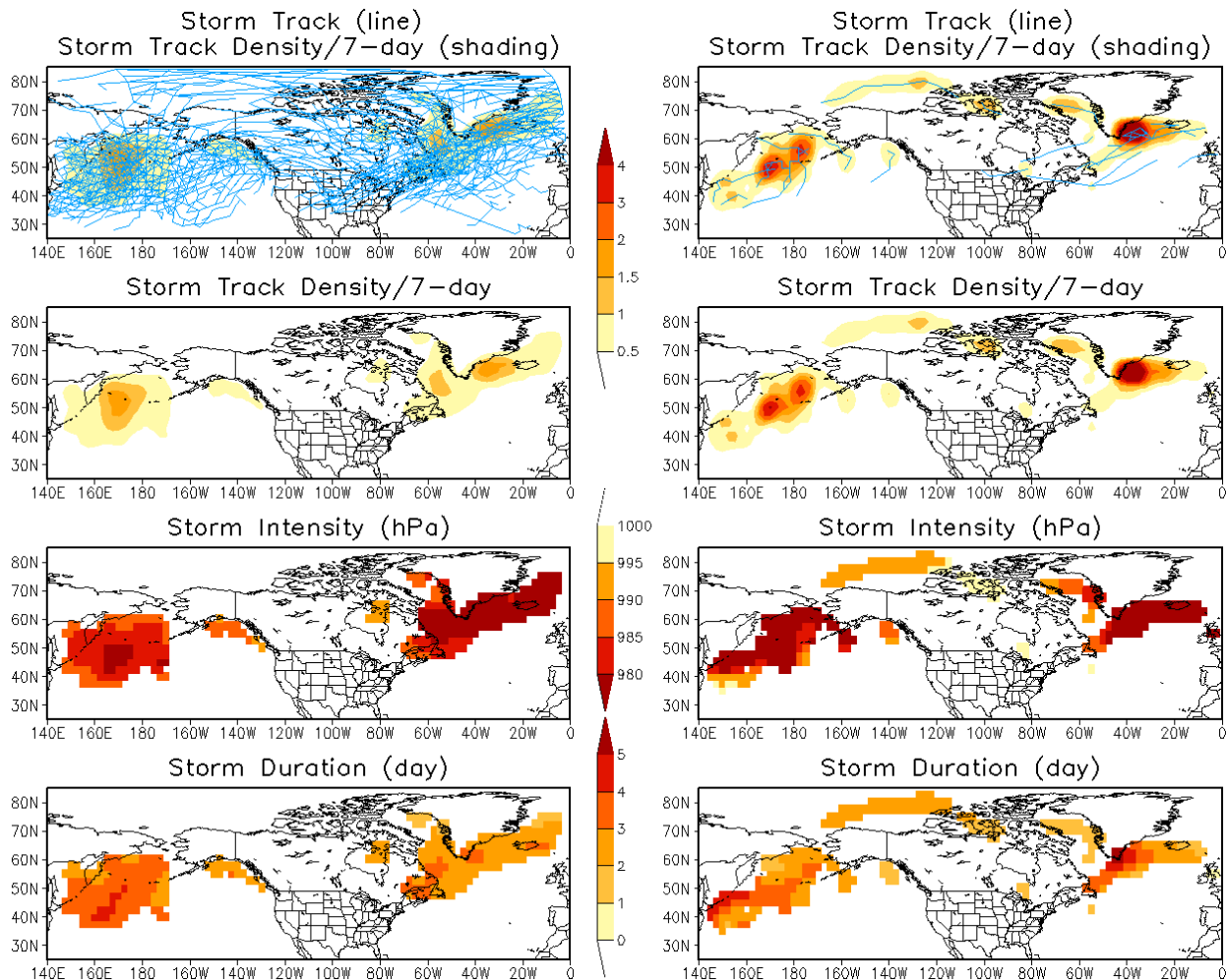


Fig. 2a Verification (right) of GEFSv12 week 3–4 forecast (left) for storm tracks, track density, storm intensity and duration with total fields. The forecasts were issued on January 30, 2022 for week 3–4 from February 13 to February 27, 2022.

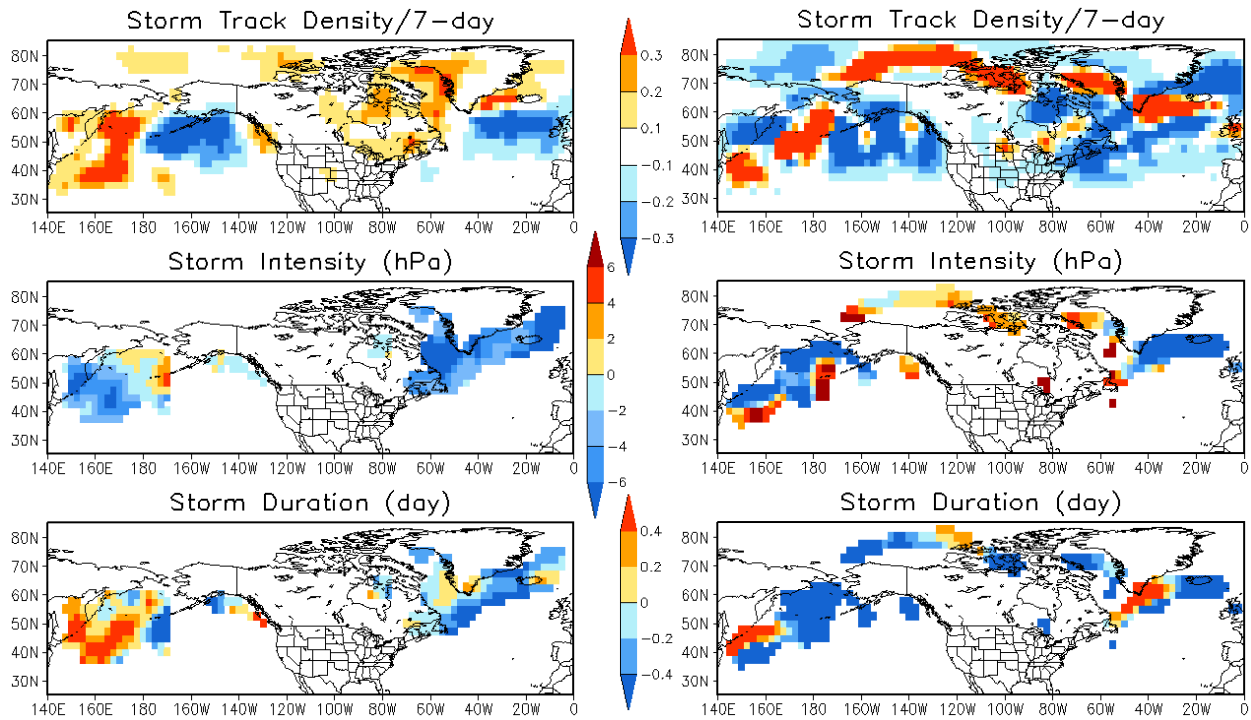


Fig. 2b Same as Fig. 2a but with anomaly fields.

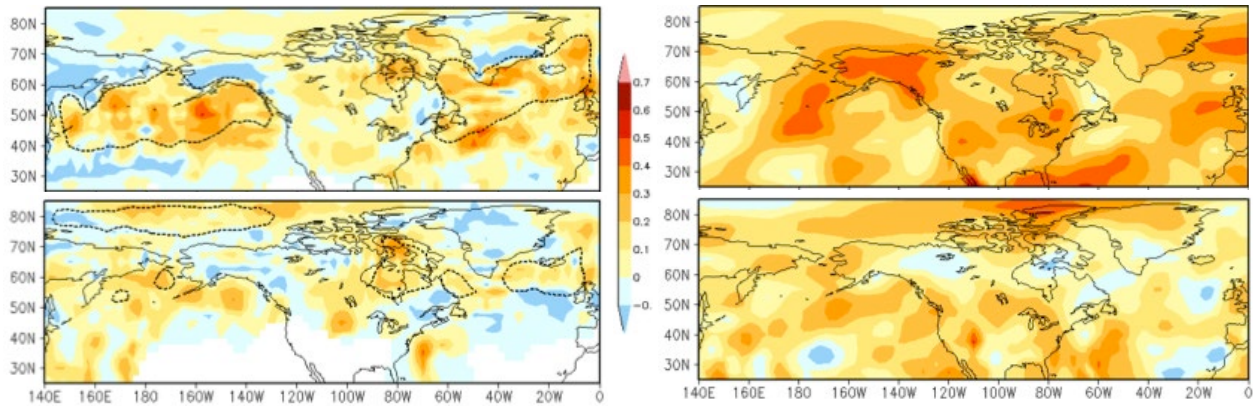


Fig. 3 Anomaly correlation of week 3-4 storm track density (left) and SLP day-to-day variance (right) between the GEFSv12 hindcasts and CFSR over the 21-year (1999–2019) hindcast period for January (top) and July (bottom). Areas circled by dashed curves represent climatological storm-active regions.

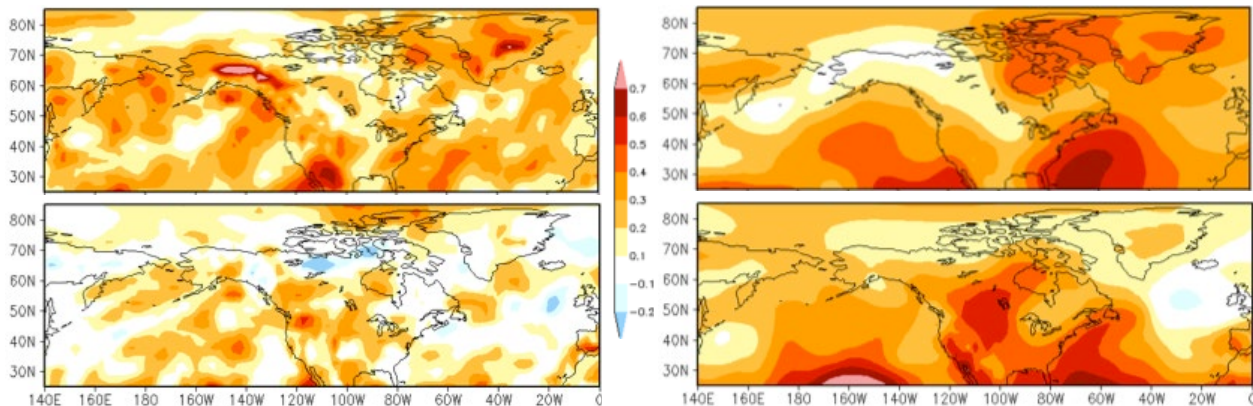


Fig. 4 Anomaly correlation of week 3-4 precipitation (left) and SLP (right) between the GEFSv12 hindcasts and CFSR over the 21-year hindcast period for January (top) and July (bottom).

CPC's Climate Assessment Database Version 2

Melissa Ou, Tom Collow, Donald Garrett

Climate Prediction Center, NOAA/NWS/NCEP, College Park, MD

1. Introduction

The Climate Assessment Database (CADB) is a suite of station based monitoring analyses products that is routinely produced by the Climate Prediction Center (CPC). It consists primarily of daily, weekly, monthly, and annual estimates of various weather parameters, based on hourly data from the Global Telecommunication System (GTS). These datasets were created based on the need for station observations updated in a more timely manner. The National Centers for Environmental Information (NCEI) is the national leading authority for environmental data, with complex algorithms for quality control (QC) and calculations of observations. However, these datasets take more time to produce due the complex nature of the production algorithms. The CADB has lighter QC performed, enabling quicker processing and production of datasets, and thus near real-time data availability for users' various needs. The original CADB software (CADBv1) was created in the early 1980s with a backend direct access database. After years of documenting CADBv1, code development, and benchmarking new results with CADBv1, CPC transitioned to the new version of CADB software (CADBv2) at the beginning of 2020, including the upgrade to using a relational database for the backend storage and web application support. Additionally, the previously static temperature and precipitation time series charts based on CADB data have been replaced with a dynamic web application.

CADBv2 continues to be occasionally upgraded and improved as needed based on upstream data changes and feedback from users. There is a broad variety of users and stakeholders of the CADB, including those internal to CPC that produce other gridded datasets and international products, external government agencies such as the Joint Agricultural Weather Facility (including usage in crop models) and U.S. Department of Agriculture (USDA), private companies, and general public users. This article will describe the available CADB datasets, changes from version 1, as well as the related upgraded base period station normals, station meta-data, and new CADB time series web application.

2. About the CADBv2

The CADB refers to a framework that consists of a suite of station based datasets that have summarized values of weather parameters over various timescales based on raw hourly station observation reports over the globe. Multiple pieces of software support the routine production of the CADB, including the summarized station observations that are estimated on daily, weekly, monthly, and seasonal timescales, which are available at <https://www.cpc.ncep.noaa.gov/products/cadb> or via FTP at https://ftp.cpc.ncep.noaa.gov/cadb_v2.

It is important to note that the CADB is considered a preliminary dataset, meaning that there is "light quality control (QC)" used to assess and summarize the data. The CADB is not intended to be a high quality controlled dataset with more complex post-processing, such as using nearest neighbor check techniques, other data sources for adjustments, and assessing climatological data for value bounds. Other datasets, such as those from NCEI or the CPC unified gauge dataset (which ingests primarily the CADB as well as standard hydrometeorological exchange format (SHEF) data) utilize more complex post-processing and QC techniques. The benefit of having lighter QC used for the CADB is that the data is available in a more timely manner, which is necessary to support many real-time operational needs, whereas datasets with more complex QC procedures, such as from NCEI, take more time to process and release updated data.

The original version of CADB software, CADBv1, was created in the 1980s in Fortran77. It was deemed necessary by CPC to rewrite the software in a more modern software language for multiple reasons including:

1) the original code was sparsely documented, thus the details of the methodology and calculations were not well understood, precluding upgrades from being done easily and making the software hard to maintain, 2) suspicion that there may be some errors in the code and/or approaches, 3) changes in upstream data formats required updates to be made in the CADB software, 4) many more stations of data are available globally now compared to the 1980s but the framework of v1 did not easily allow for expansion of station coverage, and 5) need for having software in a more modern language enabling more staff to be able to maintain and upgrade it over time.

Another related upgrade was the conversion from storing data in a binary formatted direct access file to a relational database. Updating data in the direct access file was not transparent and could easily lead to errors, making it hard to update, maintain, in addition to being difficult to view the data itself. In contrast, using a relational database to store and extract CADB data is much easier to maintain and was also a needed technology to support a new CADB temperature and precipitation time series web application, which replaced the static CADB time series chart graphics.

3. Data

Hourly station observation data from the Global Telecommunication System (GTS) is the source of input to the CADB. This hourly data is a global network of transmitted meteorological data from various sources, including manual and automated measurements of various weather conditions. The raw data reports are comprised of both surface synoptic observations (SYNOP) and meteorological aerodrome reports (METAR) formatted stations. As of 2022, there are about 12,000 stations of the GTS that are included in the CADB, where the number is also based on whether there is quality meta-data available for a station.

With the release of CADBv2, global station data was made available to the public (previously limited to the U.S.). Additionally, archived data was made available and many more parameters are included in the summary data files. Previously CADBv1 public data focused primarily on temperature and precipitation. Inclusion of more meta-data in the CADBv2 summary files allows more potential for user applications downstream, including easier plotting of values if desired. Table 1 gives an overview of the meta-data and parameters included in the daily summaries. Weekly, monthly, and seasonal summaries are also updated regularly, containing statistics based on temperature and precipitation.

Documentation is available on the CADB webpage that provides details regarding the format, range, and units of the parameters.

Table 1 List of meta-data and available parameters in the CADB dataset.

Metadata	Parameters
Station ID	Max and min temperature
Station call	Reported precipitation - Based on reported “6 hour” and “24 hour” precipitation amounts
City	Final precipitation - Complex estimated amount based on reported precipitation, weather
State	type, various adjustments based on <i>e.g.</i> latitude, <i>etc.</i> , time and value weighted
Country	Precipitation flag - Denotes source/quality of the final precipitation amount
Date	Number 6-hr precip reports - number of 6hr report values (assigned to 3-hr time steps)
Lat	Weather characters - Weather character string representing weather type for 8 times
Lon	Trace - Flag denoting 0 for no trace, 1 for trace precip
Elevation	Vapor pressure and vapor pressure deficit
	SLP (at 6Z, 12Z, 18Z, 0Z [day after valid])
	Max and min relative humidity
	Apparent temperature (heat index) and wind chill
	Wind speed at 8 times (relative to the beginning of the precip bounding period)

4. Methodology

The CADB is driven by a suite of scripts and codes that perform various levels of formatting and quality control. A diagram is shown (Fig. 1) that outlines the framework and process flow of CADBv2 that starts from ingesting raw GTS data in binary universal form for the representation of meteorological data (BUFR), ending

with CADB output provided to users. CADB data is stored on an internal relational database at the National Centers for Environmental Prediction (NCEP) that serves a variety of purposes including providing the back-end data for the newly created CADB time series web application.

The CADB rewrite project was a multi-year effort to improve the many parts of the CADB suite, including rewriting software to produce improved station meta-data and normals, leading to the overall improvement of the CADB product suite as a whole. The first year was dedicated to the understanding of the v1 code and methodology of the CADB software suite, as well as planning methodologies for the new version of the code. Subsequent work entailed an iterative process of developing code to calculate summary values for the different parameters, evaluating results by benchmarking with v1 and external sources and interacting with users to get feedback, and making further adjustments to the code.

The approach and equations used to calculate the daily summary values based on hourly data is similar to those used by CADBv1, plus modifications made based on updates in equations or methodologies that were deemed appropriate and beneficial by the CADB development team. Each time a change was made to the software, a variety of techniques were employed to determine whether this change was kept in the software including 1) evaluating the percent of stations over the globe for a sample day/week where the difference between the CADBv1 and v2 values exceeds a set acceptable threshold, 2) manual inspection of values for specific cases including looking at actual values and maps of data, 3) comparison to other external sources of summary values, and 4) feedback from users with expertise in assessing specific variables and regions. Information about any newly released versions of the software and related impacts to data are detailed in technical notes available on the home CADB summary data webpage.

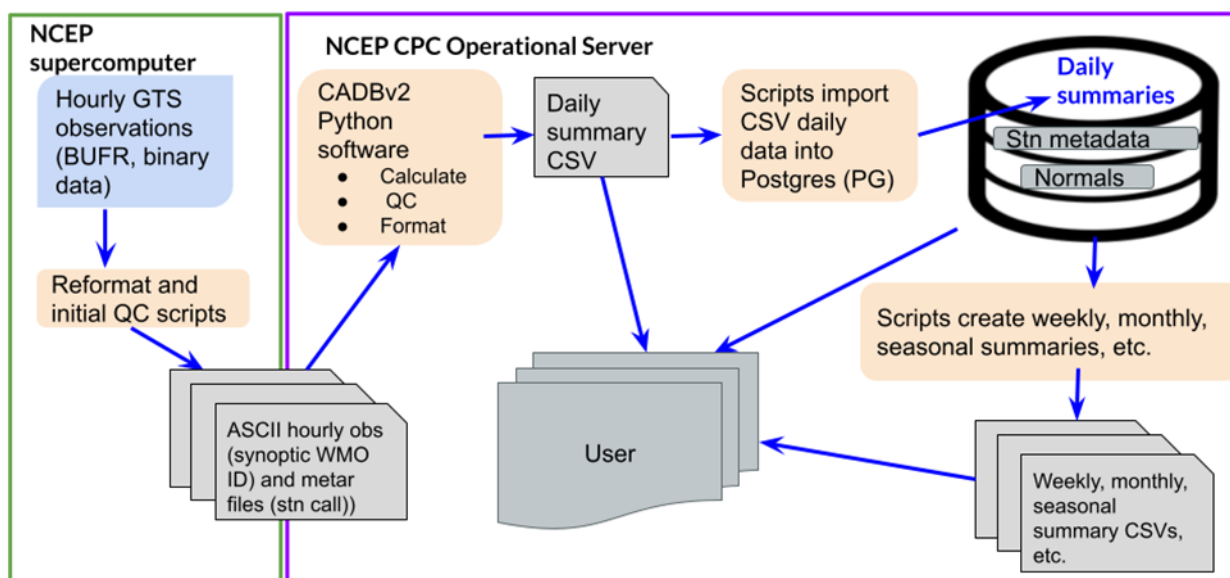


Fig. 1 Diagram of process flow indicating how raw station observations get processed and summarized to CADB output.

5. CADBv2 upgrades and changes compared to v1

Version 2 of the CADB data is available from January 1, 2020 to present. Many improvements were made to various parts of the framework. A list of some notable upgrades and changes are listed below:

- Expanded datasets of summaries available to the public including global coverage, whereas previously only U.S. data was provided. The number of stations may vary slightly day to day depending on the reports that come in.
- More parameters are available; previously the data focused primarily on temperature and precipitation.
- More meta-data available in the CADB summary, making it easier to ingest into downstream applications, such as plotting in a geographic information system (GIS).

- Well documented software suite
- Core software that calculates daily summary values is written in Python, replacing f77 code
- Many scripts doing the initial QC and formatting of the raw hourly data were improved, including retrieving more stations
- Replace direct access database in backend with relational database enabling web applications to dynamically pull data and much easier to maintain
- Software is in version control (git), promoting code collaboration and ease of maintenance and tracking changes
- In many of the CADBv1 files, METAR stations were identified as '99' followed by 3 characters. CADBv2 will now use 4 characters after '99'. Therefore, the METAR stations will change from 5 to 6 characters in total (e.g. '99CHO' will be '99KCHO'). This allows potentially more stations because the upstream data is actually reported with 4 characters, with the first character denoting the region. In CADBv1, we were using only U.S. METAR. Additionally, there are stations in the contiguous U.S. (CONUS) and Alaska that have the same 3 last characters with a different char at the beginning. CADBv1 only picked the CONUS version of the 3-char station if it was a duplicate. For example for station 'KABR' and 'PABR', with 'KABR' being the CONUS station and 'PABR' being the Alaska station. CADBv1 only chose the CONUS station, precluding the Alaska PABR station from being output. CADBv2 would now include, e.g. '99KABR' and '99PABR'.
- Summary files are formatted differently, previously a fixed width text file now formatted as a comma delimited CSV file that is easier to parse and contains header column names.
- New column with station call letters, if there is one. If not, a '-99999' is listed for a station that does not have a station call.
- Many stations have higher resolution of latitude/longitude information (up to 4 decimal places), which can be especially important for plotting in GIS.
- Archive of data available now, previously only 7 day rotating files were available. Many users requested days of data prior to the current 7 days, so this allows users to grab past data now.
- File naming convention has changed to include the valid summary date since files are no longer rotating, e.g. daily files named *daily_summary_YYYY\$mm\$dd.csv*

6. Related upgrades

6.1. Station meta-data

In addition to the core software that produces summary analyses, significant time in the CADB rewrite project was also dedicated to rewriting software to produce improved meta-data. One reason that station meta-data is important to the quality of the CADB data is that many stations require matching station call identifiers (ID) associated with the METAR data input stream with the World Meteorological Organization (WMO) formatted IDs that are used with the SYNOP input. These stations have observation data that come in through both METAR and SYNOP sources upstream. Therefore, it is essential that the different forms of identifiers are matched correctly in order to combine the hourly data for the proper stations.

Previous to the meta-data software rewrite, maintenance and upgrading of the station meta-data was more of an ad-hoc approach and completely manual. Now, there is a set of scripts that uses a complex approach to combine various sources of meta-data to produce a meta-data file which is utilized by various downstream processes, including the core CADB software that produces the summary values. A major challenge of station meta-data is that despite multiple external sources maintaining a station meta-data library, there are significant differences among them including how METAR and SYNOP IDs are matched, associated latitude and longitude, etc. There are also stations that are known to be incorrect in these meta-data sources, but these issues are only found when the CADB team or users notice erroneous summary values. Therefore, a methodology was devised to try to optimally combine meta-data using multiple sources, while retaining all the previous manual adjustments that need to be applied.

Input sources for meta-data reference include the NCEI enhanced global station meta-data list, recent real-time GTS data, legacy CADB meta-data file, and aviation weather global meta-data file. After reformatting data as necessary, multiple steps of assessment and QC are performed to determine a best guess for each of the meta-data parts with tiers of prioritizing quality of data, which was based on iterative trial and error of evaluating various stations and getting feedback over time from users based on the CADB output (e.g. a regional user identifying an incorrect station assigned) and those with meta-data expertise. This updated station meta-data list is available on the CADB webpage with the summary data.

In addition to the WMO formatted ID and associated METAR call ID (where applicable), logic is included in the software to assign an NCEI formatted station ID. It was necessary to pair NCEI formatted station IDs with the WMO SYNOP and METAR station IDs in order to assign NCEI 30-year normals which are in an NCEI ID convention.

6.2 Normals software

New Python software was created to produce quality upgraded station normals for temperature and precipitation based on various sources of available data. These CADB station normals feed into a number of downstream applications, including usage by CPC's international desk and the global CADB time series web application. This includes three sources: 1) archived CADB daily input, 2) monthly summaries from the Global Historical Climatology Network (GHCN) available on the NCEI website, and 3) NCEI monthly and daily normals. A suite of calculated monthly and daily normals are produced from these inputs, including usage of the station meta-data to pair station IDs with different formats, and various steps of QC including comparing values to the previous 30-year normals. This new software was then used to produce updated 30-year normals for the 1981-2010 base period, and 1991-2020 normals are being reviewed for release currently as well.

6.3 Time series web application

A new web application was developed in tandem with the CADB software suite rewrite, which replaces the globally available static graphics of temperature and precipitation time series charts using CADB data. This web application allows users to display charts of time series based on CADB summary values for many global stations. Charts include information such as daily values, departures from normal (anomalies), accumulated precipitation, and minimum and maximum temperatures over a customizable period. There are many more global stations available now compared to the previous static graphics, which utilize the updated station normals previously discussed for climatology information. The web application allows users to interact more with the backend data, and includes new features such as the ability to hover over chart points to get values, customizable

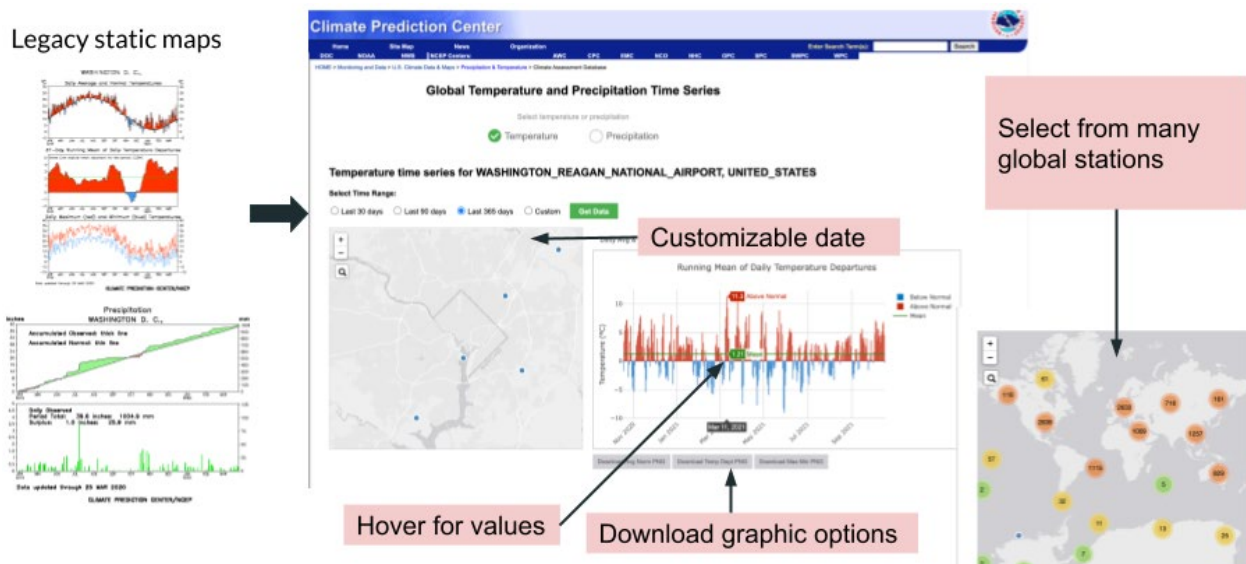


Fig. 2 Diagram showing previously available legacy CADB time series static maps (left) and new dynamic web application (center and right).

dates and periods of assessment, and downloadable graphics in multiple formats (Fig. 2). Available stations for time series charts are based on which stations have available 30-year normals to pair with CADB summary data. The time series application is available at <https://www.cpc.ncep.noaa.gov/products/timeseries>.

7. Results

Quality preliminary global station summary data is now produced with more extensive parameters and stations over the globe, including using upgraded well documented python software for many parts of the CADB framework. Archived data is saved in a relational database that allows for more reliable maintenance and easier access of past data, including the ability to serve data dynamically to the time series web application. CADB data has been evaluated and continues to be evaluated in real-time by many stakeholders and is run operationally in real-time for routine updates. Data results in general are fairly similar to v1, although there are some parameters that have more notable differences, for example the heat index which now uses an updated methodology used by the Weather Prediction Center.

Another noticeable difference between v1 and v2 is that v1 had a known dry bias in some areas of the world. It was found that v1 had assigned zero precipitation values where there shouldn't have been. It is suspected that v1 set many values to zero due to suspicion of raw station values. However, there have been many improvements to the quality of data since the inception of v1 in the 1980s, so more station precipitation reports should be deemed valid and are thus incorporated in v2. These zero precipitation values previously produced in CADBv1 were evaluated by the CADB team and stakeholders, particularly from JAWF, and deemed incorrect. The approach in CADBv2 retains more valid ingested values, thus resulting in more locations that have valid values rather than a zero assigned. This, however, has made localized areas sporadically have more “bullseye” values compared to neighboring values in space and/or time. These situations are not deemed to be systematic, and are considered a caveat to expect when dealing with preliminary data that do not use more complex QC techniques. These inherent issues are associated with sporadically occurring erroneous raw reports from the station where no systematic light QC can be applied.

8. Challenges and caveats

There are inherent challenges associated with assessing raw station data due to the nature of the inconsistency of how upstream raw data is reported at stations across the globe. As discussed in the ‘results’ section, there are many challenges in trying to devise QC that can be systematically applied to data, particularly when “bullseye” values deemed too large occur sporadically. It is also difficult to find reliable ground truth values from other sources to determine whether a value is actually erroneous. Significant caution must be taken in adding a QC technique in the CADB summary software since fixing some station values may result in degrading the quality of data for other locations. Other sources of data, such as from NCEI, should be used if higher QC data is desired.

There are inconsistencies across the global stations in terms of the level of maintenance in the measuring and reporting equipment over time, and thus the level of quality of the reported values. Also, reports are sent at many various times of a day with varying frequencies, where some stations routinely report at least once an hour if not more, whereas other stations have sparse reports for the period being assessed, which can result in a poorer quality summary value due to poor sample size of the raw reports.

9. Summary

A revamped suite of software has been developed at CPC to produce global station observation summary data with improved quality and stability, including upgraded meta-data station data, 30-year station normals, and a new time series web application. The CADB development team and other CPC teams continue to perform ongoing assessment to identify any changes that could be incorporated into the suite of software for further improvement. Publicly available CADB summary data is now in an improved user-friendly format in addition to providing new additional parameters, expansion to global stations, and access to the archive of available data.

Moisture Based Agroclimate Indices Across the Canadian Prairies Under a Changing Climate

Aston Chipanshi, Dongzhi Qi, Yinsuo Zhang, and Mark Berry

Science and Technology Branch, Agriculture and Agri-Food Canada, Regina, SK, Canada

1. Introduction: Motivations and overview of the work

In our ongoing assessment of agroclimatic indices across Canada's largest agricultural zone in the Prairie Provinces under a changing climate (Chipanshi *et al.* 2021) (Fig. 1), we present a comparative analysis of soil moisture based indices derived from a water budget model, the Versatile Soil Moisture Budget (VSMB) (Baier *et al.* 2000) under the present and future climatic scenarios. We analyzed values of available soil moisture and water stress for the summer peak crop growth period: June, July and August (JJA), when moisture limitations have their maximum impact on crop growth and final crop yield and production. Simulations of the water related indices were based on a cool season crop (spring wheat), which is typically seeded in the spring and harvested by end of September. The wheat crop in this study represents other cool season crops grown on the Canadian Prairies such as barley, oats, canola etc., which currently have the largest seeded acreage (Statistics Canada, 2019). This research was motivated by the following questions which resource based sectors are asking as effects of climate change gain prominence: 1) Will the Canadian Prairie agricultural space remain suitable in supporting the production of cool season crops under climate change? 2) Is there a likelihood that the current agricultural zone for cool season crops will shrink or expand under climate change?



Fig. 1 Map of the study area with insert showing the location of the Canadian Prairies within Canada. The dotted line (red) describes the cropped zone under the present climatic conditions.

2. Data and methodology/experimental design

Daily soil moisture down to the root zone (~120cm) and crop water stress were simulated for all of the Canadian Prairies south of the 60° parallel for the base climate period of 1981 to 2010 and for both the medium (RCP4.5) and high (RCP8.5) emission projections extending to the distant future (2071–2100) (van Vuuren *et al.*, 2011). The Versatile Soil Moisture Budget, which works as a simplified bucket model, was used to calculate the following moisture based agroclimatic indices.

- i) *Available Water Content as a fraction of the total Water Holding Capacity of the soil layer down to the root zone* - The Available Water Holding Capacities (AWHC) of the Canadian Prairie soils which define the range of soil moisture levels across the study area can be found in De Jong and Shields (1988). The AWHC values typically range from 50mm for the sandy soils to 250mm for the fine

textured soils. In order to understand the effects of climate on soil moisture depletion in more detail, the AWHC for all grid points was set to 150mm and this represents a sandy loam soil on the Canadian Prairies. The simulated water by volume from the four layers was then converted to mm equivalents by multiplying volumetric content by depth. The product was then expressed as a fraction of the water holding capacity. With this normalization, the spatial and temporal comparisons could be made.

- ii) *Crop Water Stress* - Crop water stress (CWS) was calculated as the normalized ratio of actual and potential evapotranspiration from:

$$\text{CWS} = 1 - (\text{AET}/\text{PET}) \quad (1)$$

where AET is the actual evapotranspiration and PET is the potential evapotranspiration. Ordinarily, AET is estimated from PET, which defines the maximum amount of moisture that can be lost under prevailing atmospheric conditions. In order to calculate PET at many locations in the study area, a temperature based method was used (Baier and Robertson, 1968) unlike other methods that require multiple variables to estimate PET. To estimate maximum AET from PET, a crop coefficient (K_c), which is dependent on crop stage is multiplied by PET. The actual transpiration is affected by soil moisture in the root zone. Apart from soil moisture content in the layer, the coefficient for determining AET is also affected by the density of roots. An example of how to derive soil water coefficients can be found in Ritchie *et. al* (1973).

- iii) *Climate Moisture Index* (CMI) - The Climate Moisture Index is the difference between Precipitation and Potential Evapotranspiration for determining water requirements for multiple uses such as crop growth and irrigation. It is calculated from:

$$\text{CMI} = P - \text{PET} \quad (2)$$

where P is the total precipitation (gridded daily P values in this study) and PET as calculated in Equation 1.

The VSMB requires inputs of daily maximum and minimum air temperature and precipitation. These were obtained from the 10 km gridded data set from 1981 to 2010 (<https://cfs.nrcan.gc.ca/projects/3/4>). Similar variables were needed to run the VSMB for two emission scenarios: RCP4.5 (moderate emission) and RCP8.5 (high emission) from 2006 to 2100 (<https://www.pacificclimate.org/data/statisticallydownscaled-climate-scenarios>). Six GCMs were used to generate T and P per emission scenario. The VSMB was configured to run with 4 soil layers in 30 cm increments from 0 to ~120cm.

3. Results

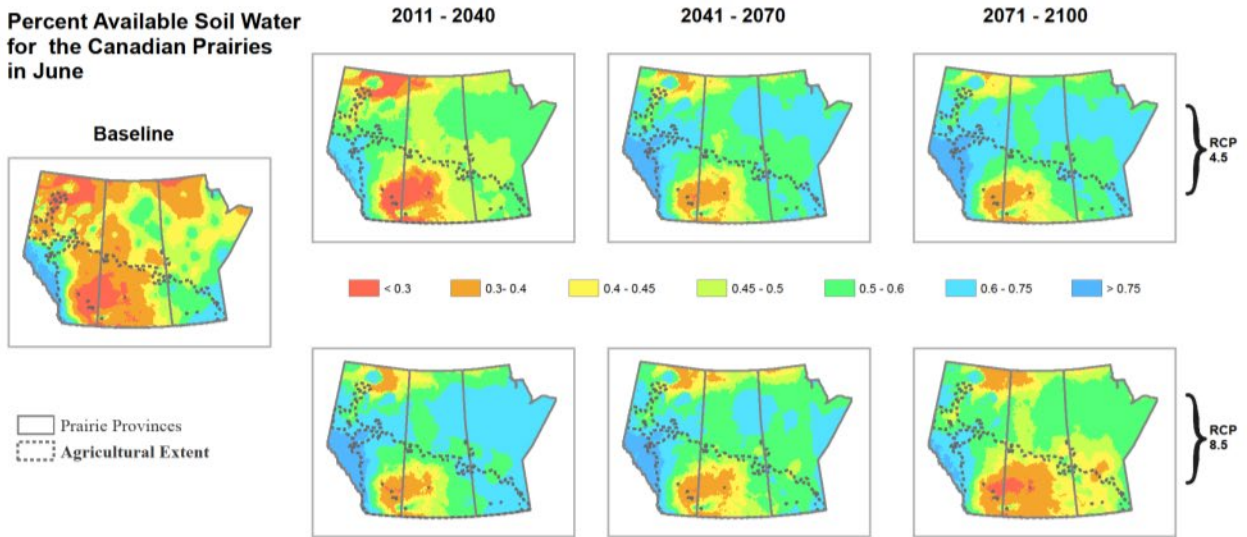
a) *Soil moisture availability in summer*

Figure 2 shows a comparison of available soil moisture for the months of June to August and compares the climatological values (shown as baseline or 1981 to 2010) with values simulated from climate change scenarios under the medium (RCP4.5) and high (RCP8.5) emission scenarios. Climate change results are presented in 30-year time steps as near future (2011 to 2040), far future (2041-2070) and distant future (2071-2100). The baseline maps show that soil moisture is usually depleted in August, which nearly coincides with the period when cool season crops reach maturity or near harvest time. Under climate change, soil moisture depletion occurs much earlier (July) irrespective of the emission scenario although depletion is amplified under the high emission scenario. The spatial extent of the area with the least moisture (red) is relatively small for August compared to July and Manitoba is less impacted compared to Alberta and Saskatchewan. The early peak in soil moisture depletion in July under climate change is in response to early summer warming that has been reported across the study area (Tam *et al.*, 2019).

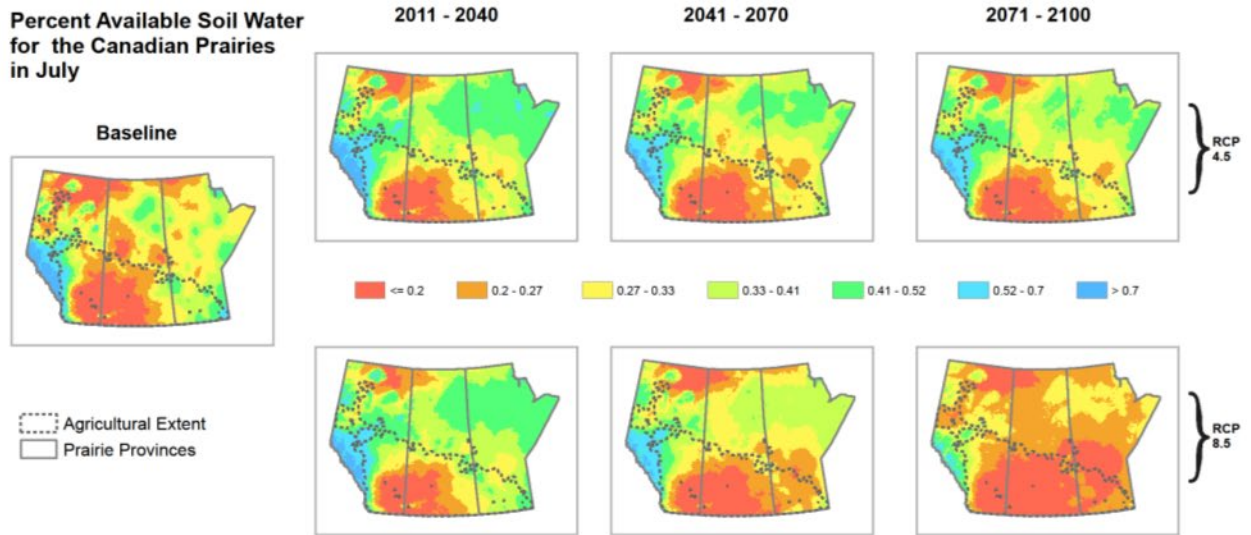
b) *Crop Water Stress*

Under climatological conditions (1981-2010), crop stress is least in July when the majority of precipitation is received across the Canadian Prairies (Fig. 3). There are exceptions from southern Alberta and southwestern Saskatchewan where pockets of stress in the order of 0.4 and above can be expected over the long term.

Percent Available Soil Water for the Canadian Prairies in June



Percent Available Soil Water for the Canadian Prairies in July



Percent Available Soil Water for the Canadian Prairies in August

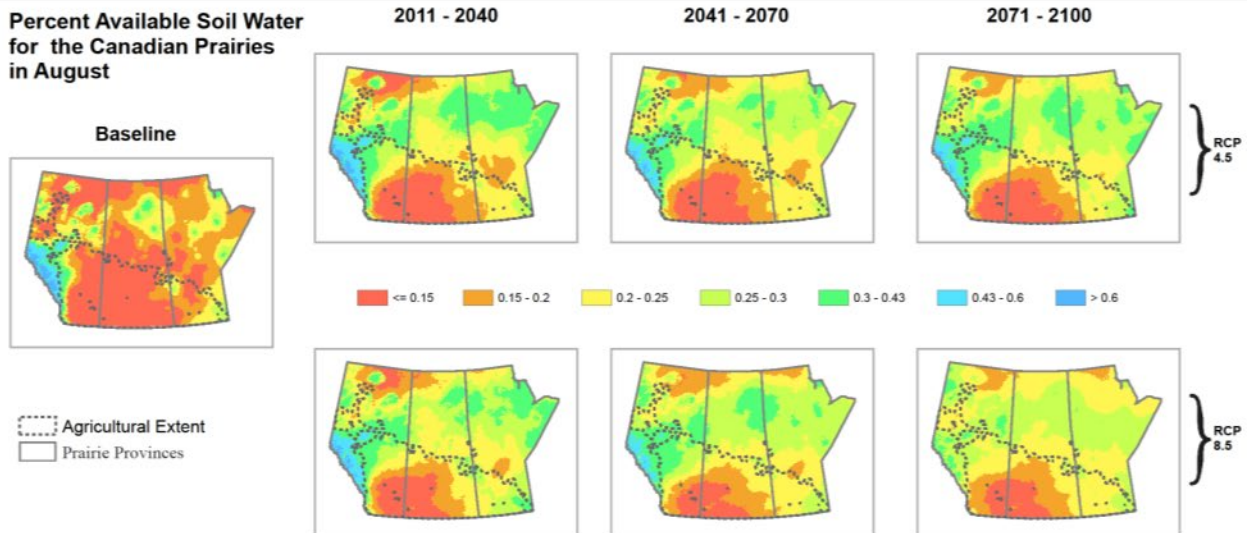
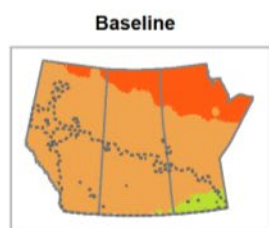
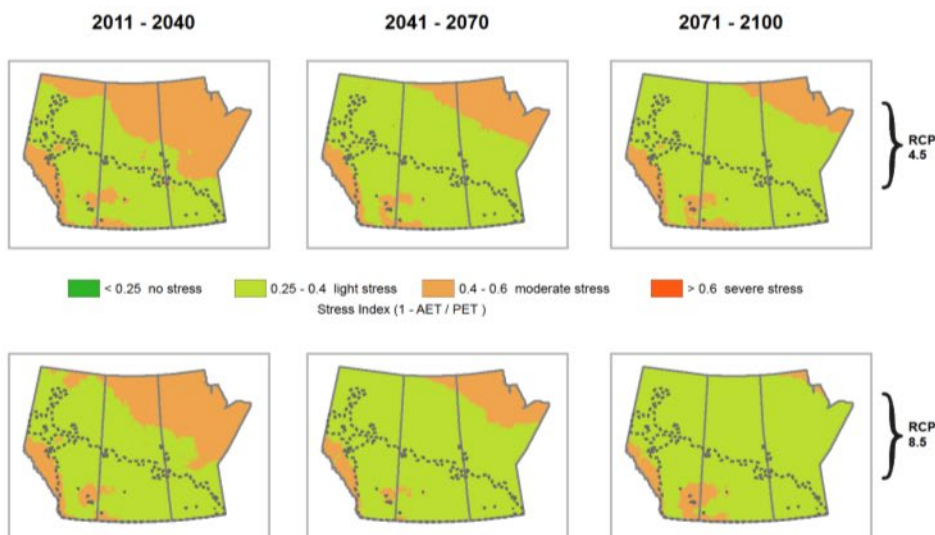


Fig. 2 The spatial and temporal variability of soil moisture during the summer months (From top to bottom panels are for June, July and August, respectively).

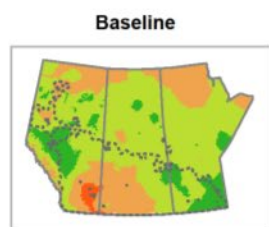
Stress Index for the Canadian Prairies in June



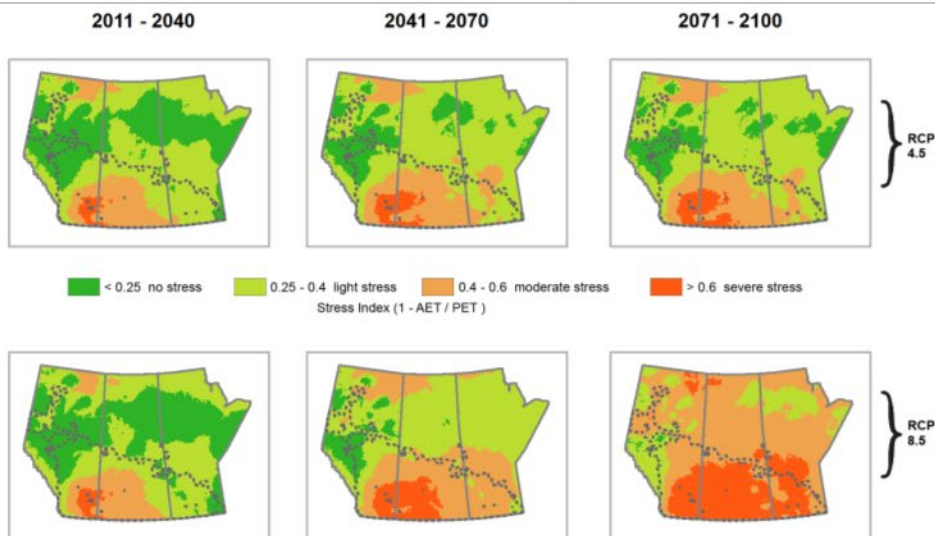
⋯ Agricultural Extent
□ Prairie Provinces



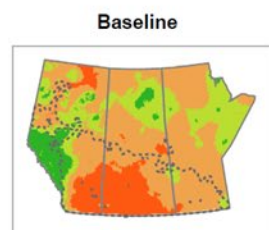
Stress Index for the Canadian Prairies in July



⋯ Agricultural Extent
□ Prairie Provinces



Stress Index for the Canadian Prairies in August



⋯ Agricultural Extent
□ Prairie Provinces

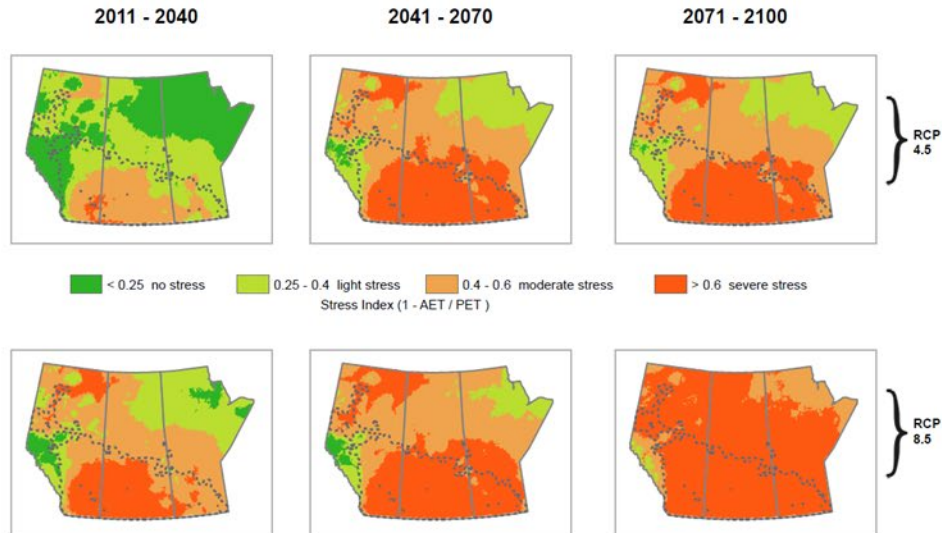


Fig. 3 Simulated crop water stress for a wheat crop under the present and future climate change scenarios (From top to bottom panels are for June, July and August, respectively).

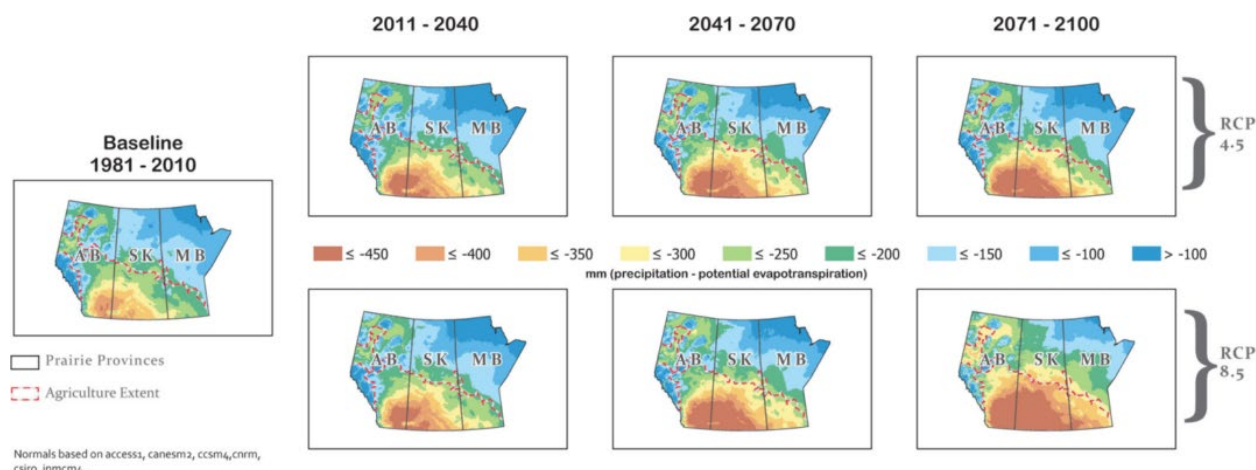


Fig. 4 The simulated Climate Moisture Index (P-PE) under the present and future climate scenarios during the Growing Season ~May to August.

Under both medium and high emission scenarios, crop stress expands beyond the agricultural zone and reaches its maximum during the month of August.

c) Climate Moisture Index

The southern agricultural zone of the Canadian Prairies experiences water deficits (the difference between Precipitation and Potential Evapotranspiration) of up to 450mm during the growing season under the observed climate (1981-2010); baseline map. Under the climate change scenarios studied, water deficits in the order of ≤ 450 mm expanded northwards and eastwards for both emission scenarios across the 30-year climate periods up to the distant future (Fig. 4). On a relative basis, the CMI index shows that Manitoba's water deficit ranks better than Alberta and Saskatchewan.

4. Summary

We found changes in the temporal and spatial distribution of the water based indices across the Prairie landscape. Chiefly, the timing of when soil moisture is least in the soil occurs early under climate change (around July) compared to the observed climate period (1981 to 2010) where soil moisture depletions occurs in August. Using the difference between precipitation and potential evapotranspiration as a proxy for water availability, moisture deficits expanded northwards and eastwards and beyond the current agricultural zone under the high emission scenario. Based on these findings, rain-fed agriculture for the traditional cool season crops like wheat, barley and canola may become riskier to grow in the distant future due to increased dryness in the soil and high evaporative losses conditioned by high temperatures. It will be possible to adjust the crop calendar to take advantage of when these risks are less intense as an adaptation measure, however a detailed analysis of the agroclimatic indices at finer scale (temporally and spatially) will provide more insight for addressing variability.

References

- Baier, W., J. B. Boisvert, and J. A. Dyer, 2000: The Versatile Soil Moisture Budget (VB) reference manual [Computer software], Eastern Cereal and Oilseed Research Centre (ECORC) Contribution No. 001553. Agriculture and Agri-Food Canada, Ottawa, ON.
- Baier, W. and G. W. Robertson, 1968: The performance of soil moisture estimates as compared with the direct use of climatological data for estimating crop yields. *Agric. For. Meteorol.*, **5**, 17-31.
- Chipanshi, A., M. Berry, Y. Zhang, B. Qian, and G. Steier, 2021: Agroclimatic indices across the Canadian Prairies under a changing climate and their implications for agriculture. *Int. J. Clim.*, doi: 10.1002/joc.7369.
- De Jong, R., and J. A. Shields, 1988: Available water holding capacity maps of Alberta, Saskatchewan and Manitoba. *Can. J. Soil Sci.*, **68**, 157-163.

- Ritchie, J. T., 1973: Influence of soil status and meteorological conditions on evaporation from a corn canopy. *Agron. J.*, **65**, 893-897.
- van Vuuren, D. P., J. Edmonds, M. Kainuma, K. Riahi, A. Thomson, K.-H. Hibbard, C. George, T. Kram, V. Krey, J.-F. Lamarque, T. Masui, M. Meinshausen, N. Nakicenovic, S. Smith, and S. K. Rose, 2011: The representative concentration pathways: an overview. *Clim. Change*, **109**, 5, doi:10.1007/s10584-011-0148-Z.
- Statistics Canada, 2019: Estimated areas, yield, production, average farm price and total farm value of principal field crops, in metric and imperial units. Table: 32-10-0359-01 (formerly CANSIM 001-0017), <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3210035901>.
- Tam, B. Y., K. Szeto, B. Bonsal, G. Flato, A. J. Cannon, and R. Rong, 2019: CMIP5 drought projections in Canada based on the Standardized Precipitation Evapotranspiration Index. *Canadian Water Resources Journal/Revue Canadienne des ressources hydriques*, **44**, 90-107, doi:10.1080/07011784.2018.1537812.

Extended Range Verification Using Economic Value

Eva Regnier and Joel Feldmeier

Naval Postgraduate School, Monterey, California

Economic value estimates can be used to evaluate forecast performance. Called Value-of-Information (VOI) when used to compare alternatives in areas such as pharmaceutical R&D and oil and gas exploration, this approach can be used much like verification metrics to identify regions, seasons, and lead times where extended-range forecasts outperform climatology for use in decision-making. This approach easily supports multi-variable performance evaluation. It also addresses some end-users' assumption that extended range forecasts aren't useful because the shorter-range forecast is more accurate.

1. Introduction

Murphy (1993) describes the conventional metrics that summarize the relationship between forecasts and observations - such as error measures, RMSE, hit rate, accuracy - as "quality" measures. Economic value, which Murphy simply calls "value", is the expected savings for decisions made using forecasts, compared to using climatology.

Quality metrics are applied in many domains of forecasting. They depend only on forecasts and ground truth, and are independent of any specific use-case. Typically used in contexts where acquiring information is costly, Value-of-Information (VOI) metrics integrate forecasts and ground truth with models of specific decision contexts. They can combine into a single measure of the impact of multiple variables, forecast products, and lead times. A further difference is that the units of quality metrics are environmental scales, e.g. length, temperature. The exception is metrics like hit rate and false alarm rate that are most similar to VOI metrics. VOI metrics are in end-user-relevant units like \$, delay, or the probability of a good or bad outcome.

The main drawback of VOI metrics is that they depend on a specific decision context. However, since VOI can show skill in forecasts that show no skill using conventional verification metrics---and vice versa---it can pay off to search for skill using VOI metrics for specific end users.

Moreover, some common characteristics of decisions made using extended-range forecasts can narrow the focus to decision models most relevant to exploiting these forecasts.

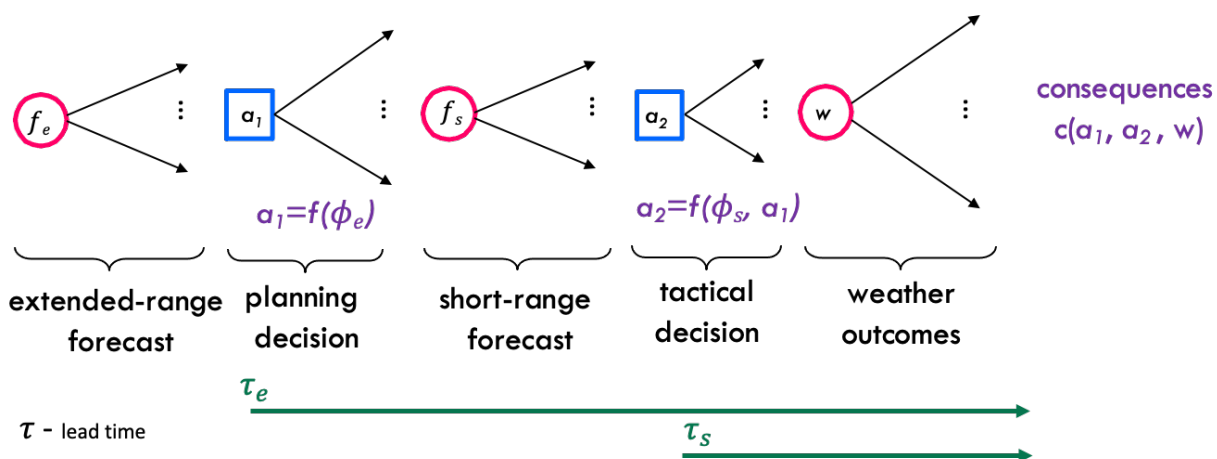


Fig. 1 Schematic of two-stage decision, where the consequences of weather outcomes are affected by planning decisions that must be made based on an extended-range forecast and tactical decisions that can be made using a shorter-range forecast.

2. Common features of extended-range decisions

Decisions that end users would make using extended range forecasts are systematically different from the ones they make with shorter-term forecasts. In general, end users will use longer-range forecasts to make decisions that take longer time to implement, such as when consequences depend on the location of a slow-moving asset. For example, if you are planning a major military exercise or other event that requires people to converge at a selected location, it would make sense use an extended-range forecast to make that choice.

An action that requires a long time to implement is also likely to be costly—and therefore only done in an unusual case, for an extreme environmental event: if it's something that the end user would need to do often, they would have invested in a way to make it less costly. The consequences of extended-range decisions may depend on variables that accumulate over time, such as precipitation, or drought.

Longer-lead decisions include stage-setting decisions such as reserving some kind of agricultural equipment, pre-positioning humanitarian response supplies or weather-related retail inventory. This means that later decisions that can be made with shorter-range forecasts can still affect the consequences. In other words, there is flexibility remaining at shorter lead times: *e.g.*, ship routes may be updated repeatedly. The value of the extended-range forecast is similar to option value.

Some of these characteristics we can use to make some generalizations about the types of value or with the types of decision contacts in that we can use some of the ways we can look for value in forecasts.

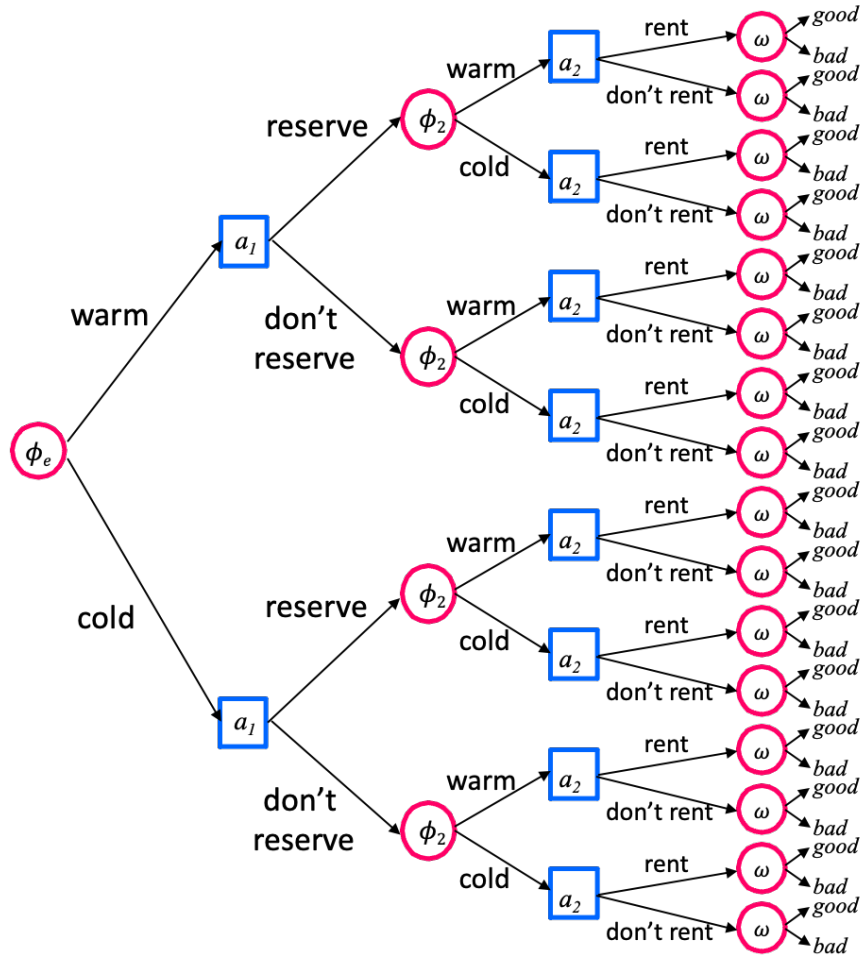


Fig. 2 Illustration of the garden party decision scenario. Time moves from left to right—decisions are made using the forecasts available before each decision.

3. Two-stage decisions: Garden Party Example

VOI allows us to measure the value of extended range forecasts while taking into account the fact that the decision maker may make related decisions using more accurate, later forecasts. We illustrate this with a simple two-stage cost-loss model (Fig. 1).

Using the notation from Figure 1,

$$\text{VOI} = E[c(a_1, a_2, w) - c(f(\phi_e), f(\phi_s), w)]$$

The first term is the consequences of reference decisions—*e.g.* the end user uses climatology to make the decision, while in the second term, the two decisions are made using forecasts. The decision policies $f(\phi)$ do not

have to be identical at the two stages. Like quality metrics, VOI depends on the relationships among ϕ_e , ϕ_s , and w . However it also depends critically on the user specific function $c(\cdot)$.

We will take the simplest two-stage decision model, a two-stage cost-loss model with two alternatives at each stage and two weather outcomes. This structure can be adapted to many decision contexts, but as an accessible narrative to explain the parameters, we'll treat it as a garden party, illustrated in Figure 2. The end user is planning a garden party for a particular date and time. The extended range forecast for that date, ϕ_e is received first---summarized as either warm (enough) or cold. At that time the end user must decide whether to reserve a tent, incurring a cost c_1 . Later, the enduser receives a second forecast ϕ_2 for the date and time of the party, and must then decide whether to rent the tent, incurring an additional cost c_2 .

The consequences of these two decisions further depend on the weather outcome w ---if the weather is *good*, the consequence is zero plus any costs incurred for the tent. If the weather is *bad*, then if the end user rented the tent, the consequence is again zero plus tent costs, but if the end user did not rent the tent, the consequence is *Loss*, normalized to one.

VOI is often measured as a percent of the value of perfect information---with perfect information, the end user would prepare for adverse weather when it occurs and never otherwise.

4. Data used to illustrate

To illustrate the use of VOI to evaluate extended-range forecasts, we use the NOAA MDL GFS MOS MEX for the extended range,¹ and MOS MAV guidance for short-range decisions.² For ground truth, we use hourly data from NCEP ISD,³ and for climatology, max average daily temperatures by day of the month average from 1981-2010,⁴ for twenty stations in the contiguous United States.

We average results over all days in April and May, 2015-2019, all valid at 1800 UTC, at the following twenty stations: Cheyenne, WY: KCYS, Detroit, MI (Wayne): KDTW, New Orleans, LA INTL: KMS, Phoenix, AZ Sky Harbor: KPHX, Seattle-Tacoma, WA: KSEA, Washington D.C. Reagan National: KDCA, KBLV, KDAY, KJAX, KMCN, KBOI, KOKC, KOMA, KORF, KPSM, KPUB, KSAN, KSAT, and KVCB.

Since the available climatology gives daily maximum temperature, and actuals are for 1800Z, we estimated the bias of using the daily maximum as a prediction of the 1800Z temperature, and removed this bias of 5.1° to estimate a climatological temperature for the valid time.

5. Effect of short-term flexibility

The costs c_1 and c_2 are a proportion of *Loss*, and $0 < c_1, c_2 < 1$, and additive, i.e. the consequence of choosing to both reserve and rent the tent = $c_1 + c_2$. However, the end user cannot rent the tent if they haven't already reserved it. More general models can also be included by changing this last assumption, as in the next section.

In this model, short-term flexibility is reflected in how much of the total cost of the tent can be held back until the short-range lead time. In other words, how large is c_2 as a proportion of $c_1 + c_2$. If $\frac{c_2}{c_1 + c_2} = 0$, there is no flexibility, and the entire cost of the tent must be committed at the extended-range lead time. If $\frac{c_2}{c_1 + c_2} = 1$, there is no cost to reserve the tent, and the extended-range forecast has no value---because it is never worthwhile to reserve the tent.

Figure 3 shows how flexibility, measured as $\frac{c_2}{c_1 + c_2}$ affects the average VOI for 18Z temperature forecasts for the stations and dates listed above, measured as a percent of the value of perfect information, with $c_1 + c_2 = 0.25$ and the weather is considered bad if the temperature is less than 60°F and good otherwise. RMSE is shown

¹ <https://vlab.noaa.gov/web/mdl/extended-range-gfs-mos>

² <https://vlab.noaa.gov/web/mdl/short-range-gfs-mos>

³ <https://www.ncei.noaa.gov/access/search/data-search/global-hourly>

⁴ <https://www.ncei.noaa.gov/products>

in black on its own axis as a reference—it is in temperature units (not shown). However, the red bar at the bottom of each plot represents the region in which the skill is negative—assuming climatology is used in the extended-range decision and everything else is identical to the with-forecast decision. In the climatology reference, the short-range decision is still made using the short-range forecast. Climatology differs by station and date, and therefore, for some dates and stations, the climatology will call for reserving the tent, and others it will not.

While RMSE skill is positive at all lead times in this example, which only goes out to 192 hours, when there is more flexibility remaining at shorter lead times, the VOI of the extended-range forecast is negative for longer lead times. The less flexibility remaining at the shorter lead time, and the bigger the commitment at the extended-range lead time, the higher the value of the extended-range forecast. In other words, the characteristics of the decision maker—in this example, flexibility summarized as $\frac{c_2}{c_1+c_2}$, affects whether the forecasts have value (skill) relative to climatology. Although not shown here, the threshold for *bad* weather and the total cost to prepare also substantially affect the VOI.

Note that the short-range lead time also affects the VOI, which makes sense as the shorter the lead time, the better the short-range forecast, but the effect is relatively small for short-range lead times of 1–3 days.

6. Conclusions

Economic value metrics can find value over and above climatology in forecasts where that show no skill using conventional verification metrics—and vice versa. In the example above, we illustrated the VOI method using lead times of only 8 days, but in other work, we find positive skill using extended range forecasts at lead times where RMSE does not.

By exploring VOI as a function of user characteristics, it is possible to find skill that would be overlooked if you use only accuracy measures. It can therefore help identify end users who can benefit from forecasts.

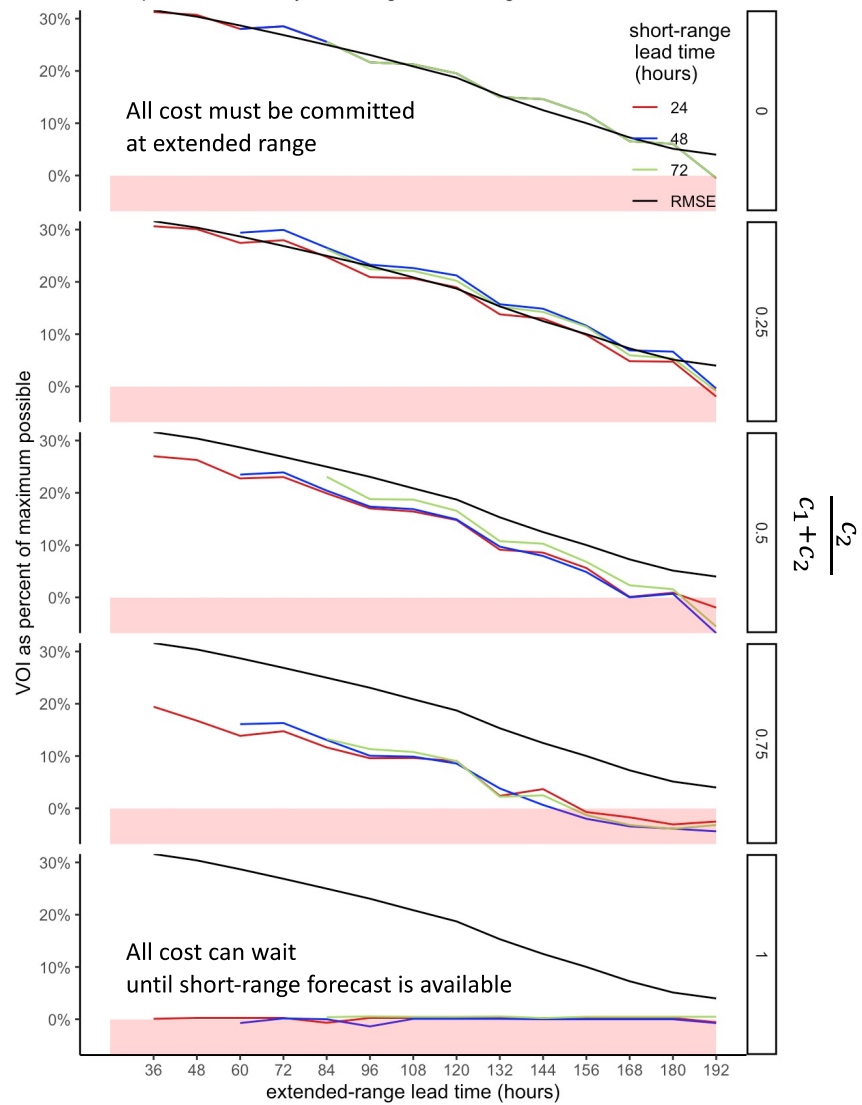


Fig. 3 VOI depends on flexibility remaining at short-range lead time. VOI, measured as a percent of the VOI for perfect forecasts, is shown as a function of the extended-range forecast lead time, the short-range lead time, for five levels of flexibility, with RMSE on its own scale shown as a reference

References

- Murphy, A. H., 1993: What is a good forecast? An essay on the nature of goodness in weather forecasting. *Wea. Forecasting*, **8**, 281-293.
- Pegion, K., B. P. Kirtman, E. Becker, D. C. Collins, E. LaJoie, R. .Burgman, R. Bell, and Coauthors, 2019: The Subseasonal Experiment (SubX): A multimodel subseasonal prediction experiment. *Bull. Amer. Meteor. Soc.*, **100**, 2043-2060, doi:10.1175/BAMS-D-18-0270.1

NWS Science and Technology Infusion Climate Bulletin

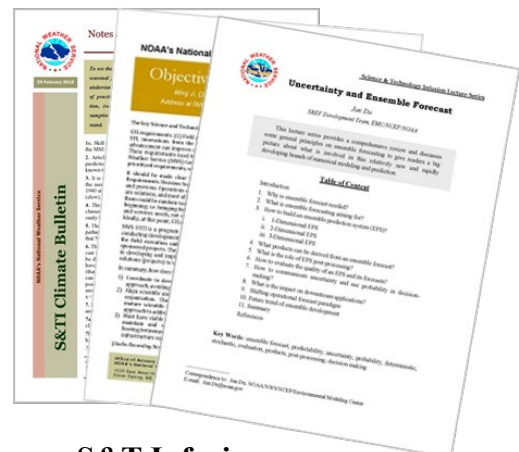
Featured Special Collections

(<https://vlab.noaa.gov/web/osti-r2o/climate>)

Climate Prediction Science and Technology Digest (2008 – 2022)



NOAA Climate Test Bed Joint Seminar Series Extended Summaries Collection Volume



S&T Infusion e-Lecture Series & Notes

