36th NOAA Climate Diagnostics and Prediction Workshop
Special Issue

Climate Prediction S&T Digest

December 2011

NWS Science &Technology Infusion Climate Bulletin Supplement

Inside this issue:

NOAA's National Weather Service

Office of Science and Technology
1325 East West Highway
Silver Spring, MD 20910

Climate Prediction Center
5200 Auth Road
Camp Springs, MD 20746

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, helping our partners in the research community to understand which R&D activities are needed to "shoot arrows at the target".

Science and Technology Infusion Climate Bulletin
http://www.nws.noaa.gov/ost/climate/STIP/index.htm

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

PREFACE

It is with great pleasure that the Climate Prediction Center and the Office of Science and Technology offer you this synthesis of the 36th Climate Diagnostics and Prediction Workshop. The CDPW continues to be an important way for the climate prediction community to stay connected. As is clearly evident in this digest, the climate community continues to make great strides in our ability to simulate and predict climate. The purpose of this digest is to help you stay informed while also ensuring that climate research advances are shared with the broader climate community and 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.

Finally, I would like to thank Dr. Jiayu Zhou of the Office of Science and Technology / NWS, for developing the digest concept and for being its champion. This partnership between OST and CPC is an essential element of NOAA climate services.

Wayne Higgins

Director, Climate Prediction Center
National Centers for Environmental Prediction
NOAA's National Weather Service

CONTENTS

OVERVIEW

The 36th Climate Diagnostics and Prediction Workshop was held in Fort Worth, Texas, on 3-6 October 2011. The workshop was hosted by the National Weather Service (NWS) Southern Region Headquarters and the National Climatic Data Center (NCDC), and cosponsored by the Climate Prediction Center (CPC) of the National Centers for Environmental Prediction and NWS Climate Services Division. The American Meteorological Society was a cooperating sponsor.

A diverse group of about 132 scientists from more than 70 domestic and international institutes gathered to explore current operational climate prediction capabilities, identify opportunities for advances, and discuss new products needed to support regional decision makers.

The workshop addressed the status and prospects for advancing climate monitoring, assessment, and prediction with emphasis in five major themes:

1. Changes in weather and climate extremes on timescales from daily to decadal;
2. Performance of NOAA coupled models on timescales from daily to decadal, including reanalysis, reforecasts and multi-model ensembles;
3. Status and prospects for improved understanding and more realistic simulation and prediction of drought/pluvial, including objective drought tools and impacts on water resources;
4. Prospects for improved understanding and prediction of warm season North American hydroclimate variability;
5. Development and delivery of climate information that meets the evolving needs of users, including business, government, resource managers, scientists, sectoral stakeholders, and private citizens.

This Workshop is continuing to grow and provides a stimulus for further improvements in climate prediction, monitoring, diagnostics, applications and services.

1. ANNUAL REVIEW OF WEAHRE & CLIMATE

AND CLIMATE OPERATION

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

The Climate Prediction Center's 2010-11 Seasonal Forecasts and a Look Ahead to 2011-12

Mike Halpert

Climate Prediction Center, NOAA/NWS/NCEP, Camp Springs, MD

The fall, winter and spring of late 2010 and 2011 witnessed a moderate to strong La Niña across the tropical Pacific Ocean, which shaped CPC's seasonal outlooks for those seasons. La Niña often results in drier-than-average conditions across much of the southern part of the continuous United States and wetter-than-average across parts of the North, particularly in the Pacific Northwest and in the Ohio Valley. Temperatures are often above average across the South and below average across the North, mainly from the northern Great Lakes region westward. These effects normally result in probabilistic seasonal forecasts which favor these impacts in these regions and the forecasts for the winter 2010-11 and spring 2011 were no exception (Figs 1 and 3).

Precipitation forecasts for September – November 2010 through April – June 2011 all scored at least 30% better than a climatological forecast, the longest streak (eight) of very successful forecasts since CPC began issuing forecasts in this style in 1995 (*i.e.* Fig. 1, right). The success of these forecasts was directly tied to the La Niña, as these forecasts heavily leveraged the expected impacts from the episode.

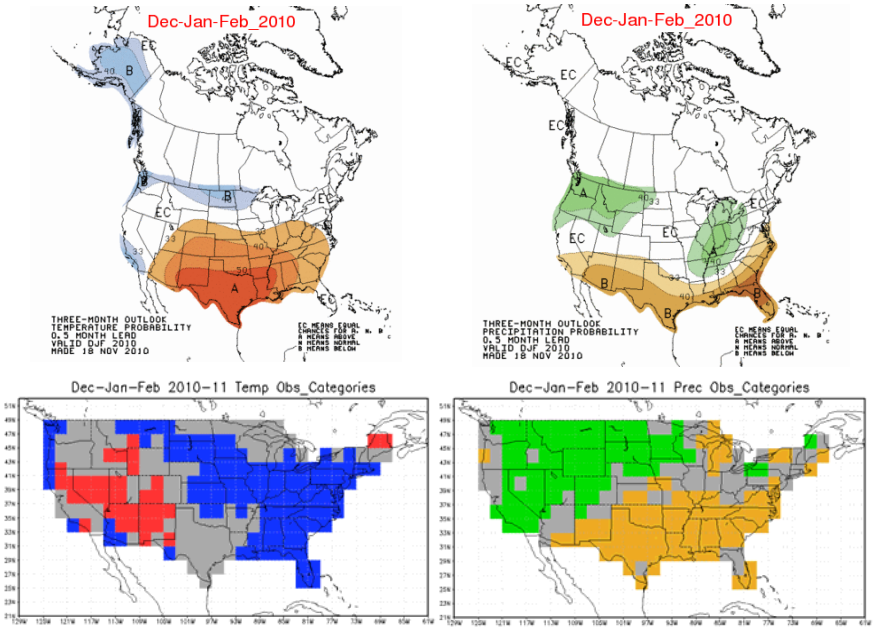


Fig. 1 December 2010 – February 2011 forecasts (top) and categorical verification (bottom).

In contrast, the temperature forecasts during the heart of the winter (November – January, December – February, and January – March) were not as successful (*i.e.* Fig. 1, left), with Heidke skill scores near or below zero for all 3 forecasts. The obvious question to ask is what caused the disparity in skill between the temperature and precipitation forecasts and the answer highlights some of the unpredictable factors that often influence the climate during a particular season. While the La Niña dominated the precipitation signal, a strong negative phase Arctic Oscillation (AO) teleconnection pattern developed in mid-November (Fig. 2) and persisted largely unabated through the middle of January, at which point another teleconnection pattern, the Pacific-North American (PNA) pattern became strongly positive for 2-3 weeks (not shown). The influence of these two teleconnection patterns during the winter, which included the second lowest December AO value (Fig. 2a) in the historical record dating back to 1950 (exceeded only by December 2009), resulted in temperatures during the winter that were either normal or below normal across much of the eastern two-thirds of the country. This result was at odds with CPC's seasonal forecasts, which favored above average temperatures across much of the southern and central parts of the nation.

Correspondence to: Mike Halpert, Climate Prediction Center, NCEP/NWS/NOAA, Camp Springs, MD 20746; E-mail: Mike.Halpert@noaa.gov.

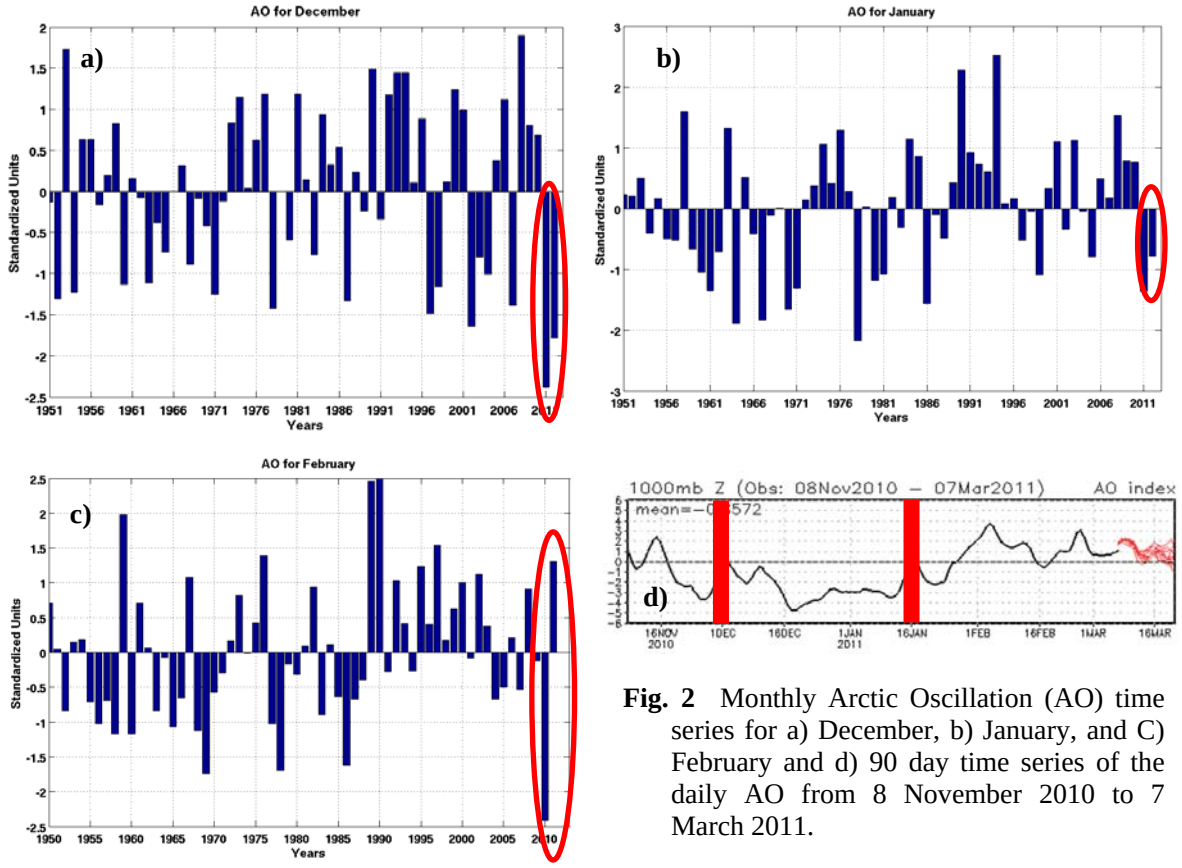


Fig. 2 Monthly Arctic Oscillation (AO) time series for a) December, b) January, and C) February and d) 90 day time series of the daily AO from 8 November 2010 to 7 March 2011.

By early February, the AO became positive and remained neutral or positive during the remainder of the winter (Fig. 2d) and throughout the spring. This transition, along with the continuation of La Niña conditions, resulted in a string of successful temperature forecasts beginning with February – April 2010 continuing into early summer. The verification for these forecasts was quite consistent with previous La Niña episodes, with colder-than-average temperatures observed across much of the northern Plains, the northern Rockies, the Pacific Northwest and warmer-than-average temperatures across much of the southern part of the nation during spring (Fig. 3).

As we head into the winter of 2011-12, La Niña conditions have again developed throughout the Tropical Pacific, as sea surface temperature departures more than 0.5°C below average spanned the central and eastern tropical Pacific. This has helped to shape the outlooks for the winter, with CPC's seasonal outlooks again generally favoring wetter and colder than average conditions across much of the northern part of the country and drier and warmer than average conditions across much of the South.

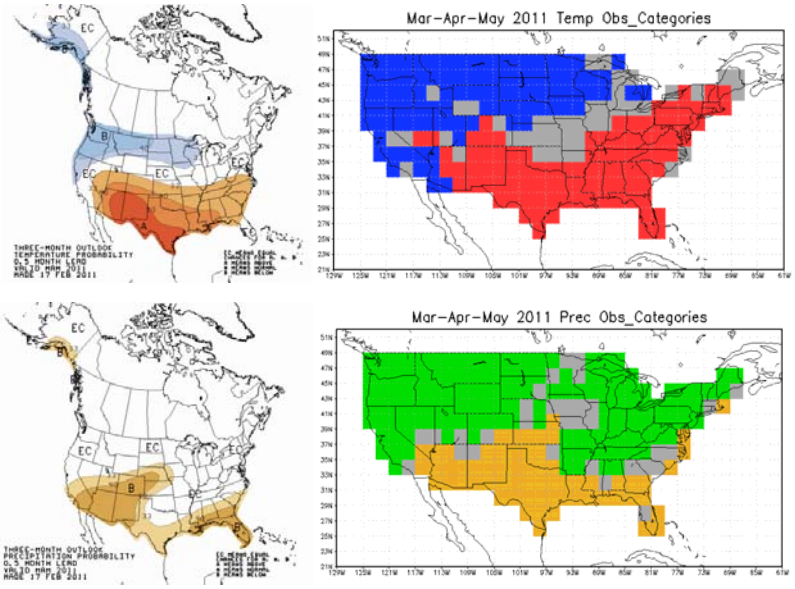


Fig. 3 March – May 2011 forecasts (left) and categorical verification (right) of temperature (top) and precipitation (bottom).

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Tropical ISO and Extratropical Extreme Weather during the 2009-2011 ENSO Cycle

Bin Wang

Department of Meteorology and International Pacific Research Center,
School of Ocean and Earth Science and Technology, University of Hawaii at Manoa Honolulu

J-.Y. Moon, Kazuyoshi Kikuchi, J.-Y. Lee, and S.-Y. Yim

International Pacific Research Center,
School of Ocean and Earth Science and Technology, University of Hawaii at Manoa Honolulu

The tropical intraseasonal oscillation (ISO) shows distinct variability centers and propagation patterns between boreal winter and summer. To accurately describe and monitoring the state of the ISO at any particular time of a year, a *bimodal ISO index* is used. It consists of Madden-Julian Oscillation (MJO) mode with predominant equatorial eastward propagation and Boreal Summer ISO (BSISO) mode with prominent northward propagation and large variability in off-equatorial monsoon tough regions. The spatial-temporal patterns for the MJO and BSISO modes are identified with the extended empirical orthogonal function analysis of 31 years (1979-2009) OLR data for the December-January-February and June-July-August period, respectively. The details are referred to Kikuchi *et al.* (2010). The dominant mode of the ISO at any given time can be judged by the proportions of the OLR anomalies projected onto the two modes. The proposed bimodal ISO index provides objective and quantitative measures on the annual and interannual variations of the predominant ISO modes. From December through April the MJO mode dominates while from June through October the BSISO mode dominates. May and November are transitional months when the predominant mode changes from one to the other (Fig. 1). The fractional variance reconstructed based on the bimodal index is significantly higher than the counterpart reconstructed based on the Wheeler and Hendon's index (Wheeler and Hendon 2004). The bimodal ISO index provides a reliable real time monitoring skill. The method and results provide critical information in assessing models' performance to reproduce the ISO and developing further research on predictability of the ISO and are also useful for a variety of scientific and practical purposes.

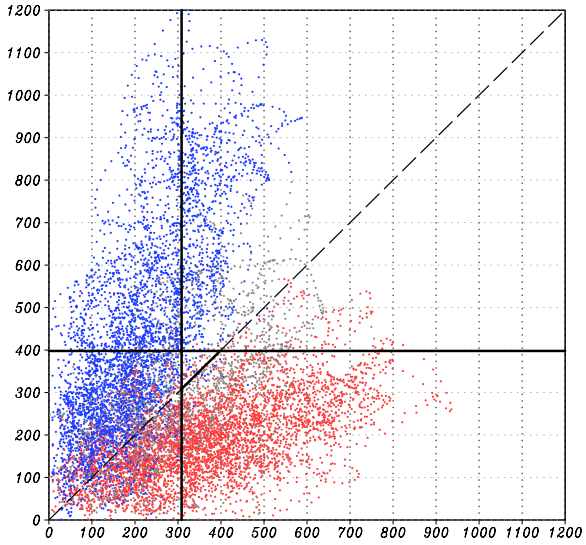


Fig. 1 Scatter plot of ISO amplitude in terms of the BSISO mode (x-axis) and the MJO mode (y- axis) for the period 1979–2009. Solid lines parallel to the x and y axes represent one standard deviation of each ISO mode during the period each EEOF analysis was performed. Dashed line represents the situation when both the MJO mode and the BSISO mode have the same variance. Three seasons are marked by different colors (June–October in red, December–April in blue, otherwise in gray).

The tropical ISO during the past ENSO cycle (December 2009 to August 2011) is analyzed using the bimodal index. The DJF 2009/2010 is a mature phase of El Niño while the DJF 2010/2011 is a mature phase of La Niña. The JJA 2010 is a developing phase of La Niña while the JJA2011 is a decaying phase of La Niña. Due to strong regulation of the ENSO, the MJO is amplified during DJF 09/10 at the El Niño mature phase, while

Correspondence to: Bin Wang, Department of Meteorology and International Pacific Research Center, School of Ocean and Earth Science and Technology, University of Hawaii at Manoa, Honolulu HI 96822; E-mail: wangbin@hawaii.edu.

suppressed during DJF 10/11 when the La Niña matures (Fig. 2).

The pulsation of the strong MJO during the winter of 2009/10 had strongly modulated the winter storms in the United States through an atmospheric teleconnection, resulting in a number of record-breaking snowfall events registered in the eastern United States (Moon *et al.* 2011). The intraseasonal variation of OLR had reached maximum strength over the eastern subtropical Pacific near Mexico and the second largest over the equatorial central Pacific since 1979/80. The convection over these two regions experienced a remarkable wet-dry-wet cycle during the 60-day period from late December to mid-February; correspondingly, the daily snowfall over the eastern US exhibited a cohesive wet-dry-wet cycle (Fig. 3). As the MJO convection reached the central Pacific, a teleconnection pattern extends to North America, resulting in a westward-tilted deep anomalous trough anchored over the eastern US, producing a low-level

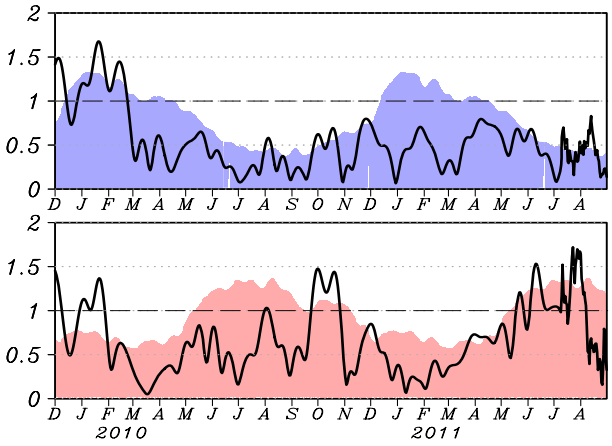


Fig. 2 Bimodal description of the tropical ISO for the period December 2009 to August 2011. The upper and lower panels show the amplitudes of the MJO and BSISO modes, respectively. The shading represents climatological annual cycle. Note the large amplitude of MJO during DJF 2009/2010.

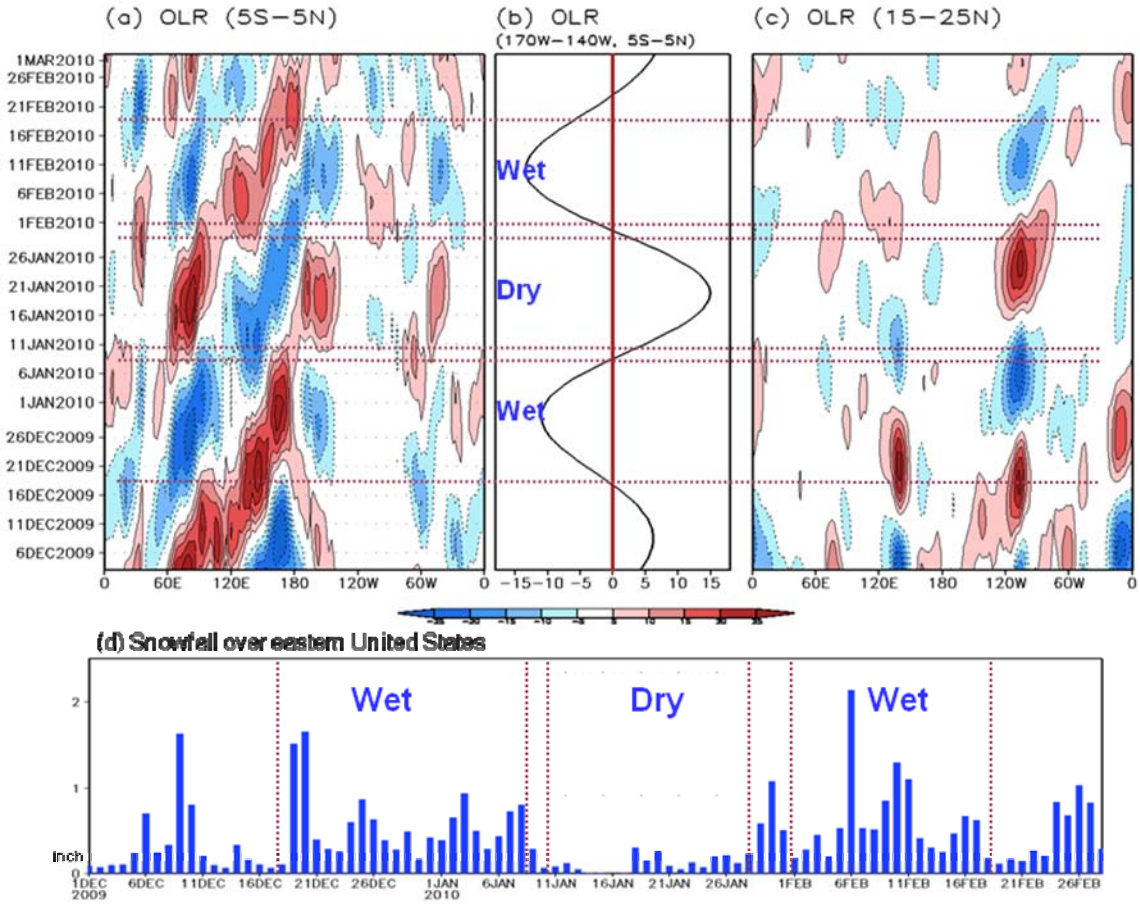


Fig. 3 MJO during DJF 2009/10 and snowfall in eastern US. (a) 30-60 day OLR anomalies along the equator (5ºS-5ºN) as a function of time. (b) Time series of the 30-60 day OLR anomalies averaged over the equatorial central Pacific (5ºS-5ºN, 170ºW-140ºW). (c) The same as in (a) except along the band between 5ºN and 15ºN. (d) snowfall amount averaged over the eastern US.

pressure dipole anomaly with an anticyclone (cyclone) centered at the US west (east) coast (Fig. 4). The convection over the Indian Ocean varied in phase with the central Pacific convection, reinforcing the extratropical atmospheric teleconnection pattern. As a result, the enhanced high-latitude cold air penetrated southward, affecting the central and eastern US. Meanwhile, warm moist air was transported from the tropical central Pacific through Mexico to the southern US along with the upper-level subtropical westerly jet. These enhanced warm moist air from the tropics and the cold air transportation from the high-latitude are also supported by the existing El Niño and negative AO (NAO), respectively, which reinforced the MJO teleconnection over the eastern US during wet cycle. As such, the eastern US was located in a convergence zone between the enhanced cold air from the high-latitude and the warm moist air supplied from the subtropics, resulting in favorable conditions for extremely heavy snowfall.

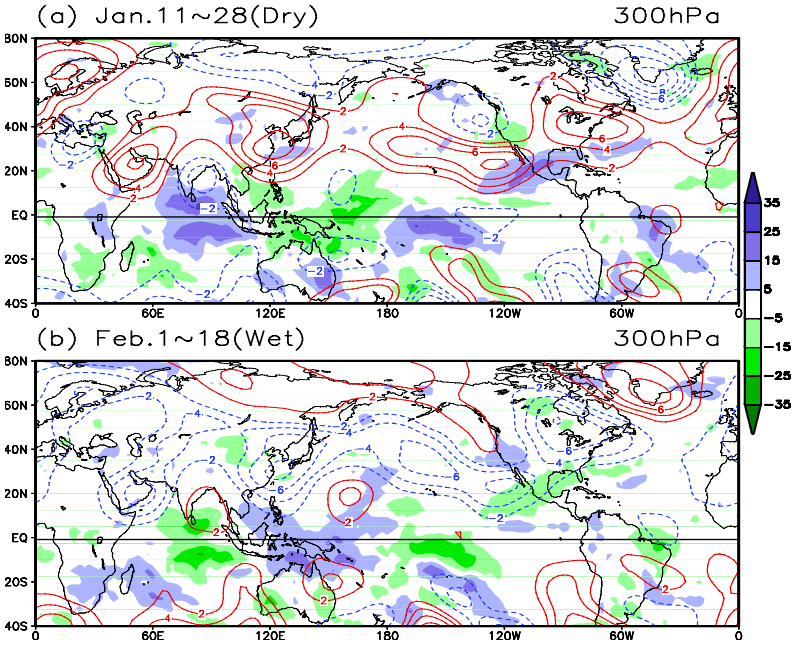


Fig. 4 Intraseasonal anomalies of OLR and 300hPa streamfunction on the days when convection over the tropical central Pacific in (a) dry and (b) wet phases. The OLR is shaded. The 300hPa streamfunction is drawn in contour.

During the La Niña development summer, *i.e.*, JJA 2010, the BSISO is generally weak except one event in October 2010, which is related to a supper typhoon activity in the western North Pacific. During the La Niña decaying phase (JJA 2011), the BSISO activity tends to be normal. There was a strong event in June 2011. This event had caused flooding conditions over the Meiyu-Baiu front in East Asia. Seoul also experienced unusual floods.

References

Moon, J-Y, B. Wang, and K.-J. Ha, 2011: ENSO regulation of MJO teleconnection. *Climate Dynamics*. In press.

Kikuchi, K., B. Wang, and Y. Kajikawa, 2011: Bimodal representation of the tropical intraseasonal oscillation. *Clim Dyn*, DOI 10.1007/s00382-011-1159-1

Wheeler, Matthew C., and Harry 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.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Skill of Real-time Seasonal ENSO Model Predictions during 2002-2011 — Is Our Capability Increasing?

Anthony G. Barnston [1], Michael K. Tippett [1], Michelle L. L'Heureux [2], Shuhua Li [1], and David G. Dewitt [1]

[1] *International Research Institute for Climate and Society, The Earth Institute of Columbia University, Palisades, New York*

[2] *Climate Prediction Center, NCEP/NWS/NOAA, Camp Springs, Maryland*

1. Introduction

In this study, real-time model predictions of ENSO conditions during the 2002-2011 period are evaluated and compared to skill levels documented in studies of the 1990s. ENSO conditions are represented by the Niño3.4 SST index in the east-central tropical Pacific. The skills of 20 prediction models (12 dynamical, 8 statistical), that have been displayed on the ENSO prediction plume of the International Research Institute for Climate and Society (IRI) since 2002, are examined. Over the last two to three decades, our ability to predict ENSO variations at short and intermediate lead times has presumably gradually improved due to improved observing and analysis/assimilation systems, improved physical parameterizations, higher spatial resolution, and better understanding of the tropical oceanic and atmospheric processes underlying the ENSO phenomenon. Studies in the 1990s showed moderate ENSO prediction capability, with forecast versus observation correlations of about 0.6 for 6-month lead predictions for the Niño3.4 region (Barnston *et al.* 1994). This study reviews the recent model performances, and reexamines the question of the relative performance of dynamical and statistical models. We also compare the skills of the 9 years of real-time predictions to those of longer-term (30-year) hindcasts from some of the same models. The ENSO prediction models studied here are listed in Table 1.

The ENSO predictions issued each month from February 2002 through January 2011 are examined for multiple lead times for future 3-month target (*i.e.*, predicted) periods. Figure 1 shows the variability of the Niño3.4 anomaly from 1981 to 2011, highlighting the recent 9-year period of the current study. Although there were some moderate ENSO events, no very strong events occurred. The last target period is January-March 2011, while the earliest target period is February-April 2002 for the shortest lead time and October-December 2002 for the longest lead time. The forecast data from a given model consist of a succession of running 3-month mean SST anomalies with respect to the climatological means for the respective predicted periods, averaged over the Niño3.4 region.

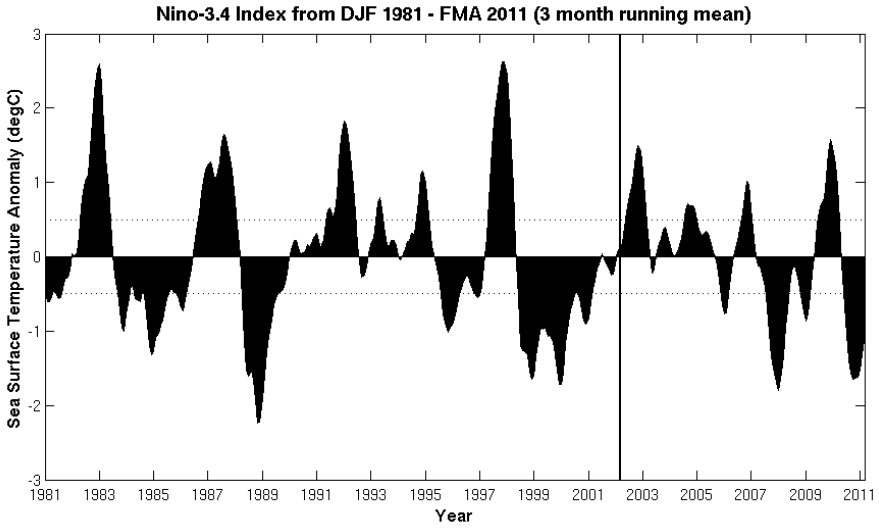


Figure 1 Time series of running 3-month mean SST anomaly with respect to the 1981-2010 period climatology in the Niño3.4 region for 1981-2011, highlighting the 2002-2011 study period.

Correspondence to: Anthony G. Barnston, International Research Institute for Climate and Society, The Earth Institute of Columbia University, Palisades, New York; E-mail: Tonyb@iri.columbia.edu.

Predicted periods begin with the 3-month period beginning immediately after the latest available observed data, and continue for increasing lead times until the longest lead time provided by the given model, to a maximum of 9 running 3-month periods. Here, lead time is defined by the number of months of separation between the latest available observed data and the beginning of the 3-month forecast target period. Although anomalies were requested to be with respect to the 1971-2000 climatology, some prediction anomalies were with respect to means of other periods, such as from 1982 to the early 2000s for some dynamical predictions. Adjustments for these discrepancies were not conducted, nor were model bias corrections attempted. Only the ensemble mean of the dynamical model forecasts is considered as a deterministic prediction.

Dynamical Models	**Model type**
NASA GMAO	Fully coupled
NCEP CFS	Fully coupled
Japan Meteorological Agency	Fully coupled
Scripps Hybrid Coupled Model (HCM)	Comprehensive ocean, statistical atmosphere
Lamont-Doherty	Intermediate coupled
Australia POAMA	Fully coupled
ECMWF	Fully coupled
UKMO	Fully coupled
Korea Met. Agency SNU	Intermediate coupled
Univ. Maryland ESSIC	Intermediate coupled
IRI ECHAM/MOM	Fully coupled, anomaly coupled
COLA Anomaly	Anomaly coupled
COLA CCSM3 (too short a record)	Fully coupled
Météo France (too short a record)	Fully coupled
Japan Frontier FRCGC (short record)	Fully coupled
Statistical Models	**Method and predictors**
NOAA/NCEP/CPC Markov	Markov: Preferred persistence and transitions in SST and sea level height fields
NOAA/ESRL Linear Inverse Model (LIM)	Refined POP: Preferred persistence and transitions within SST field
NOAA/NCEP/CPC Constructed Analogue (CA)	Analogue-construction of current global SSTs
NOAA/NCEP/CPC Canonical Correlation Analysis (CCA)	Uses SLP, tropical Pacific SST and sub-surface temperature (not used beginning in 2010)
NOAA/AOML CLIPER	Multiple regression from tropical Pacific SSTs
Univ. British Columbia Neural Network (NN)	Uses sea level pressure and Pacific SST
Florida State Univ. Multiple Regression	Uses tropical Pacific SST, heat content, winds
UCLA TDC Multi-level Regression	Uses 60N-30S Pacific SST field

Table 1 Dynamical and statistical models whose forecasts for Niño3.4 SST anomaly are included in this study. Note that some models were introduced during the course of the study period, or replaced a predecessor model.

The Reynolds-Smith version 2 optimal interpolation (OI) observed SST data averaged over the Niño3.4 region (5°N-5°S, 120°-170°W) is used as the verification data, using the 1981-2010 period to define the anomalies.

2. Results

a. Real-time Predictive Skill of Individual Models

Time series of the running 3-month mean observed SST anomalies in the Niño3.4 region and the corresponding predictions by 23 prediction models at 0-, 2-, 4- and 6 month lead times are shown in Fig. 2, showing that the models generally predicted the variations of ENSO with considerable skill at short lead times, and decreasing skill levels with increasing lead times. Figure 3 shows the temporal correlation between model predictions and the corresponding observations as a function of target season and lead time, with a separate panel for each model. The correlation skill patterns of the models appear roughly comparable. All indicate a northern spring predictability barrier (Jin *et al.* 2008), with short lead prediction skills having a relative minimum for northern summer, extending to later seasons at longer lead times. Relative to the statistical models, Fig. 3 shows higher correlation skills by many of the dynamical models for seasons in the middle of the calendar year that generally have lowest skill. By contrast, for seasons having highest skills (*e.g.* northern winter target seasons at short to moderate lead times), skill differences among models and between model types appear small.

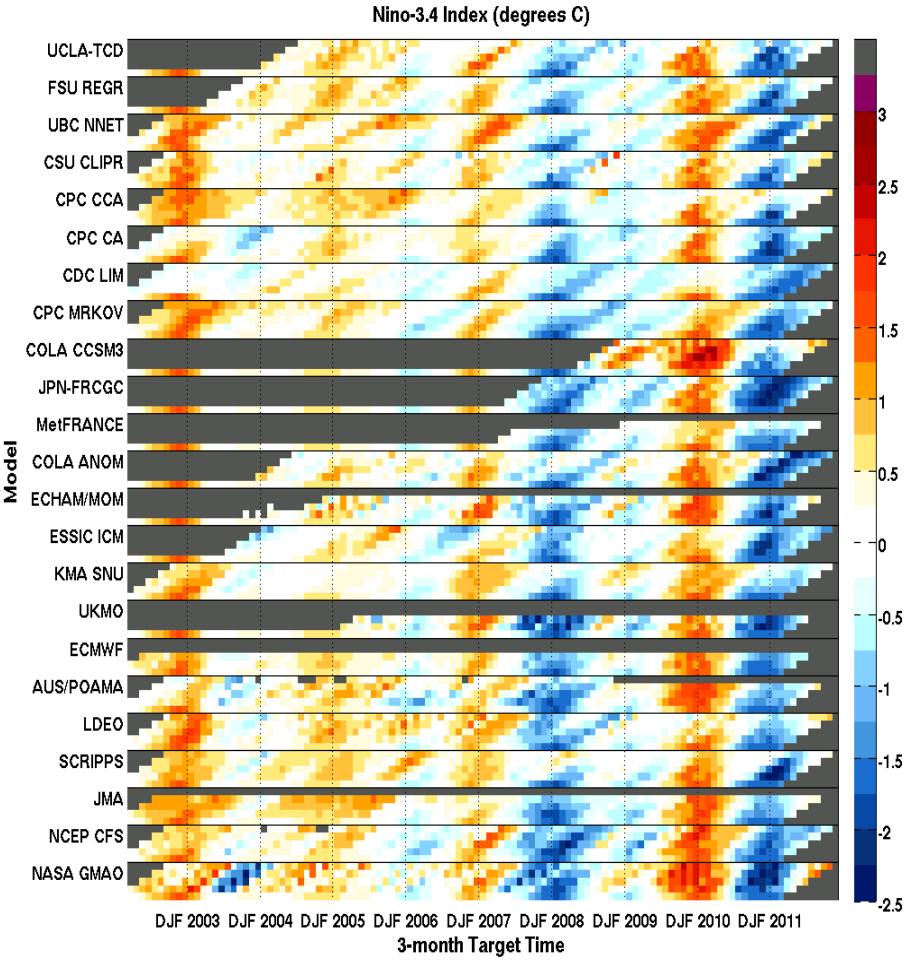


Figure 2 Time series of running 3-month mean Niño3.4 SST observations (°C anomaly), and corresponding model predictions for the same 3-month period from earlier start times at 0-, 2-, 4- and 6-month leads. Data for each model are separated by thin black horizontal lines. The first 8 models at the top are statistical models. For each model, the bottom row shows the observations, and the four rows above that row show predictions at the four increasing lead times. Vertical dotted lines demarcate calendar year changes, separating Nov-Dec-Jan from Dec-Jan-Feb. Observations span from Feb-Mar-Apr 2002 to Jan-Feb-Mar 2011, while forecasts at longer lead times start and end with later seasons. Black shading indicates missing data.

Figure 4 shows individual model correlation skills as a function of lead time for all seasons combined, while the top and bottom panels of Fig. 5 show skills for the pooled target seasons of NDJ[1] , DJF and JFM, and for MJJ, JJA and JAS, respectively. Overall, model correlation skills at 6-month lead range anywhere

[1] Seasons are named using the first letter of the three constituent months; e.g. DJF refers to Dec-Jan-Feb.

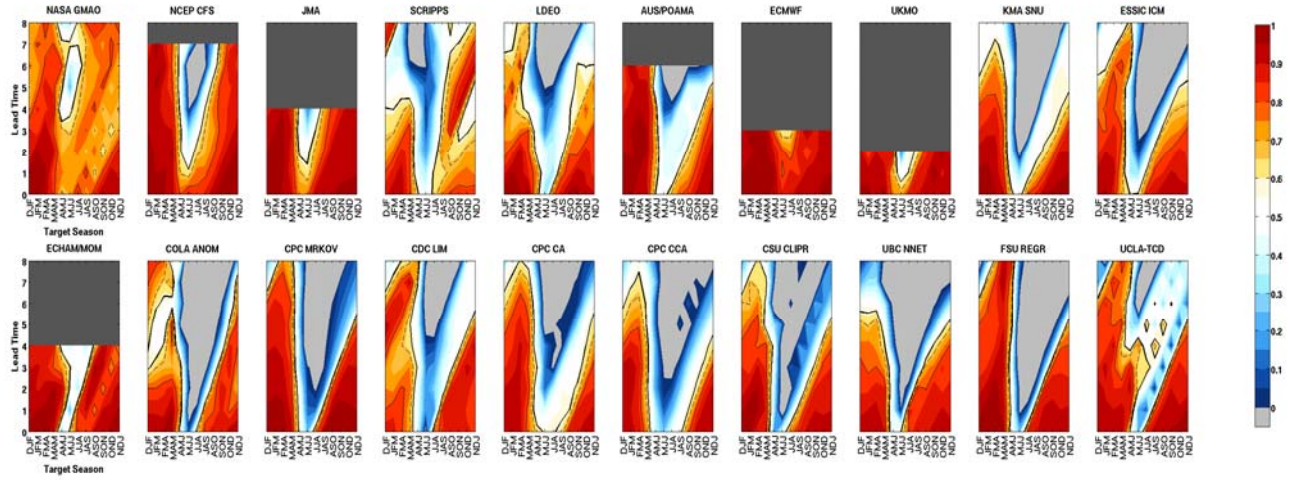


Figure 3 Temporal correlation between model forecasts and observations as a function of target season (horizontal axis) and lead time (vertical axis). Each panel highlights one model. The first 12 models are dynamical, followed by 8 statistical models. Thick solid contour shows the 90% significance level, dashed contour the 95% level, and thin solid contour the 99% level

from zero to about 0.7 for all seasons combined, while predictions for the northern winter season range from 0.4 to 0.9, and for the northern summer season from below zero to 0.55. The model skill levels for all seasons combined (Fig. 4) differ from one another noticeably at all lead times. Averaged over all seasons, skills average somewhat lower than the 0.6 level found at 6-month lead in earlier studies. However, a small number of current models, some of which do not predict out to 6 months lead, have shorter-lead skill levels that would exceed a 0.6 correlation if their forecast range were extended, and if their skill followed a downward slope with increasing lead time averaging that shown by other models having longer maximum lead times. Examples of models with such good or potentially good skill include ECMWF, NASA-GMAO, JMA; NCEP-CFS skill approximately equals 0.6. However, two caveats in the comparison of skills of today's models against models of 10 to 20 years ago include (1) the ENSO variability during the 2002-2011 period will be demonstrated to have been more difficult to predict than that over the 1981-2011 in general; and (2) the current set of predictions were made in real-time, while those examined in previous studies were partly hindcasts. Both factors will be examined further below.

One reasonably might ask whether the skill differences at any lead time are sufficient, for a 9-year period, to statistically distinguish among the performance levels of some of the models. Because ENSO episodes last up to a year, we assume (perhaps conservatively) that we have only about one independent sample per year. The existence of statistically significant differences between skills of any pair of individual models requires very large sample skill differences—larger than those found here. However, the statistical significance of skill differences between dynamical and statistical model types is more tractable, and is addressed below.

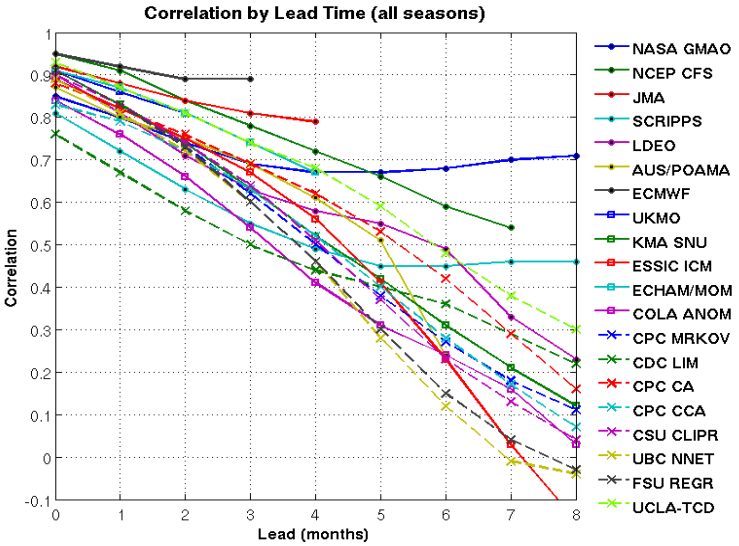


Figure 4 Temporal correlation between model forecasts and observations for all seasons combined, as a function of lead time. Each line highlights one model. The 8 statistical models are shown with dashed lines and the cross symbol.

The correlation between model

predictions and observations reflects purely the discrimination ability of the models, since biases of various types do not affect this metric. However, such prediction biases (*e.g.*, calibration problems involving the mean or the amplitude of the predictions) are also part of overall forecast quality, despite being correctible in many cases. To assess performance in terms of both calibration and discrimination, root mean square error (RMSE) is examined. Here the RMSE is standardized for each season individually. Standardization scales RMSE so that climatology forecasts (zero anomaly) result in the same RMSE-based skill (of zero) for all seasons, and all seasons' RMSE contribute equally to a seasonally combined RMSE. Figure 6 shows RMSE as a function of lead time for all seasons together. The ECMWF model has the lowest RMSE over its range of lead times. For lead times greater than 2 months, persistence forecasts have higher RMSE than that of any of the models. There is clearly some comparability between correlation skill (Fig. 4) and RMSE (Fig. 6), with models having highest correlation tending to have low RMSE. However, exceptions are discernible, due to the effects of mean biases and amplitude biases (not shown).

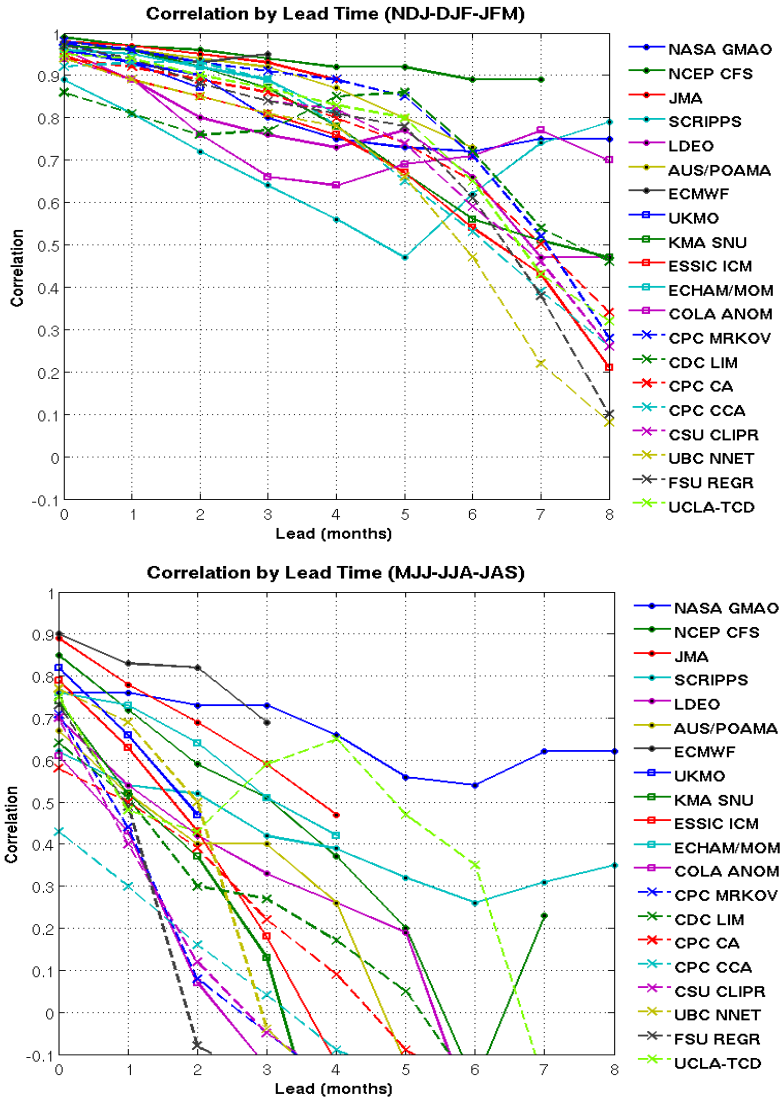


Figure 5 (top) Temporal correlation between model forecasts and observations for Nov-Dec-Jan, Dec-Jan-Feb and Jan-Feb-Mar as a function of lead time. Each line highlights one model. The 8 statistical models are shown with dashed lines and the cross symbol. (bottom) As in top, but for May-Jun-Jul, Jun-Jul-Aug and Jul-Aug-Sep.

Establishing statistical significance of skill differences between dynamical and statistical models for specific times of the year is difficult for a 9-year study period. However, the fairly large number of models can be used to help overcome the short period length. Models are ranked by correlation skill for each season and lead time separately, using the 9 year sample. Systematic differences in the ranks of the dynamical and statistical models are identified using the Wilcoxon rank sum test (Wilcoxon 1945). Additionally, the average correlation of the dynamical and statistical models is compared using a standard t-test, applied to the Fisher Z equivalents of the correlations (Ramseyer 1979). The p-values resulting from these two statistical approaches are shown in Table 2. Although the difference-in-means test generally yields slightly more strongly significant results than the rank sum test, the season/lead patterns of the two approaches are similar. Significant differences, in which dynamical models tend to outperform statistical models, are found at short lead time for the target periods near May-Jul-Jul, the seasons just following (and most strongly affected by) the northern spring predictability barrier. This significance pattern migrates to later target periods with increasing lead time, following the target periods corresponding to the fixed forecast start times of April or May. For forecasts whose lead times do not traverse the northern spring barrier, statistical versus dynamical skill differences are not significant.

Wilcoxon rank sum test (field significance p=0.034)													
Lead	**DJF**	**JFM**	**FMA**	**MAM**	**AMJ**	**MJJ**	**JJA**	**JAS**	**ASO**	**SON**	**OND**	**NDJ**	**All**
0	0.32	0.19	0.76	0.28	**0.01**	**0.01**	0.09	0.95	0.22	0.41	0.95	0.76	0.70
1	0.88	0.25	0.22	0.32	0.06	**0.003**	**0.02**	0.17	-0.68	0.34	0.68	1.00	0.64
2	1.00	0.76	0.32	0.19	0.17	**0.04**	**0.01**	**0.01**	0.54	-0.73	0.38	0.94	0.22
3	1.00	-0.93	0.32	0.14	0.36	0.19	0.14	**0.01**	**0.01**	0.25	1.00	-0.74	0.12
4	-0.33	-0.37	0.79	0.29	0.48	0.48	0.18	0.25	**0.01**	**0.004**	0.36	-0.21	0.16
5	-0.09	-0.29	-0.40	0.67	0.92	0.60	0.67	0.30	0.34	**0.03**	**0.002**	0.92	0.21
6	0.60	-0.60	-0.75	1.00	0.40	0.46	0.75	-0.83	0.75	0.46	**0.05**	**0.05**	0.25
7	**0.02**	1.00	-0.35	1.00	-0.64	0.20	0.82	0.91	0.70	0.73	0.25	**0.03**	0.35
8	**0.05**	**0.02**	1.00	-0.52	-0.44	1.00	0.19	-0.61	-0.66	-0.88	1.00	**0.05**	0.61

t-test for mean difference (field significance p=0.026)													
Lead	**DJF**	**JFM**	**FMA**	**MAM**	**AMJ**	**MJJ**	**JJA**	**JAS**	**ASO**	**SON**	**OND**	**NDJ**	**All**
0	0.27	0.19	0.65	0.22	**0.003**	**0.001**	0.06	0.48	0.41	0.65	0.74	0.46	0.49
1	0.50	0.37	0.12	0.16	**0.04**	**0.000**	**0.01**	0.10	-1.00	0.32	0.73	-0.85	0.29
2	-0.90	0.54	0.33	0.06	0.23	**0.04**	**0.001**	**0.01**	0.16	-0.86	0.29	0.93	0.12
3	-0.98	0.65	0.13	0.13	0.71	0.20	0.07	**0.002**	**0.001**	0.11	-0.90	-0.65	0.09
4	-0.35	-0.39	0.82	0.40	0.26	0.78	0.29	0.12	**0.002**	**0.000**	0.22	-0.19	0.11
5	-0.16	-0.47	-0.31	0.56	-0.93	0.55	0.87	0.17	0.18	**0.004**	**0.001**	0.66	0.18
6	0.34	-0.80	-0.73	-0.62	0.34	0.30	0.70	0.97	0.66	0.28	**0.01**	**0.02**	0.18
7	**0.01**	0.37	-0.39	-0.60	-0.67	0.26	0.42	0.65	0.42	0.51	0.17	**0.01**	0.15
8	**0.04**	**0.02**	0.52	-0.37	-0.45	-0.83	0.13	0.77	0.48	0.66	0.70	0.29	0.34

Table 2 Statistical significance results (2-sided p-values), by target season and lead time, for differences in temporal correlation skill of dynamical versus statistical models: (top) Wilcoxon rank sum test for correlation skills, and (bottom) t-test of difference in means of Fisher Z equivalents of the correlations skills. Entries statistically significant at the 0.05 level are shown in bold. Negative sign indicates cases when statistical models have higher ranks (or means) than dynamical models. P-values are shown to 3 decimal places when p<0.005; 0.000 indicates p-value <0.0005.

Although significant differences are noted for specific seasons and leads, there is a multiplicity of candidate season/lead combinations, and 5% of the 108 candidates (*i.e.*, 5 or 6 of them) are expected to be significant by chance. In the case of the Wilcoxon test, 20 entries are significant, and for the difference-in-means test 20 entries are significant. To assess the field significance of the collective result (Livezey and Chen 1983), Monte Carlo simulations are conducted in which the model type is randomly shuffled 5,000 times, maintaining the actual number of dynamical and statistical models for the given lead time, and the set of local significances is regenerated. Using the sum of the z or t values of all 108 cells as the test statistic, the percentage of the 5,000 randomized cases that exceeds the actual case is determined. The z or t values are

taken as positive when the correlation of the dynamical models exceeds that of the statistical models, and negative for the opposite case. Resulting field significances are 0.034 and 0.026 for the Wilcoxon rank test and t-test, respectively, indicating significantly low probabilities that the set of local significances occurred accidentally. This finding suggests that the circumstance under which local significance is found, namely forecasts impacted by the northern spring predictability barrier being more successful in dynamical than statistical models, is meaningful and deserves fuller explanation.

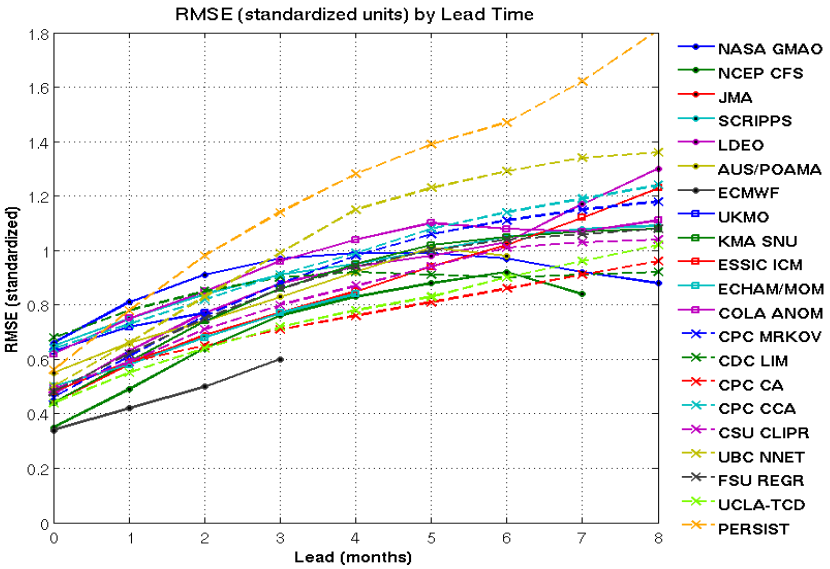


Figure 6 RMSE in standardized units, as a function of lead time for all seasons combined. Each line highlights one model. The 8 statistical models and the persistence model are shown with dashed lines and the cross symbol.

A likely reason that dynamical models are better able to predict ENSO through the time of year when transitions (dissipation of old events and/or development of new events) typically occur is their more effective detection, through the initial conditions, of new evolution in the ocean-atmosphere system on a relatively short (intramonth) time-scale—evolution that may go unnoticed by statistical models that use monthly or seasonal means for their predictor variables. Statistical models might be able to compete better against dynamical models if they used finer temporal resolution, such as weekly means.

Statistical models need long histories of predictor data to develop their predictor-predictand relationships. This need presents a problem in using the 3-dimensional observations in the tropical Pacific, such as the data from the Tao-Triton array, dating from the 1990s. (However, some subsurface tropical Pacific data do date back 10 or more years earlier in the eastern portion of the basin, and are available in the GODAS product.) This shorter data history precludes robust empirical definition of their predictive structures, and thus they are often omitted in statistical models. Although comprehensive dynamical models require a data history sufficient for verification and as a basis for defining anomalies, such a history is not basic to their functioning, and real-time predictions are able to take advantage of improved observing systems as they become available, potentially resulting in better initial conditions. While use of such crucial data suggests that dynamical models should be able to handily outperform statistical models, dynamical models have been burdened by problems such as initialization errors related to problems in data analysis/assimilation, and biases or drifts stemming from imperfect numerical representation of critical air-sea physics and parameterization of small-scale processes. As these weaknesses have improved, some comprehensive dynamical models have begun demonstrating their higher theoretical potential. This improvement will likely continue (Chen and Cane 2008).

b. Real-time Predictive Skill versus Longer-Period Hindcast Skill

Because 9 years is too short a period from which to determine predictive skill levels with precision, one reasonably might ask to what extent the performance levels sampled here could be expected to hold for future predictions. To achieve more robust skill estimates, a commonly used strategy is to increase the sample of predictions by generating retrospective hindcasts—"predictions" for past decades using the same model and procedures as in real time, to the extent possible. Cross-validation schemes are often used with statistical models, where varying sets of one or more years are withheld from the full data set, and the remaining years are used to define the prediction model which then is used to forecast the withheld year(s). In practice, there is

no comparable procedure applied in dynamical model development, and model parameter choices are often made using the same data used to evaluate skill.

Fourteen of the 20 models whose 9-year real-time forecast performance was discussed above (6 dynamical, 8 statistical) have produced hindcasts for the approximately 30-year period of 1981 (or 1982) to 2011. To assess the consistency of their skills during the longer period and the 9-year period of real-time predictions, the temporal correlation between hindcasts and observations is examined as a function of target seasons and lead time. Figure 7 shows a comparison of the correlation skills for the 9-year real-time predictions (as in Fig. 3) and the 30-year hindcasts for the subset of models having both data sets. Although the correlation plots are roughly similar, inspection shows generally higher hindcast skill levels for all of the models. Why do the hindcasts have higher skills? One explanation is that the 2002-2011 period may have been more difficult to predict than most of the longer period. Another explanation is that skills tend to be higher in hindcasts than in real-time predictions because the cross-validation designs may still allow inclusion of some artificial skill.

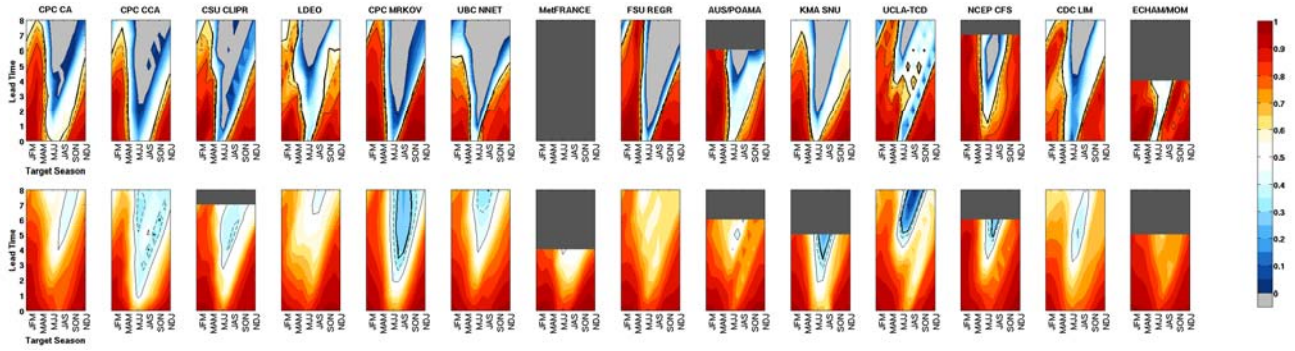


Figure 7 Temporal correlation between model forecasts and observations as a function of target season and lead time for (top) real-time forecasts (as in Fig. 3) and (bottom) hindcasts for the 1981-2010 period for models having long-term hindcasts. Thick solid contour shows the 90% significance level, dashed contour the 95% level, and thin solid contour the 99% level for sample sizes of 9 (top) and 30 (bottom).

To assess the relative difficulty of the recent 9-year period, the time series of uncentered correlation skills[2] of sliding 9-year periods, each phased 1 year apart, are examined for the 22 running periods within 1981-2010. The resulting time series of correlation are shown in Fig. 8 (top), for lead times of 3 months and 6 months. It is clear that for all models, and for both lead times, the 2002-2010 period, as well as the early to middle 1990s, posed a greater predictive challenge than most of the last three decades. As noted earlier in Fig. 1, one distinguishing feature of the 2002-2011 period is a lower amplitude of variability (no very strong events). The feature may be expected to reduce the upper limit of correlation skill by reducing the signal part of the signal-to-noise ratio. If the noise component remains approximately constant, and signal strength is somewhat restricted as during 2002-2010, then the correlation is reduced. The bottom inset of Fig. 8 (top) shows the 9-year running standard deviation of the observed Niño3.4 SST anomalies, with respect to the 1981-2010 mean. The correlation between the running standard deviation and the model average skill is about 0.8 for both 3- and 6-month lead predictions, confirming a strong relationship between signal strength and correlation skill.

The average of the anomaly of the 2002-2011 correlation with respect to that over 1981-2010 is approximately -0.14 (0.61 vs. 0.75) for 3 month lead forecasts and -0.23 (0.42 vs. 0.65) for 6-month lead forecasts. The -0.23 “difficulty anomaly” for 6 month lead forecasts is of greater magnitude than the deficit in skill of the real-time predictions during 2002-2011 compared with the approximately 0.6 skill level found in earlier studies, suggesting that today’s models would slightly (0.65 versus 0.60) outperform those of the 1990s if the decadal fluctuations of the nature of ENSO variability were taken into account.

[2] For the uncentered correlation, the 9-year means are not removed, so that standardized anomalies with respect to the 30-year means, rather than the 9-year means, are used in the cross-products and the standard deviation terms.

To examine the signal versus skill relationship with more precision, a 3-year time window is used in Fig. 8 (bottom), the bottom inset again indicating the running standard deviation. Within the 2002-2010 period, the subperiod of 2003-2007 is a focal point of low skill and low variability. The correlation between the 3-year running standard deviation and model average skill is again about 0.8 for both 3- and 6-month lead predictions, confirming a strong linkage between signal strength and correlation.

A second cause of the recent real-time predictions having lower correlation skill than the 30-year hindcasts is that using a period for which the verifying observations exist may permit inclusion of some artificial skill not available in real-time predictions. Attempts to design the predictions in a manner simulating the real-time condition (*e.g.* cross-validation) reduce artificial skill, but subtle aspects involving predictor selection often prevent its total elimination. Another impediment to the skill of real-time predictions includes such unavoidable inconveniences as delays in availability of predictor or initialization data, computer failure or other unforeseen emergencies, or human error. While this factor may seem minor, experience with the ENSO prediction plume has shown that such events occur more than once in a while.

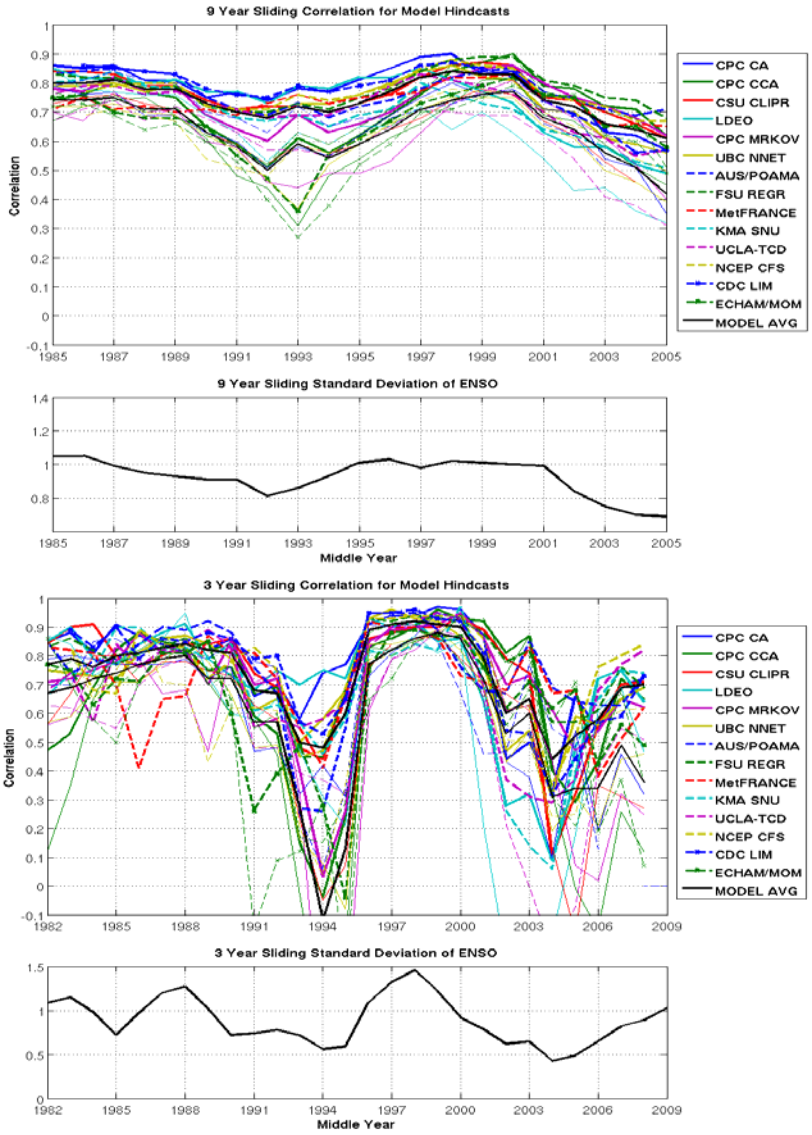


Figure 8 (top) Time series of uncentered correlations between predictions of given individual models and observations for sliding 9-year periods, phased 1 year apart, for the 21 or 22 running 9-year periods within 1981-2010. Correlations for forecasts at lead times of 3 months (thick lines) and 6 months (thin lines) are shown. Inset below main panel shows the standard deviation of observed SST anomalies for each sliding 9-year period, with respect to the 1981-2010 mean. (bottom) As in top, except for sliding 3-year periods for the 27 or 28 running 3-year periods within 1981-2010.

3. Concluding remarks

Verification of the real-time ENSO prediction skills of 20 models (12 dynamical, 8 statistical) during 2002-2011 indicates skills somewhat lower than those found for the less advanced models of the 1980s and 1990s. However, this apparent retrogression in skill is explained by the fact that the 2002-2011 study period was demonstrably more challenging for ENSO prediction than most of the 1981-2010 period, due to a somewhat lower variability. Thirty-year hindcasts for the 1981-2010 period yielded average correlation skills of 0.65 at 6 month lead time (slightly higher than the 0.6 found in studies from the 1990s), but the real-time predictions for 2002-2011 produced only 0.42. The fact that the recent predictions were made in real-time, in contrast to the partially hindcast design in the earlier studies, introduces another difference with consequences difficult to quantify but more likely to decrease than increase the recent

performance measures. Thus, based solely on the variability of 9-year correlation skills of the hindcasts within the 30-year period, ENSO prediction skill is slightly higher using today's models than those of the 1990s (0.65 versus about 0.6 correlation). Decadal variability of ENSO predictability can dominate the gradual skill improvements related to real advances in ENSO prediction science and models.

Unlike earlier results, the sample mean of skill of the dynamical models exceeds that of statistical models for start times between March and May when prediction has proven most challenging. The skill comparison by model type passes a field significance test for all seasons and leads collectively, at the p=0.03 level.

A likely reason for the better performance of dynamical than statistical models is a more effective detection and usage in dynamical models, through their initial conditions, of new information in the ocean-atmosphere system on a short (intramonth) time-scale—information that may not play a role in statistical models that use longer time means for their predictor variables. Statistical models may have potential for higher skill if their predictors were designed with finer temporal resolution. Statistical models need long histories of predictor data to develop their predictor-predictand relationships, but the valuable 3-dimensional observations across most of the tropical Pacific (e.g. from the Tao-Triton array) began only in the 1990s, precluding a robust empirical definition of their predictive relationships, and thus they are often omitted in statistical models, putting them at a relative disadvantage.

References

Barnston, A. G., and Coauthors, 1994: Long-lead seasonal forecasts – Where do we stand. *Bull. Amer. Meteor. Soc.*, **75**, 2097-2114.

Chen, D., and M. A. Cane, 2008: El Niño prediction and predictability. *J. Comput. Phys.*, **227**, 3525-3640.

Jin, E. K., and Coauthors, 2008: Current status of ENSO prediction skill in coupled ocean-atmosphere models. *Clim. Dyn.*, **31**, 647-664.

Livezey, R. E., and W. Y. Chen, 1983: Field significance and its determination by Monte Carlo technicques. *Mon. Wea. Rev.*, **111**, 46-59.

Ramseyer, G. C, 1979: Testing the difference between dependent correlations using the Fisher *Z*. *The J. of Experimental Education,* **47,** 307-310.

Wilcoxon, F., 1945: Individual comparisons by ranking methods. *Biometrics Bulletin,* **1**, 80–83. doi:10.2307/3001968.

2. DROUGHT

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Whither U.S. Drought: Much Ado About Nothing

Kelly Redmond
Western Regional Climate Center and Desert Research Institute, Reno NV

1. Introduction

Borrowing from Shakespeare, the phrase "much ado about nothing" seems appropriate for drought. In many ways, precipitation amounts to almost nothing in the more severe droughts. In reality, of course, most droughts are associated with extended periods of less than usual precipitation, but not necessarily zero, and they often result from seemingly modest deficiencies of "only" 25 or 50 percent. The southern plains (Texas-Oklahoma-New Mexico) drought of 2011 is somewhat unusual in that many locations did actually receive next to nothing in terms of precipitation over an extended period. The paradoxical quality and the resulting play on words stems from the greater and greater attention given as precipitation becomes less and less.

Drought typically is associated with the *redistribution* of water. The recently ended Water Year 2011 (October 2010 through September 2011) shows this dramatically (Figure 1). Very dry conditions in the southern plains and southeast U.S. are "balanced" by very wet conditions in other parts of the country: the southwest (cool season), the Missouri River basin (warm season), and the northeast (tropical storm). Major impacts resulted from both the wet and dry anomalies.

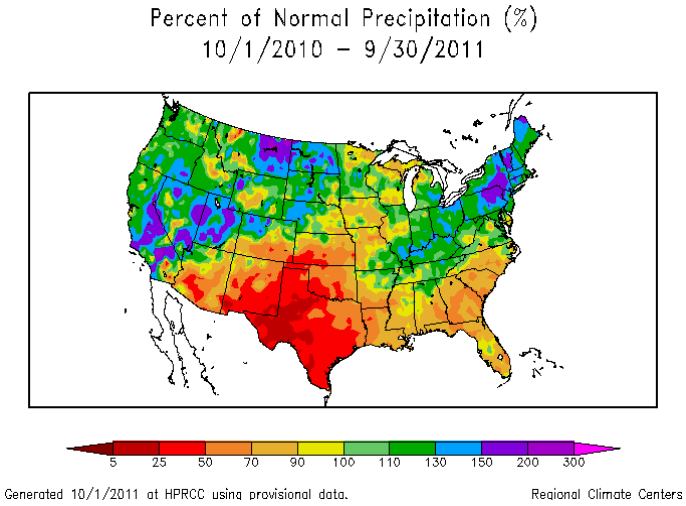


Fig. 1 Precipitation departure (from 1971-2000 mean) during Water Year 2011 (October 2010 through September 2011) based on surface measurements.

As a general rule, the most serious droughts are associated with the loss of the main precipitation season(s), in climates that have pronounced seasonal cycles. In addition, even in regions with a more even distribution of precipitation within the year, some seasons are more effective in supplying groundwater recharge and streamflow; almost always this is the cool season. This is especially the case when snowpack is an important hydrologic factor. A few dry months in a normally dry season do not have the consequences of dry months during (for example) the winter recharge season, or the summer monsoon, or the tropical storm season, depending on the part of the country. Figure 2 shows typical examples of seasonal cycles for Texas and for Arizona. With a double peak during the year, there is some opportunity to recover if the first peak is deficient. The worst droughts involve the loss of both peaks. In the Arizona case, one peak is in winter and the other is in summer. The climate causes and teleconnections that lead to a loss of winter precipitation may be very different from those that lead to a loss of summer precipitation, and thus unless they share a common source, are often relatively uncorrelated with each other.

2. Defining drought

The definition of drought has been the subject of animated discussion for well over a century, and invariably arises during attempts to depict its status and establish relationships to impacts (Redmond 2002). This situation is likely not destined to end any time soon. What appears to be a simple and straightforward

Correspondence to: Kelly Redmond, Western Regional Climate Center and Desert Research Institute, Reno NV; E-mail: kelly.redmond@dri.edu.

exercise turns out in reality to be vexingly complicated. A myriad of indicators are available to describe drought (*e.g.* Heim, 2002) and more are steadily proposed and created. For a definition to have much value, it should be applicable in the widest possible range of circumstances, and cover all the types of droughts experienced around the world. It is relatively easy to formulate definitions that fit particular geographic and sectoral circumstances. It is in the attempt to develop a universally applicable definition that we are forced to look at this problem in its generality.

The fundamental concept of drought involves a water balance. Through climatological and meteorological processes a supply of water is produced. Through climatological, meteorological, ecological, and social processes, a demand for that water has come about. Note the addition of added factors on the demand side. When supply (broadly defined) is unable to meet demand (also broadly defined), over extended periods, needs for water by human and natural systems are increasingly unable to be fully met. At some point, the accumulated imbalance begins to result in impacts. This "some point" varies, sometimes very greatly, depending on the specific circumstances of each of the many sectors affected by shortages in their own particular water budget. The consequences of these shortages are described as "impacts" and can be expressed in socioeconomic terminology. (The situation for "natural systems" is more ambiguous, and is tied to various kinds of value judgments.) The reason we care about drought is that these shortages (negative water budget) produce impacts. Because of this, drought is thus defined by its impacts. For human systems, if there is no impact, there is no drought. For natural systems, every excess or deficiency in the moisture budget results in some kind of adjustment, and in this case the definition is more murky. Because of dependence on the specific path of causation from climate to impact, the same geophysical drivers need not always lead to the same level of impact, and thus of drought.

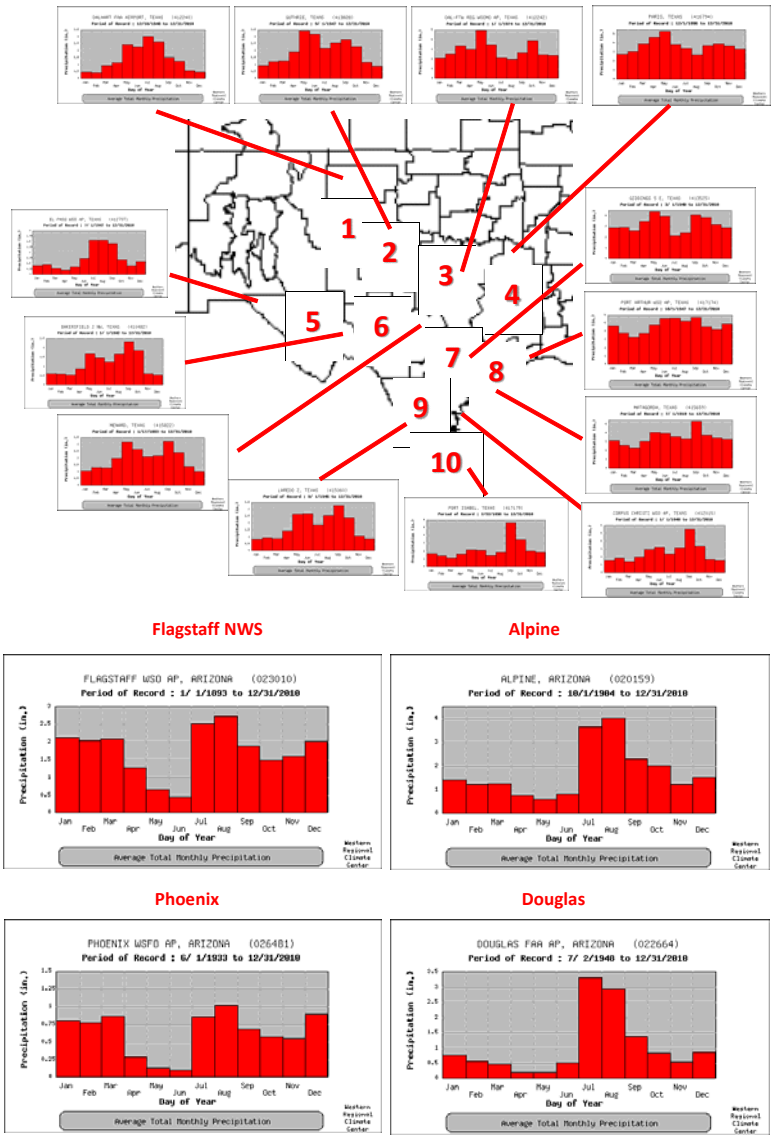


Fig. 2 Qualitative depiction of typical seasonal cycles of precipitation in Texas (top) and in Arizona (bottom). Based on period of record. Graphs cover January to December, and are to different scales.

The reliance on a water budget approach implies that both supply and demand matter. Demand, as used here, includes such factors as evapotranspiration, as well as human, animal, and vegetative demand. The factors that affect supply certainly include precipitation, but also other climatic elements such as temperature, wind, radiation, and humidity. These factors are even more relevant to demand. The history of drought, over decades or centuries, is affected by the separate histories of both supply and of demand. The presence of numerous buffers, and of different temporal lags (*e.g.*, relating to snowpack formation and melt timing, or of groundwater recharge), and of spatial separation (snow, or its lack, in Colorado affects water supplies in Los

Angeles months or years later), complicates matters further. These issues have themselves been discussed elsewhere, and are beyond the immediate scope of these remarks.

Metrics for drought that are capable of taking all of these real-world effects into account (in order to be practical and useful) are the subject of longstanding discussion. Their formulation is clearly not straightforward.

3. The Drought Monitor

The U.S. Drought Monitor (Svoboda *et al.*, 2002) is both a process and a product. The process is an extended weekly electronic national conversation from a pool of up to nearly 300 participants (typically 30-50 voices in a given week). The product is the map itself and associated descriptive material. The Drought Monitor developed in rather a grass-roots fashion, starting as an experimental product in 1999 arising from a drought around the nation's capital. An important characteristic is that it is owned by everybody and by nobody, with contributions from a variety of federal and state agencies and individuals. The impact-based definition of drought is adopted to insure that indications from physical measurement are corroborated by indicators of affected sectors. The Drought Monitor is thus a combined social and physical endeavor with real-world grounding. This is especially necessary where the Drought Monitor forms the basis for resource allocation decisions, an increasingly common situation. In many cases the Drought Monitor constitutes a late arrival, especially in western U.S. settings where the complexities of drought are greater, and have been understood and addressed by a variety of interagency approaches over many decades. An outstanding issue is how to represent drought in heavily managed systems.

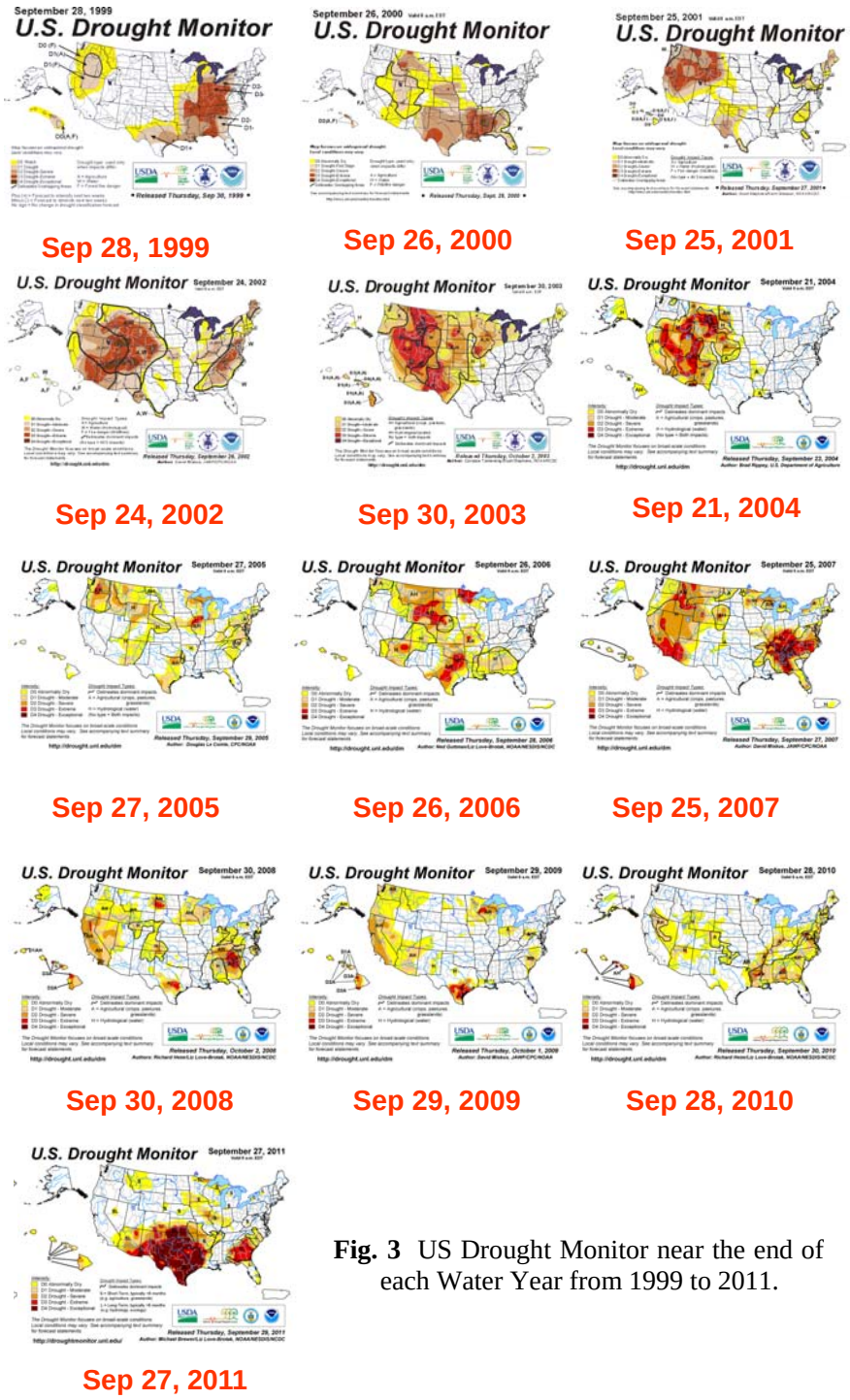


Fig. 3 US Drought Monitor near the end of each Water Year from 1999 to 2011.

The email "conversation" referred to above has proven to be a rich, vibrant, interesting, varied, and generally quite intelligent discussion. This has provoked a variety of real-world, practical, and intellectual challenges to addressing the many dimensions of drought.

Summarized in Figure 3 is an annual snapshot of the Drought Monitor near the end of each Water Year. This shows that over this interval drought has been present somewhere in the United States for the entire time it has been produced (also true of months not shown).

4. The Standardized Precipitation Index

Approximately 30-40 products and tools are first examined by the weekly Drought Author (the person charged each week with assembling the map; there are about 10 such individuals) to provide an initial map. This is then critiqued and modified through community input. Even though drought monitoring requires information beyond merely precipitation, this does remain the most important primary element. One of the tools consulted is the Standardized Precipitation Index (SPI), which uses a long history (in this case about 115 years) to create a suite of five main products, shown in Figure 4. Depending on the basic underlying climate and its typical seasonal cycle, all of these maps have relevance in the assessment of drought status. The SPI itself represents a mapping of the percentile, obtained either from the empirical distribution (usually) or a fitted distribution (sometimes) of accumulated precipitation over a particular duration, typically 1 month to 6 years, onto a normal distribution, and expressed in terms of standard deviations. The percentile map and the SPI map are thus the two most spatially comparable products in the SPI suite. The SPI (see McKee *et al.*, 1993, 1995) was created to explicitly reflect the fact that a region can be simultaneously in excess and in deficit, at different time scales (see Figure 5 for examples). This feature of the SPI is under-appreciated and not utilized to full potential. In part this may a consequence of the lack of a good method for showing the temporal history that has led to the present situation, in the form of a spatial representation of such histories.

5. Concluding comments

In addition to the complexities of representing spatial variations in the recent history of "meaningful" precipitation (not all precipitation is equal in value), other spatial issues are present that are pertinent to drought. Jerome Namias (citation unavailable) once noted that droughts that are long in duration tend to be large in spatial extent. Contrast two significant recent droughts: The southeast drought of 2007 developed rapidly in about February and lasted about 10 months during the worst conditions, and especially affected a relatively small area that just

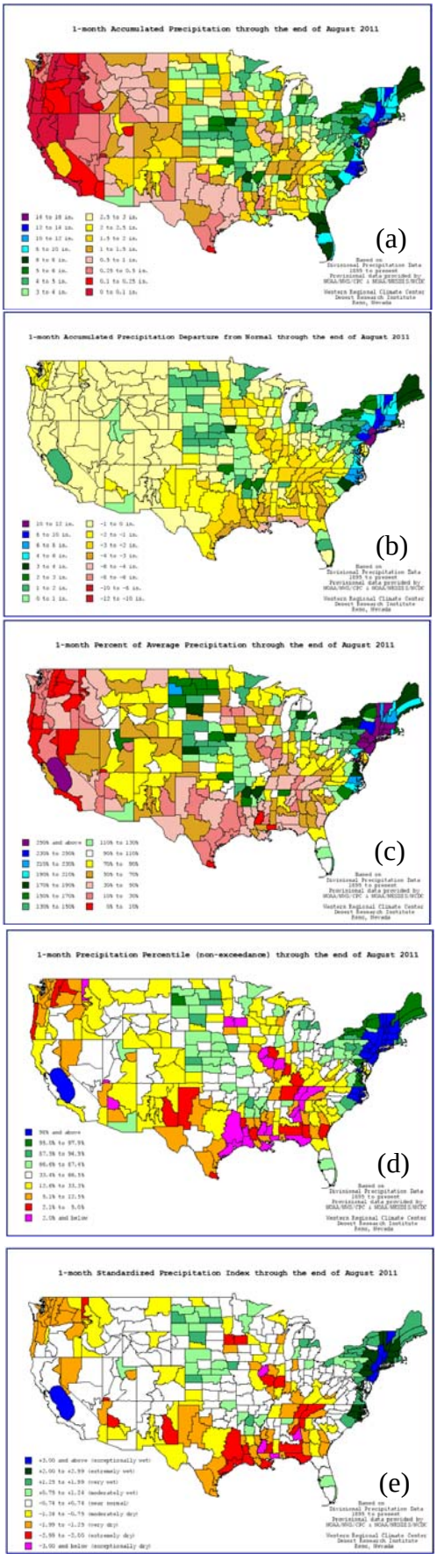


Fig. 4 The five precipitation quantities most requested by water managers and hydrologists in understanding the status of drought (McKee *et al.*, 1993, 1995). These are routinely computed in generating SPI values. One-month quantities for August 2011 are shown, (a) accumulation, (b) accumulation departure, (c) accumulation percentage of average, (d) accumulation percentile, and (e) a standardized measure (the SPI).

happened to be the source of the Atlanta GA water supply. A second drought that also affected water supplies to millions of people in the Southwest developed in the late 1990s, peaked in 2002, but continued to have effects for at least several more years, and has still not entirely abated and may yet re-intensify. This drought affected entire state-sized regions in the headwaters of the Colorado River, which has several years of buffering capacity in its two large reservoirs, Powell and Mead. This drought extended over about a decade and has affected a large area. Multi-year droughts require a multi-year cause, typically suspected to be ocean conditions elsewhere on the globe and the subject of many investigations.

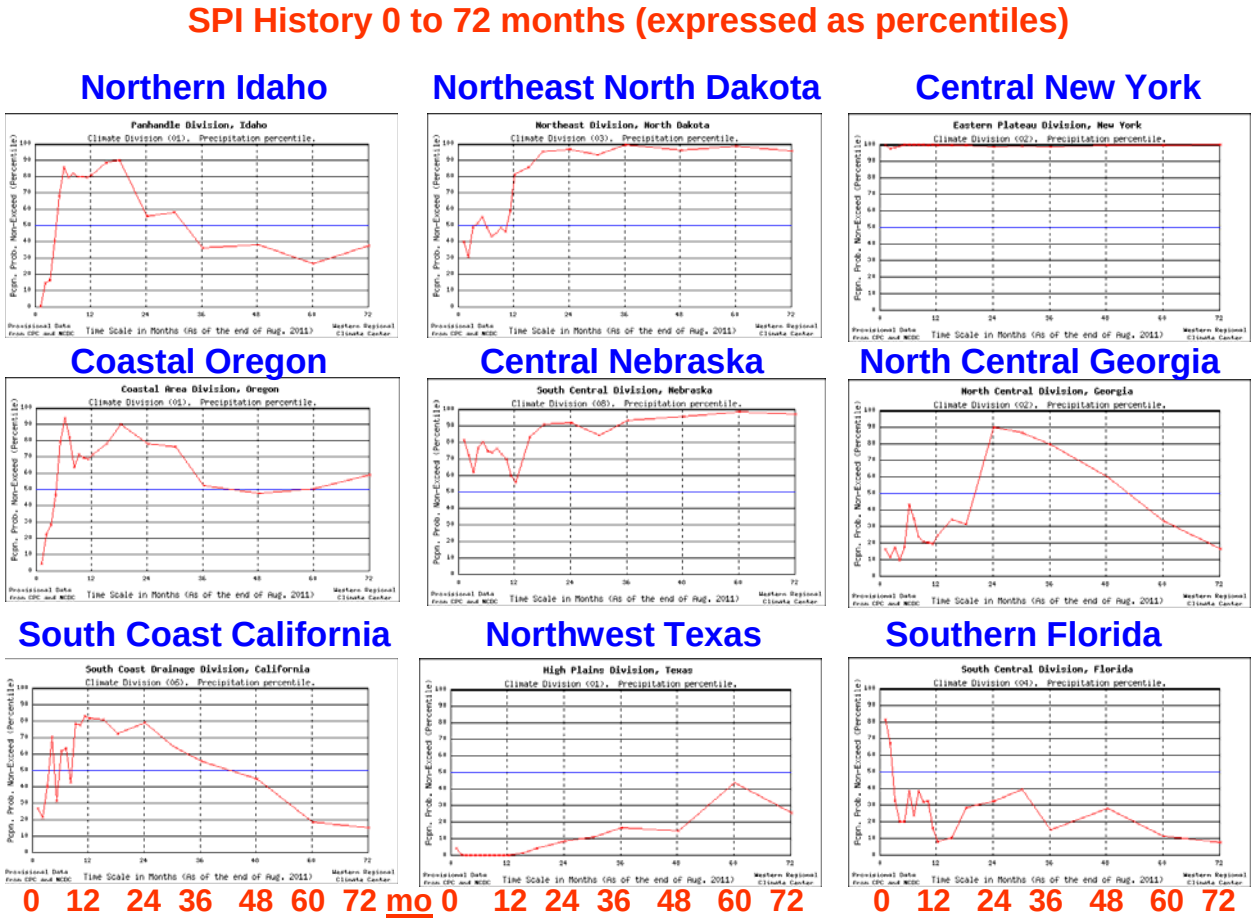


Fig. 5 Examples of the precipitation percentile (scale 0-100) at different time scales for different parts of the U.S. for the period ending at the end of August 2011. Time scales range from past 1 month to past 72 months.

Some of the outstanding issues of drought are as follows:

- We have made a lot of progress in better monitoring, a larger array of tools and products, and improved physical understanding of some drought causes. Social science aspects of drought need more attention and need to be folded into drought activities in the U.S.
- There is still a strong need for granularity and resolution in both the monitoring and prediction of drought. This is especially true in mountain environments, where much runoff is generated in relatively small source regions, and climatic heterogeneity is extreme (sharp spatial gradients in temporal histories).
- We need to give continued and increased attention to understanding, measuring, and depicting the state of the entire water budget, to cover both the supply and demand sides of drought.

- In some places, major fractions of the annual water supply are delivered in a few big events, such as large winter storms or tropical storms. Their occurrence or non-occurrence may greatly affect seasonal precipitation totals and status. The role of significant weather events (an example is atmospheric rivers) in affecting seasonal and multi-year drought status, and elucidation of the climate-weather connection, are greatly in need of further sustained attention.
- We must continue to press for better understanding of multi-year and decadal scale variability. A diversity of methods is needed, including paleoclimate inferences.
- There are many different styles ("flavors," Bumbaco and Mote, 2010) of drought. Most of these tie to the role of differential seasonal contributions to longer-term droughts. There are many nuances that lead to different impacts, and this needs to be better and more widely understood.
- There is really no one-size-fits-all approach to drought.

References

Bumbaco, K.A., and P.W. Mote, 2010: Three recent flavors of drought in the Pacific Northwest. *J. Appl. Meteor. Clim.*, **49**, 2058-2068.

Heim, R.R. Jr., 2002: A review of Twentieth-Century drought indices used in the United States. *Bull. Amer. Meteor. Soc.*, **83**, 1149-1165.

McKee, T.B, N.J. Doesken, and J. Keist, 1993: The relationship of drought frequency and duration to time scales. Preprings, Eighth conf. on applied Climatology, Anaheim CA, Amer. Meteor. Soc, 179-184.

McKee, T.B., N.J. Doesken, and J. Keist, 1995: Drought monitoring with multiple-timescales. Preprints, Ninth Conf. on Applied Climatology, Dallas, TX, Amer Meteor. Soc., 233-236.

Redmond, K.T., 2002: The Depiction of Drought: A commentary. *Bull. Amer. Meteor. Soc.*, **83**, 1143-1147.

Svoboda, M., D. LeComte, M. Hayes, R. Heim, K. Gleason, J. Angel, B. Rippey, R. Tinker, M. Palecki, D. Stooksbury, D. Miskus, and S. Stephens, 2002: The Drought Monitor. *Bull. Amer. Meteor. Soc.*, **83**, 1181-1190.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Hydro-climatological Drought Analyses and Projection Using Meteorological and Hydrological Drought Indices: A Case Study in Blue River Basin, Oklahoma

Lu Liu[1,2], Yang Hong[1], Christopher N. Bednarczyk[3],
Bin Yong[4,1], James E. Hocker[2], Mark A. Shafer[2]

[1]*School of Civil Engineering and Environmental Science, University of Oklahoma*
[2]*Southern Climate Impacts Planning Program, Oklahoma Climate Survey, University of Oklahoma*
[3]*National Weather Center Research Experiences for Undergraduates Program and Valparaiso University*
[4]*State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, China*

1. Introduction

Understanding the characteristics of the historical droughts will be beneficial to reveal the possible impacts of the future climate changes on such climatic extreme phenomenon (Edwards and McKee 1997). However, drought is difficult to be quantified due to its dependence on regional differences, needs, and disciplinary perspectives (McKee *et al.* 1993). For the past decades, various drought indices have been developed to assimilate thousands of bits of data on rainfall, snowpack, streamflow, and other water supply indicators into a comprehensible big picture (Heim 2002; Wells *et al.* 2004; Jain *et al.* 2010). Therefore, the overarching goal of this study is to reconstruct the past drought situation and assess the future drought risks for a drought-prone Blue River Basin, Oklahoma, under a likely changing climate from three aspects, *i.e.*, intensity, duration and extent. This basin is located in Southeastern Oklahoma with a drainage area of 676 square miles (Fig. 1). It is a relatively small basin but has experienced several severe droughts in the past half century. The first objective of this study is to construct the past drought conditions and predict future drought scenarios for Blue River Basin using three types of drought indices, the Standardized Precipitation Index (SPI) (McKee *et al.* 1993), Palmer Drought Severity Index (PDSI) (Palmer 1965) and Standardized Runoff Index (SRI) (Vasiliades *et al.* 2010). These types are meteorological drought index, hydro-meteorological index and hydrological index, respectively. The second objective is to examine the relationships among the three indices. The third objective is to find the suitable drought index for Blue River Basin under a changing climate.

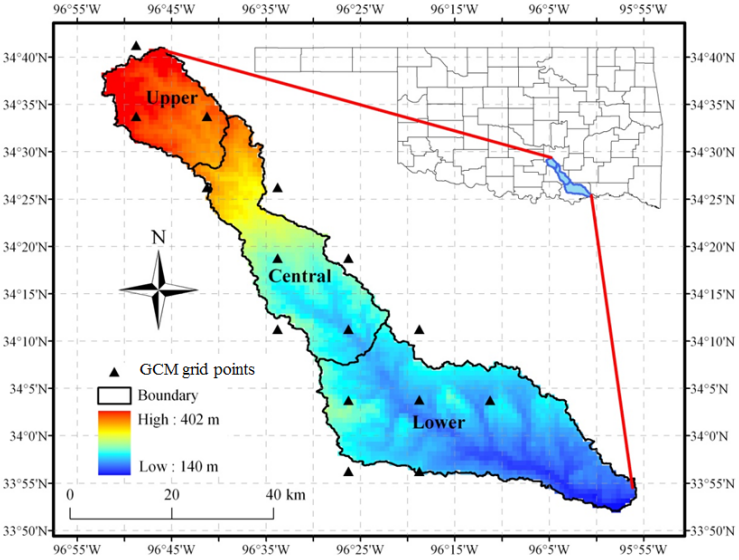


Fig. 1 The study area: Blue River Basin in Oklahoma.

2. Data, model and drought indices

2.1 Climate Data

Climate data of the study region were first extracted and later modeled for SPI, PDSI and SRI calculation. For this study, the observational data used were the gridded National Climatic Data Center (NCDC) Cooperative Observer station data, described by Maurer *et al.* (2002). The data cover the time period from 1950 to 1999 in a monthly time step. The observation data have climate variables of surface temperature (°C)

Correspondence to: Lu Liu, School of Civil Engineering and Environmental Science, University of Oklahoma; E-mail: liuemma2@ou.edu.

and monthly precipitation (mm/day). The data domain covers the continental U.S. and portions of southern Canada and northern Mexico at a 1/8° (~12 km) resolution. Projection data are archived from the World Climate Research Programme's (WCRP's) Coupled Model Intercomparison Project phase3 (CMIP3) multi-model dataset. CMIP3 has temperature and precipitation projections under three CO2 emission scenarios (A2, A1B and B1) for the period of 2010 to 2099, and this data share the same resolution and coverage with the NCDC observation data. The two scenarios of the 21st century for future greenhouse gas emissions used in this study were A2 and A1B as defined in the IPCC Special Report on Emissions Scenarios (IPCC 2007). According to IPCC (2007), scenario A2 is a higher emission path and describes a higher population world where technological change and economic growth are more fragmented and slower. Scenario A1B is a middle emission path known as business-as-usual and describes a balanced world where people do not rely too heavily on any one particular energy source.

2.2 Thornthwaite Monthly Water Balance Model

The hydrological model used to simulate the hydrologic process and generate runoff output for SRI calculation is the Thornthwaite Monthly Water Balance Model driven by a graphical user interface. It is named after C.W. Thornthwaite who used water budget in climate classification (Thornthwaite 1948). An updated description is given by McCabe and Markstrom (2007). Input for this model is monthly temperature and precipitation. Outputs from the model include potential evapotranspiration (PET), soil moisture, actual evapotranspiration (AET), snow storage, surplus, and runoff total.

3. Results and Discussions

3.1 Runoff Calibration

The Thornthwaite Monthly Water Balance Model was calibrated to generate future runoff under the A1B scenario. Manual calibration of the monthly water resources model was done to get the best agreement between observed and modeled runoff for the period of June 1936 through August 2006. The parameters used modeled the calibration period well with the Nash-Sutcliffe coefficient of efficiency being 0.78 and a root mean square error of 12.9. The modeled runoff was used to simulate past SRI and project future SRI.

3.2 Past and Future Drought

Blue River Basin is located within the state of Oklahoma, which survived the 1930s Dust Bowl and 1950s severe drought. Historical records show that Oklahoma experienced four major droughts in the 20th century: 1909-18, 1930-40, 1952-58, and 1962-72. According to Oklahoma Climatological Survey (OCS), “the drought of the 1930s is associated with the Dust Bowl of the Great Plains, when socioeconomic conditions, agricultural practices and drought forced the largest emigration of Oklahomans in state history.” “Statistically or meteorologically, for much of the ABRFC/Oklahoma, the droughts of the 1950s were more severe (record low SPI and PDSI) than the 1930s. However, the human toll (Socioeconomic Impact) was less severe.” “The lessons of the 1930s helped the next generation cope with the worse droughts of the 1950s: preparedness and mitigation embodied in crop selection, conservation strategies, and sound business decisions.” (Arndt 2002)

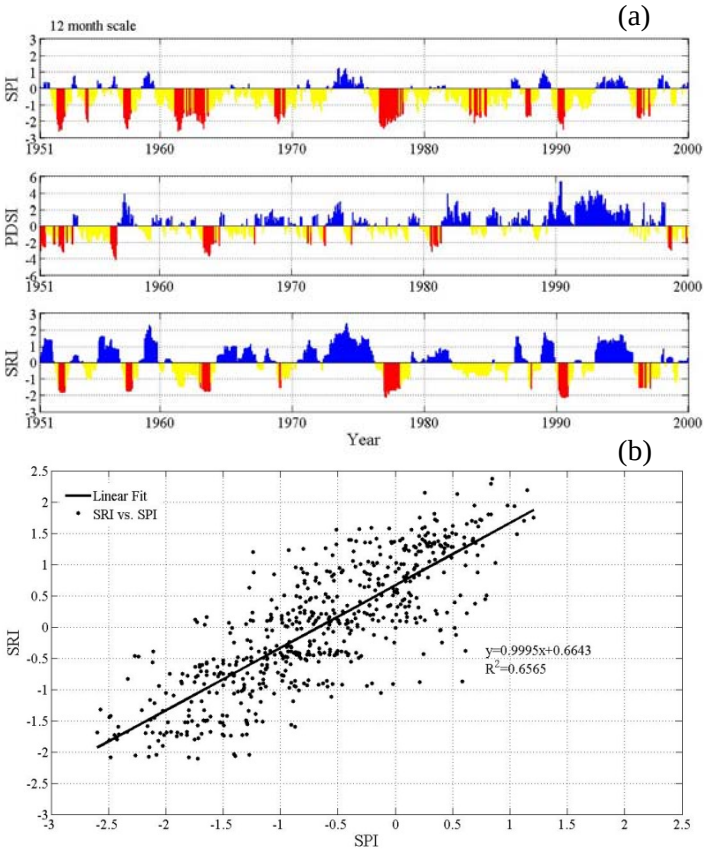


Fig 2. (a) Historical time series variation of SPI, PDSI and SRI (b) Scatter plot of SPI vs. SRI and the linear regression line.

As can be seen from Figure 2a top panel, SPI shows massive droughts from the beginning of 1950s to the end of the 1960s. 1950s droughts are interrupted by several wet spells which slightly relieved the drought severity. As shown on the figure, there was a long duration extreme drought between 1960 and 1965. This drought lasted almost 5 years without being interrupted by occasional wet spells. The drought was released slightly by some above average rains in 1965 and formed into another extreme drought by the end of 1960s. Other than the other mega drought near 1980, Blue River Basin experienced some normal dry spells and some wet spells after 1980.

PDSI provides somewhat different descriptions on Blue River Basin drought history (Fig. 2a, middle panel). The 1950s and 1960s drought were captured roughly, but the onset and severity were slightly different. The 1970s drought on SPI was totally missed on PDSI panel, and Blue River Basin is mostly under wet conditions after 1980 except for one severe drought around 1981. SRI, in this case, is very similar to SPI in terms of severity and timing of droughts. Figure 2b shows that SPI and SRI are highly correlated with correlation coefficient (CC) of 0.81. SRI is found out to have two month lag time from SPI (CC reaches highest value), which means hydrological drought doesn't happen until two months later after the meteorological drought took place. Although the wet spells look more significant on the SRI panel, SRI successfully captures all the major droughts except the one in mid 1980s. In general, droughts shown on SRI mostly have shorter duration than the same ones shown on SPI, and wet periods are longer than those on SPI panel.

For drought projection using SPI, ensemble mean monthly precipitation should not be used because the averaging process diminishes the monthly variation of precipitation which could generate misleading outputs. Therefore, one of the 16 GCMs --- GISS-ER is selected, because simulation from GISS-ER is proved to match the observation for the period of 1950-1999 with the most similarities from a statistical point of view from a previous study (Liu *et al.* 2011). Hence, future SPI is calculated based on GISS-ER projection data with the most confidence. Projections of drought conditions in Blue River Basin display quite some differences among the three indices (Fig. 3a). SPI indicates one minor drought in the early 2020s, and the frequency and intensity of drought appear to increase substantially after 2050. PDSI and SRI also display much more droughts after 2050, although PDSI and SRI show more similarity in time series. PDSI and SRI time series has correlation coefficient of 0.78 (Fig. 3b) and they do not exhibit any time lag from one another (CC reaches highest value at 0 month lag time). More drought events are displayed on PDSI panel than on the SRI panel, and severe droughts on PDSI are indicated to be more severe (PDSI<-5) than those on SRI. Blue River Basin is constantly under wet conditions before 2050 for both PDSI and SRI, with decreasing trend of wetness from

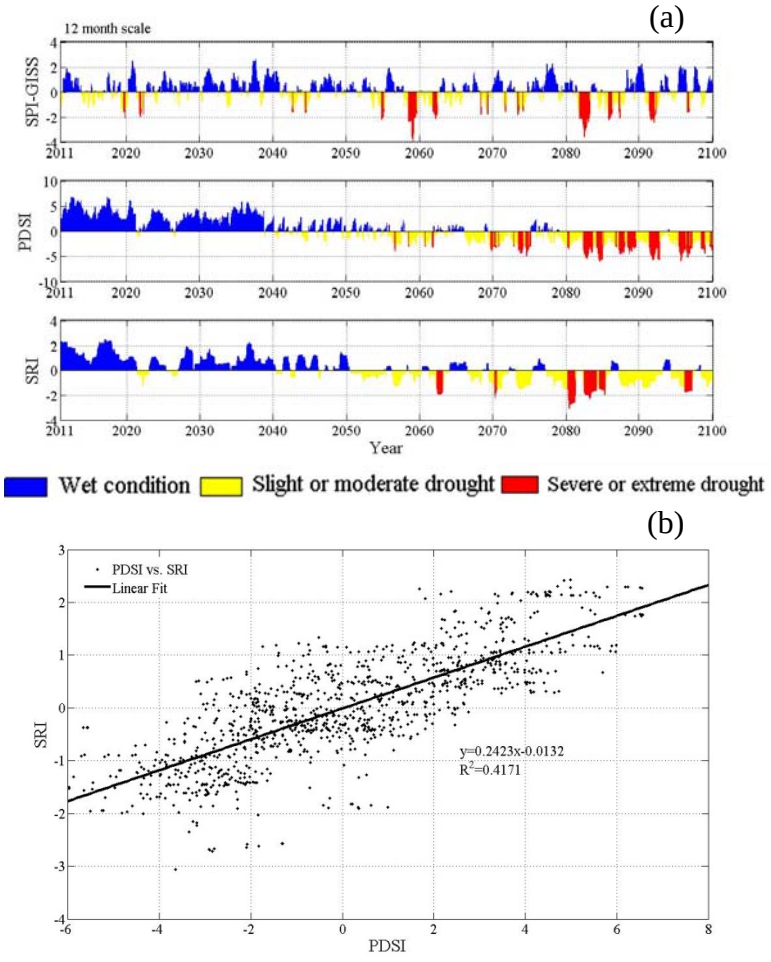


Fig. 3 (a) Future time series variation of SPI, PDSI and SRI (b) Scatter plot of SPI vs. SRI and the linear regression line.

2011 to 2050. It is not surprising to see that both PDSI and SRI demonstrate more severe and frequent drought after 2050, although the magnitude of droughts is not exactly the same. Based on Thorthwaite Monthly Water Balance Model projection, Blue River Basin is going to have increasing trend of ET and decreasing trend in total runoff under A1B scenario. Actual ET is going to increase by up to 8% on average and runoff is going to reach below -10% of change by the end of the 21st century. Apparently, more water is going out as ET and less water will be available for surface runoff, so there is less recharge back to the ground.

Future droughts are getting more severe and frequent from SPI, PDSI and SRI perspectives. Table 1 shows the areas affected by severe or extreme droughts based on the basin division. Historically, the affected areas are almost equally distributed among the upper, central and lower Blue River Basin. Both SPI and PDSI display an average around 3% of the areas affected by severe/extreme droughts from 1950 to 1999, note that the lower region is slightly less affected than the upper and central regions. In terms of projection, SPI shows an average of 25.6%, 22.9% and 20.6% of areas affected in the upper, central and lower Blue River Basin throughout the 21st century. Results from PDSI are displayed in two half centuries: the first half century sees almost no droughts while the second half century sees an average of 23.5%, 23.5%, 22.2% of areas affected by severe or extreme droughts. Overall speaking, both SPI and PDSI exhibit larger areas being under severe or extreme droughts in the future period.

Blue River Basin	1950-1999		2010-2099			
	SPI	PDSI	SPI	PDSI (2010-2049)	PDSI (2050-2099)	PDSI (total)
Upper	3.5%	3.5%	**25.6%**	0.5%	**23.5%**	13.3%
Central	3.5%	2.8%	**22.9%**	0.5%	**23.5%**	13.6%
Lower	3.2%	2.7%	**20.6%**	0.83%	**22.2%**	12.8%

Table 1. Percent of areas under severe/extreme droughts for upper, central and lower Blue River Basin.

4. Conclusions

This study analyzed the historical drought of the Blue River Basin in the past 50 years and projected the possible drought status of the future 90 years under the A1B scenario, a very likely future climate in Southern US based on previous studies. Three types of drought indices (SPI, PDSI and SRI) all capture the major droughts documented in history. In terms of timing and severity, SPI and SRI perform better and exhibit higher inner correlation. The results projected by SPI, PDSI and SRI under the business as usual A1B scenario suggest that more drought events might occur in the second half of the 21st century. This could be caused by the fact that the precipitation predicted by GISS-ER shows a descending trend, while the temperature is slowly but constantly increasing after 2010. Moreover, the ET projected by the Thornthwaite Monthly Water Balance Model also has a significant increasing trend under such a warming climate. In the projection period, PDSI and SRI have a similar performance because they both take into account the factors of soil moistures and ET. Therefore, the drought of future Blue River Basin is very likely going towards a more frequent and more intensive situation.

In this study, SRI appears to be a better indicator for the study basin because: (1) SRI considers the changing climate factor which is playing a rather significant role in the future drought management; (2) Compared to PDSI which also considers temperature change, SRI provides drought information from a hydrological point of view, which makes more sense and is more applicable to water resources managers and local farming business; and (3) SRI is proved to function well in this research both for the past drought record reconstruction and for the future drought risk assessment under a changing climate.

In summary, this study found that the three indices (*i.e.* SPI, PDSI and SRI) captured the recorded droughts for the past 50 years and also suggested very likely more frequent and more severe droughts in the future 90 years over the Blue River Basin. This study also found that SRI has better agreements with the other two indices, with high CC of 0.81 (0.78) and 2-month (no appreciable) lag time from SPI (PDSI) over 1950-

2099 time period across the basin. Although this study recommended the PDSI and SRI are the more suitable indices to assess the future drought risks under an increasingly warming climate by taking into account of the ET and soil moisture in Blue River Basin, for further comprehensively analyzing the droughts, more drought indices from ecological and socioeconomic perspectives should be also investigated and inter-compared to provide a more complete picture of drought risks and its potential impacts on the dynamically coupled nature-human systems.

References

Arndt, D.S., 2002: The Oklahoma Drought of 2001-2002. Oklahoma Event Summary. Oklahoma Climatological Survey.

Edwards, D.C., and T.B. McKee, 1997: Characteristics of 20th century drought in the United States at multiple time scales. Climatology Report Number 97–2, Colorado State University, Fort Collins, Colorado.

Heim, R.R. Jr., 2002: A review of twentieth-century drought indices used in the United States. *Bull. Amer. Meteor. Soc.*, **83**, 1149–1165.

IPCC, 2007: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. [S. Solomon, D. Qin, M. Manning, M. Marquis, K. Averyt, M. M.B.Tignor, H.L. Miller, Jr. and Z. Chen (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jain, S., R. Keshri, A. Goswami, and A. Sarkar, 2010: Application of meteorological and vegetation indices for evaluation of drought impact: A case study for Rajasthan, India. *Natural Hazards*, **54**, 643-656.

Liu, L., Y. Hong, E.J. Hocker, A.M. Shafer, M.L. Carter, J.J. Gourley, N.C. Bednarczyk, and P. Adhikari, 2011: Analyzing projected changes and trends of temperature and precipitation in the Southern U.S. from 16 downscaled global climate models under different emission scenarios, *Theoretical and Applied Climatology*, (submitted)

Maurer, E.P., A.W. Wood, J.C. Adam, D.P. Lettenmaier, and B. Nijssen, 2002: A long-term hydrologically-based data set of land surface fluxes and states for the conterminous United States, *J. Climate*, **15**, 3237-3251

McCabe, G.J., and S.L. Markstrom, 2007: A monthly water-balance model driven by a graphical user interface. U.S. Geological Survey Open-File Report 2007-1088, 12 pp

McKee, T.B., N.J. Doesken, and J. Kleist, 1993: The relationship of drought frequency and duration to time scales. Preprints, 8th Conference on Applied Climatology, pp. 179–184. January 17–22, Anaheim, California.

Palmer, W.C., 1965: Meteorological drought. Research Paper No. 45., U.S. Weather Bureau

Thornthwaite, C.W. 1948: An approach toward rational classification of climate. *Geographical Review*, **38**, 55-94

Vasiliades, L., A. Loukas, and N. Liberis, 2010: A water balance derived drought index for Pinios River Basin, Greece. *Water Resources Management*, **25**, 1087-1101

Wells, N., S. Goddard, and Co-authors, 2004: A self-calibrating Palmer Drought Severity Index. *J. Climate* **17**, 2335-2351

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

The Introduction of an In-depth Drought Monitoring Process in the Upper Colorado River Basin

Rebecca Smith

Department of Atmospheric Science, Colorado State University Fort Collins, CO

1. Introduction

Drought in the Upper Colorado River Basin (UCRB) can affect millions of people and many acres of farmland. The majority of water supplies in the western United States start as mountain snowfall. The snowfall in the UCRB eventually provides water for seven states. Therefore, it is essential to monitor the UCRB for climate and precipitation variability.

The U.S. Drought Monitor (USDM) has been the primary tool for tracking drought changes in the UCRB (and the rest of the country) since 2000. Before 2010, only a few experts would occasionally contribute to the weekly monitor for areas in the UCRB—however this was rarely coordinated with other experts in the region. Local users have said they use the USDM to assess large scale conditions, but don't find it useful on a local scale. Based on user needs though, it seems that the USDM could be the most comprehensive tool at their disposal, but improvements are needed.

In February 2010, an increased monitoring effort began as part of the National Integrated Drought Information System's (NIDIS) Upper Colorado River Basin Pilot project. This increased monitoring included the implementation of weekly summaries and webinars—both of which monitor all water variables in the UCRB and culminate in a "consensus" of recommended changes to the weekly USDM depiction based on input from local experts. The goals of this increased monitoring effort (and the NIDIS-UCRB pilot) are 1) to determine critical thresholds for when a drought begins and ends, 2) to evaluate drought indices, and 3) to develop a drought "early warning" system. With the accomplishment of these goals, it is the hope that local users' needs will be better met via the USDM weekly depictions.

This study focuses on three key drought indices (the standardized precipitation index, the Palmer drought severity index, and the surface water supply index) and compares them over time to the U.S. Drought Monitor drought categories. Ideally, the comparison will show a close relationship between the drought indices and the USDM, and hopefully the analysis will also show an improvement in that relationship after the introduction of the increased monitoring in February 2010.

2. Data and drought indicators

Because of the relatively short timespan of post-increased-monitoring, this study will focus on drought and precipitation data in Colorado (due to the lack of drought conditions in the UCRB and the prevalence of severe drought in southeast and southern Colorado since February 2010). For this study, 110 National Weather Service Cooperative Network Stations (COOP stations) are used to calculate the standardized precipitation index (SPI, McKee *et al.* [1993]) as one drought indicator. These stations contain daily precipitation data from 1981 – 2011. Missing daily data could be replaced with zeros if the three closest neighboring stations reported no precipitation for the day. The data in this study are only included when less than 5% of the data for a given month are missing. Weekly Palmer Drought Severity Index (PDSI) statistics were downloaded from the Climate Prediction Center (2011) for the Rio Grande climate division in southern Colorado as another drought indicator. Finally monthly surface water supply index (SWSI) values were obtained from the Natural Resources Conservation Service (2011) for the Rio Grande watershed as a third drought indicator.

Correspondence to: Rebecca Smith, Colorado State University, Department of Atmospheric Science, Fort Collins, CO 80523-1371; E-mail: smithre@atmos.colostate.edu.

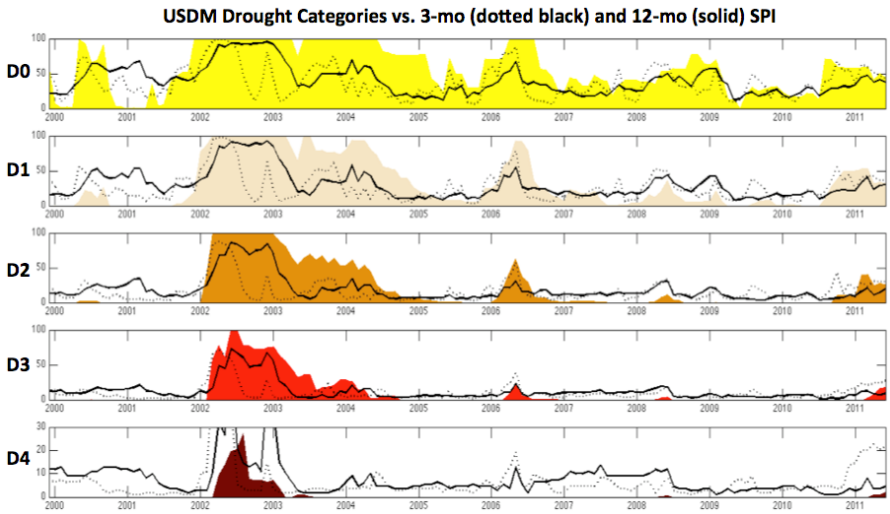


Fig. 1 Percent of Colorado in each drought category (colors) compared to the percent of COOP stations with an SPI percentile matching the associated drought category threshold.

Weekly USDM archive data were downloaded for Colorado statistics (2011). The USDM archive data simply provides the percent of the state (or county) that was currently in each drought category for every week since 1/1/2000 through 7/31/2011. Drought categories are drawn based partly on objective, quantitative analysis of many different drought indicators and water variables. Each category's threshold is determined by a specific percentile (*e.g.* a specific location could be in the D0 category [abnormally dry] if one of the drought indicators is at the 30th percentile—meaning a 3 out of 10 year occurrence). D4 droughts are the most rare and occur the least often in history (with indicators at the 0 – 2nd percentile, or a 1 in 50 year event). Other important factors for the USDM drought category determinations are more subjective and are based on user input and local impacts. This can make it more difficult to quantitatively assess how increased monitoring has improved the USDM depictions.

The SPI (first introduced in 1993 by McKee *et al.*) transforms precipitation data from its typical gamma distribution to a Gaussian distribution. This allows for easier analysis and comparison to the USDM as it puts precipitation anomalies into percentiles. For this study, the 3-month SPI is calculated for short-term anomalies and the 12-month SPI is calculated for long-term anomalies at all 110 COOP stations in Colorado at a monthly timescale.

The PDSI was developed by Wayne Palmer in the 1960s and uses temperature and precipitation data in a formula to determine dryness. It is best used to depict long-term moisture deficiency or excess for all the major climate divisions in the U.S. An index value near 0 indicates near normal conditions, and the more negative the index, the more severe the drought in that location is.

The SWSI, developed in the 1980s, is an algorithm that combines water supply (from reservoirs) and streamflow forecasts into one index for a given watershed. Negative values of the SWSI indicate that the watershed is short in water supplies.

3. Comparing the U.S. drought monitor with drought indicators

In Figure 1, it can be seen that there is good correlation between the USDM drought categories and the SPI timeseries. Overall, the drought categories are better correlated with the long-term SPI than the short-term SPI. Also of note, there is almost always D0 somewhere in CO throughout time, and in general, there is a higher percentage of D0 in CO than the percentage of stations meeting the D0 threshold in terms of SPI. This is also true many times for D1 and D2. This is not the case for the more severe D-categories—there is almost always a higher percentage of stations with SPIs meeting the D4 threshold than there is D4 in the state.

After February 2010, increased monitoring began. When comparing before and after, there is little change in the D0 – D2 panels. However, the D3 – D4 panels tend to show spikes in the SPI timeseries without corresponding spikes in the D-category timeseries before the increased monitoring period. After the increased monitoring, spikes in the SPIs are better correlated with increases in the D3 – D4 percentages.

In the USDM timeseries in Figure 2 (top), there are two instances that D4 has been introduced in the Rio Grande basin—first during the drought of 2002 and also during the summer of 2011. The weekly PDSI

timeseries (Fig. 2, bottom) shows both of these times were indeed at D4 intensity. The weekly PDSI also shows that three other periods would have qualified for D4 intensity—all of these times were before increased monitoring began. For example, in 2006 the PDSI indicated that D4 could be introduced into the basin, but the USDM timeseries only shows a D2 intensity.

There is not as strong of a relationship when comparing the end-of-month USDM percentages in the Rio Grande basin with the monthly SWSI index for the basin (not shown). In general, when the USDM timeseries indicates D2 or greater drought through much of the basin, the SWSI tends to be negative. However, there are periods when the SWSI is negative but this is not captured in the USDM timeseries or the other drought indicators. Therefore, the SWSI does not appear to be as good of an indicator for the Rio Grande basin.

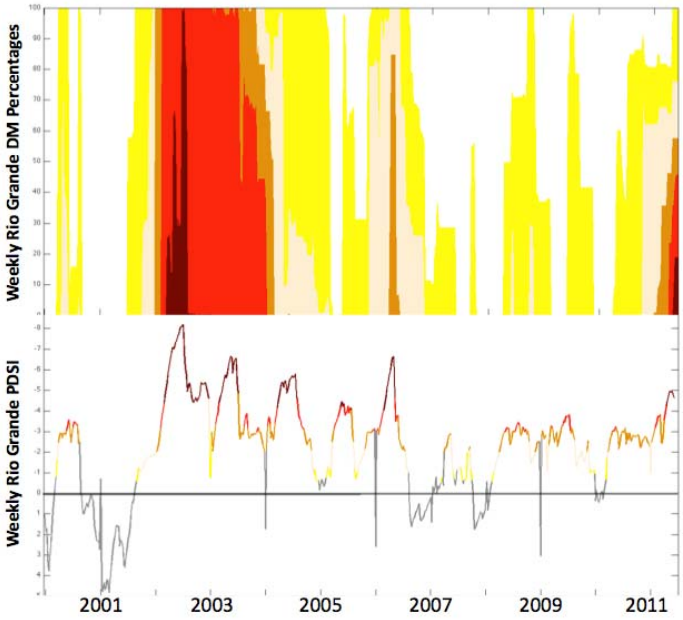


Fig. 2 Percent of Rio Grande basin in each drought category (top) compared to the PDSI value for the basin (bottom) color-coded for its associated drought category threshold.

4. Conclusions

The SPI has been and continues to be a good indicator for drought conditions throughout Colorado. At this time, it appears to be the best drought index to use for the USDM, as percentiles for the SPIs match up well with thresholds for the drought categories. It is also a better matrix for the USDM as it is higher resolution than the PDSI—while the PDSI gives an index for an entire basin (which can be useful for larger scale assessment) the SPIs are station specific and can therefore give a more accurate picture of conditions within a basin.

Both the SPI and the PDSI show that the D4 category (the 0 – 2nd percentile) occurs more often in the data than in the USDM depictions. The SPI shows that the D0 category occurs less often in the data than in the USDM depictions. Both of these results could be based on a variety of different factors. First, both indices rely on precipitation and don't take into account soil moisture, streamflow, vegetation or evapotranspiration. These other water variables also weigh heavily on the USDM depiction. Also, impacts and local user input are two qualitative variables that can have an effect on the USDM.

More time will be needed to better assess how the increased monitoring has modified the depiction of drought in the USDM. A longer timeseries (and more data points) could reveal more useful information. Also, an objective way to compare impacts over time with the USDM would also be beneficial. Use of crop data or river calls could probably shed more light on the relationship between impacts and the USDM, and whether or not that relationship has changed since the introduction of increased monitoring. Even so, it is evident with just this study that increased monitoring has helped improve local depictions in the USDM.

References

McKee, T. B., N. J. Doesken, and J. Kleist, 1993: The Relationship of Drought Frequency and Duration to Time Scales. *Eight Conference on Applied Climatology*, Anaheim, CA.

U.S. Drought Monitor, 2011: U.S. Drought Monitor Archives. http://droughtmonitor.unl.edu/archive.html.

Climate Prediction Center, 2011: Monitoring and Data: Drought Monitoring. http://www.cpc.ncep.noaa.gov/products/monitoring_and_data/drought.shtml.

Natural Resources Conservation Service, 2011: Historical SWSI, Colorado NRCS. http://www.co.nrcs.usda.gov/snow/fcst/watershed/current/monthly/data/historicalswsi.html.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

A Method for Monitoring Meteorological Drought at Daily Scale

Er Lu

Iowa State University, Ames, IA
Nanjing University of Information Science & Technology, Nanjing, China

1. Introduction

Drought is difficult to monitor, and various indexes have been proposed to detect the drought as well as flood (*e.g.*, Dracup *et al.* 1980, Wilhite and Glantz 1985, Keyantash and Dracup 2002, Redmond 2002, Svoboda *et al.* 2002). The index previously widely used in the United States is the Palmer Drought Severity Index (PDSI) (*e.g.*, Palmer 1965, Karl and Knight 1985, Heddinghaus and Sabol 1991), which is based on the anomalies of the supply and demand components in the water balance equation. In addition to precipitation, the PDSI also needs to use temperature and other local hydrological quantities. It has been found that the PDSI has many significant limitations (*e.g.*, Alley 1984, Karl and Knight 1985, Smith *et al.* 1993, Willeke *et al.* 1994, Kogan 1995, McKee *et al.* 1995, Guttman 1998) and is not satisfied in the operational monitoring (*e.g.*, Hayes *et al.* 1999). Mo and Chelliah (2006) made some modifications to improve the PDSI.

The index currently widely used is the Standardized Precipitation Index (SPI) (McKee *et al.* 1993; McKee *et al.* 1995), which measures meteorological flood and drought and uses only precipitation. It compares the precipitation of a year with historical records and uses probabilities to indicate if the precipitation of the year is greater or less than the median precipitation of the historical records. Using the SPI, Hayes *et al.* (1999) examined the 1996 drought in the southern plains and southwestern United States. The SPI can determine whether a specific year of an area is a flood or drought year when compared with historical records, but cannot identify if this is a flood or drought area in space.

The major limitation of the SPI is that it can only be used to determine the flood and drought at scales of a month and longer. The SPI computed by the National Climate Data Center has scales ranging from one month to 24 months. Within the period of a scale, it treats the precipitation days equally and uses the simple average of the precipitation to indicate the general flood and drought situation of the period. The SPI does not take into account the precipitation before the period, hence cannot be used to measure flood and drought at shorter timescales (*e.g.*, less than a week). When this index is used to determine the flood and drought extent of the recent three days, for example, if there is no precipitation in the three days, then these three days may be regarded as a drought period when compared with historical records. However, if there was a very strong precipitation just before this period, then the recent three days may still be in a flood state.

Why should flood and drought be measured over the period of a scale, and the days of the period be treated equally? Although the general flood and drought situations of long periods (*e.g.*, several months or years) are useful for long-term planning, decision-makers of agriculture and other water-related departments also hope to know the present day flood and drought situation of a place or area, as well as the flood and drought tendency in the coming days (based on the prediction of precipitation). An index is thus required to measure the daily flood and drought extent. With the daily index, the start, duration, breaks, and strength of flood and drought can be determined. The interannual variabilities and long-term changes of these timings and the strength can further be evaluated. In reality, the flood and drought extent is a hydro-meteorological state of a land-atmosphere system and thus should possess a value at every moment. Once the daily values of the flood and drought extent are determined, the means of a month and longer periods can be calculated.

The purpose of this note is to provide physical considerations for developing such an index to measure the daily flood and drought extent from precipitation with the aim to determine the timings and strength of flood and drought. The rationale is that the flood and drought extent of a day depends on both the precipitation of

Correspondence to: Er Lu, Nanjing University of Information Science & Technology, Nanjing, China; E-mail: lu_er@hotmail.com.

the day and the precipitation of the previous days, but the influences from the precipitation of the previous days are decayed, which is due to the "demands" of the water balance, including the surface runoff, evapotranspiration, groundwater flow, and percolation.

2. Physical considerations of the approach

a. Theoretical framework

The variation process of flood and drought extent can be understood as in the way that the flood extent of a day is achieved through superposing the present day precipitation on the basis of the flood extent of the previous day. If there is large precipitation in the present day, flood extent may increase. If the large precipitation is persistent and lasts for several days, the flood extent will keep increasing, though gradually, and reach large values. Then in the dry days after the persistent precipitation, the flood extent will go down gradually to smaller values. This is due to the inherent decay mechanism of the soil-land surface system, which is caused by the demands of the water balance. Among the demands, runoff can be important during or just after the precipitation days, but in the later dry days evapotranspiration can be important.

Based on the above understanding, a simple physical model can be used to describe the variation of the flood extent. It uses only precipitation, but the general effect of the demands of water balance is considered. The model can be expressed as

$$\frac{df(t)}{dt} = -bf(t) + P(t), \tag{1}$$

where t is the time with the present moment being 0, and $f(t)$ the flood extent at time t in a location with its change being forced by precipitation $P(t)$. The total effect of the decay of flood extent from the demands of runoff, evapotranspiration, groundwater flow, and percolation is represented with $-bf(t)$, in which $b > 0$ measures the strength of the decay.

Integrating equation (1) with t from $-\infty$ to 0 yields

$$\left[e^{bt} f(t)\right]_{-\infty}^{0} = \int_{-\infty}^{0} e^{bt} P(t)dt, \tag{2}$$

or

$$f(0) = \int_{-\infty}^{0} e^{bt} P(t)dt, \tag{3}$$

which indicates that the flood extent at the present moment depends on the precipitation of the moment and all the earlier time, but the influence of the earlier time precipitation is reduced.

To use daily precipitation data, denote $t = -n\Delta t$, where n is the number of the day prior to the present day, and Δt = 1 day. Then (3) can be written as

$$f_0 = \Delta t \sum_{n=0}^{\infty} a^n P_n, \tag{4}$$

where $a = e^{-b\Delta t} < 1$ represents the strength of the contribution of the previous day precipitation P_1 to the present day flood extent f_0.

For practical use, (4) can be truncated as

$$f_0 = \Delta t \sum_{n=0}^{N} a^n P_n, \tag{5}$$

where N is the number of the earlier days used in the calculation, and can be determined from the parameter a and the c, a value much smaller than 1 (*e.g.*, 1%), indicating the precision of the truncation (the contribution fraction of the last day precipitation). With $a^N = c$, N can be determined as $N = \ln c / \ln a$.

The index being proposed for measuring the daily flood and drought extent can be in two forms. One is the *cumulation of reduced precipitation* (CRP), being defined from (5) as

$$CRP \equiv \sum_{n=0}^{N} a^n P_n \,, \tag{6}$$

which means that the flood extent of a day is contributed by the precipitation of the day on the basis of the reduced precipitation of the earlier days. The contribution strength of the earlier day precipitation relies on the parameter a. The CRP measures the absolute flood and drought extent, and can be used to study the spatial variability of the flood and drought extent.

The other form of the index is the *weighted average of precipitation* (WAP), which is defined as

$$WAP \equiv \sum_{n=0}^{N} a^n P_n \Big/ \sum_{n=0}^{N} a^n \,. \tag{7}$$

The SPI and the present day precipitation can be regarded as the two extreme cases of the general form of the WAP (in which a is between 0 and 1, and N can vary). When the contribution parameter a tends to be zero, the present day flood extent WAP will become the present day precipitation P_0. When a tends to be 1, WAP will be equivalent to the simple average of the precipitation of the present day and some earlier days as used in the SPI. The different length of the period N +1 (e.g., one or several months), over which the average is made, represents the different timescales of the SPI.

The WAP in (7) can further be written as

$$WAP = \sum_{n=0}^{N} w_n P_n \,, \tag{8}$$

where the weight $w_n = (1-a)a^n / (1-a^{N+1})$ can be simplified as

$$w_n = (1-a)a^n \,, \tag{9}$$

since $a^{N+1} = ac << 1$. The WAP measures the relative flood and drought extent of a specific location or area, and can be used to study the temporal variability of the flood and drought extent. For seasonal variations of a place, the start, duration, and strength of flood and drought can be determined. The interannual variabilities and long-term changes of the timings and strength can then be calculated.

b. Examples of using the WAP index

Since the major issue of flood and drought extent is the evaluation of its temporal variability at a place (*e.g.*, the start, duration, and strength of flood and drought and their interannual variations), examples of using the WAP index are provided here. Figure 1 presents an ideal time series of daily precipitation *P*, and shows how the calculated daily flood index WAP responds under different values of contribution parameter *a*. There are two episodes of large precipitation (10mm/day) in the 65-day period, and the remaining days in the period and the earlier days before day 1 all have small precipitation (1mm/day). In the first a few days after day 6, although precipitation has become large, the index WAP increases gradually from its

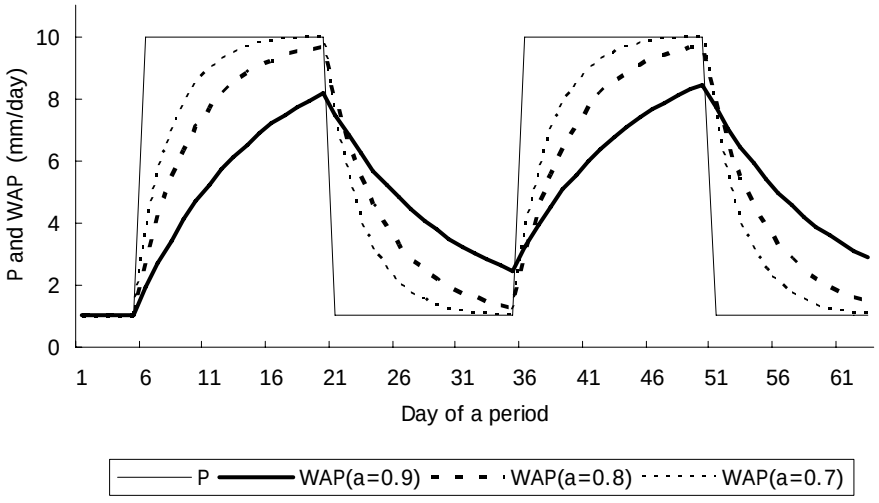


Fig. 1 An ideal time series of daily precipitation P and the calculated index WAP with the contribution parameter a being 0.9, 0.8, and 0.7.

small value of 1mm/day because of the small precipitation in the earlier days. The increase of WAP is relatively slow with *a* being 0.9. For smaller *a*, the increase of WAP is more rapid, so its curve is closer to that of *P*. As mentioned above, when *a* tends to be zero, the curve of WAP will tend to be the same as that of precipitation. With the persistent large precipitation, WAP increases steadily and becomes much larger at day 20. Similarly, WAP decreases gradually in the first a few days after day 21, when precipitation has become small. After a persistent small precipitation till day 35, WAP becomes very small. The decrease of WAP is also more rapid for smaller *a*. So, with the smoothing effect of the index, flood (large WAP) can occur after persistent strong precipitation, and drought (small WAP) may appear after persistent small or no precipitation.

Figure 2 shows an example of analyzing the seasonal and interannual variations of the flood and drought extent by use of the index WAP. The area studied is the upper Mississippi River basin that covers part of Minnesota, Wisconsin, Iowa, Illinois, and Missouri (*Lu et al.* 2009). The data, distributed by the National Climatic Data Center (www.ncdc.com), include the daily precipitation of over 500 observation stations in the basin from 1979 through 2004. A simple average of the daily precipitation is made over the stations of the basin. With taking *N* as 365, daily values of the index WAP from 1980 to 2004 are obtained. The seasonal cycles of the index in 1988 and 1993, when there were severe drought and flood respectively in the basin (*Glantz* 1988; *Changnon* 1996), and the averaged seasonal cycle of the 25 years are illustrated in Figure 2 with the parameter *a* being 0.9 and 0.8, respectively. It shows clearly the contrast of the flood and drought extent between the warm seasons of the two years. The different values of the parameter do not influence much the detections of the flood and drought. The values of the index in most of the days in the warm season of 1993 (1988) are greater (less) than the climatic values. The advantage of this index is that, compared with the SPI, the starting date and the duration of the flood and drought as well as the breaks in the duration can be determined with certain criteria. The comparison of the seasonal cycles of the two years with the climatic seasonal cycle indicates that, roughly, the flood of 1993 and the drought of 1988 both start from the first half of April. The flood of 1993 lasts until the end of September, and the drought of 1988 maintains for longer time. The breaks of the flood and drought can also be captured.

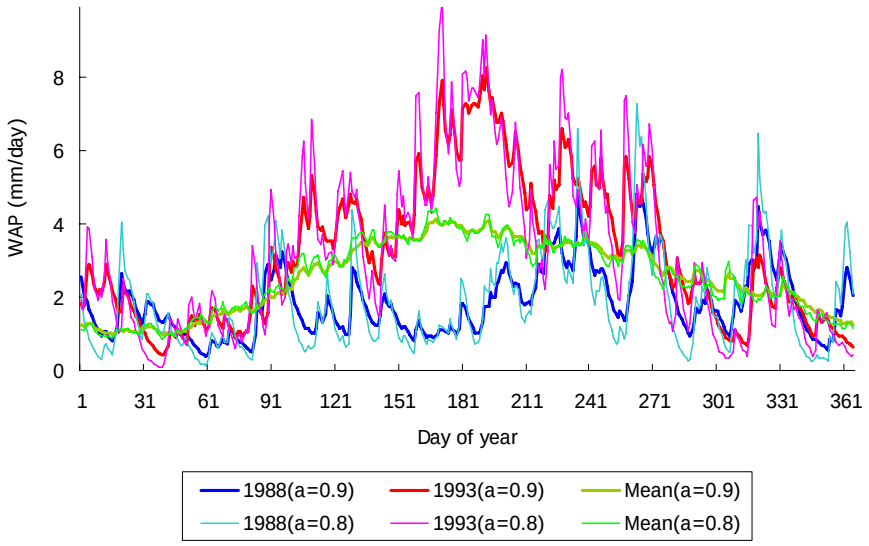


Fig. 2 Seasonal cycles of the daily index WAP in 1988, 1993, and the year averaged over 1980-2004 calculated from the observed daily precipitation averaged over the upper Mississippi River basin with the contribution parameter *a* being 0.9 and 0.8.

c. Contribution and weight parameters

The spatial variability of flood and drought extent can be examined with the index CRP, which measures the absolute flood and drought extent, and its value is dependent on the choice of the contribution parameter *a*. For meteorological flood and drought, which are caused by precipitation and consider the general characteristics of soil and land surface (*i.e.*, the general decay effect of the runoff, evapotranspiration, and percolation), an identical parameter can be used for different locations. Different values of *a* (*e.g.*, 0.9, 0.8, or 0.7) can be used respectively to evaluate the flood and drought extent, and this is somewhat like the different scales used in the SPI. If the specific characteristics of soil and land surface of each place need to be considered, then different values of *a* can be used for different places. The *a* can be chosen empirically since, after all, the flood and drought extent is what needs to be quantified, and there is no existing data of daily

flood and drought extent to fit the parameter. It can be chosen by determining, based on the knowledge of the memory decay strength of the local soil-land surface system, how long it needs to take to make the influence of precipitation reduced to half. For example, in the places where the time of reduced-to-half is a week, the *a* can be taken as 0.9. With this value for the parameter, the influence of precipitation will be reduced to a quarter after 2 weeks and 1% after a month and half. This is in general consistent with the statement that the anomaly of soil moisture caused by a heavy rain or a dry period may take weeks to months to dissipate (Mahanama and Koster 2003).

For a specific location or area, the temporal variability of the flood and drought can be studied by using the index WAP. Compared with CRP, the relative flood and drought index WAP is less affected by the parameter *a*. When a increases, a^n increases for the earlier days but (1 - *a*) decreases, hence the weight w_n does not change much. Moreover, based on the change rate of w_n with a derived from (9)

$$\frac{dw_n}{da} = \left(\frac{1-a}{a} n - 1 \right) a^n , \qquad (10)$$

the change rates are negative in the recent days when $n < a/(1-a)$ (*e.g.*, $n < 9$ for $a = 0.9$) but are positive in the earlier days when $n > a/(1-a)$ (*e.g.*, $n > 9$ for $a = 0.9$). Thus, overall, the WAP does not change much with *a*. This property has been demonstrated in Figure 2. The difference of the curves of WAP with different *a* (0.8 and 0.9) is that the day-to-day variation of WAP with l arger *a* is smoother than that with smaller *a*. Nevertheless, the seasonal cycles of WAP are very close and can give the same results in detecting the flood and drought, including their start, duration, breaks, and strength.

The *a* of being 0.9 is therefore suggested for the general use of the index WAP. Taking the truncation precision *c* as 1%, *N* is calculated to be 44. The index thus becomes

$$WAP = 0.1 \sum_{n=0}^{44} 0.9^n P_n . \qquad (11)$$

d. Determining the start, duration, and strength

Zeng and Lu (2004) designed a method to determine objectively the monsoon onset and retreat dates from seasonal cycles of precipitable water with certain criteria. A similar approach can be used to determine the start and end dates (and thus the duration and strength) of the flood and drought from seasonal cycles of WAP. For a given location or area, the daily precipitation data of the 26 years, for example (as in Figure 2), are first converted into daily WAP values that include 25 years. The maximum and minimum of the WAP of each year, and thus their averages over the 25 years, are then calculated. The difference ranging from the averaged minimum to the averaged maximum can be divided into several grades, representing a classification from the severest drought to the severest flood. The start and end dates of the flood and drought for a specific severity grade can be determined with proper criteria, *e.g.*, when the daily WAP exceeds (is lower than) the grade for 3 consecutive days for the start (end). The strength of the flood or drought averaged over the determined flood or drought duration can be calculated. With the start, duration, and strength of flood and drought determined for each year, their interannual variabilities and long-term changes can be studied. By using this methodology, which has been partly applied in Figure 2, evaluations for different regions of the globe will be presented separately.

3. Summary and discussions

The SPI is commonly used for detecting flood and drought. When it is calculated, a timescale longer than a month needs to be selected first. The days in the period of the scale are treated equally with precipitation simply averaged over the period, while the precipitation in the days before the period is not considered. Because of these, the SPI can determine only the general flood and drought situation of a long period (*e.g.*, several months or years) but cannot be used for short scales (*e.g.*, a week or less). However, the operational monitoring and decision-making do require an index to measure the daily flood and drought extent. With a daily index, the start, duration, and strength of the flood and drought of a year can be determined, and their interannual variabilities and long-term changes as well as the associated mechanisms can be studied. As an

objective state of the land-atmosphere system, the flood and drought extent should have an instantaneous value and not just have an overall condition of a long period.

Physical considerations of developing such an index for measuring daily flood and drought extent is provided in this note with a simple physical model. It uses only precipitation as the SPI does, but it also considers the demands of the water balance as the PDSI does. The principle of the physical model is that the flood extent is forced by precipitation, but dissipated by the demands of the water balance. Among the demands, runoff, groundwater flow, and percolation can be important during or just after a precipitation, while evapotranspiration can be important in the later dry time. What differs from the PDSI is that in this study, the overall dissipation or decay effect of these demands is represented with a parameterization of the flood extent. To be simple, but still fairly reasonable, the parameterization is currently taken as a linear form. The defined index, based on the solution of the physical model, has two forms, and both are easy to calculate.

The cumulation of reduced precipitation (CRP) can be used to analyze the spatial variability of flood and drought extent. The value of CRP can be affected by parameter ***a*****,** the contribution strength of the earlier day precipitation. For meteorological flood and drought, which stress precipitation and the general effect of the demands, the parameter can be identical for different locations. If local hydrological and geographic conditions need to be considered for each specific location, the parameter can be given empirically based on the local characteristics, since the daily flood and drought extent is what needs to be quantified.

The advantage of the weighted average of precipitation (WAP) is that, as shown both analytically and from the observation, its value is less affected by the parameter *a*. The WAP measures the relative flood and drought extent as the SPI does, and can be used to evaluate the temporal (*e.g.*, seasonal and interannual) variations of the flood and drought extent of a specific location or area, which is more important than the spatial comparison of the extent. The SPI makes simple average of precipitation, so a timescale is required. The WAP makes weighted average of precipitation with the weight decreasing with the number of the days past, hence no timescale is needed. Figure 2 shows that the results of the start, duration, and strength of flood and drought and their year-to-year variations calculated with two different values of *a* are very close. The *a* of being 0.9 is finally suggested for the calculation of the WAP.

Modifications could be made in future studies to the current forms of the index to better reflect the flood and drought situations in places with different hydro-climate and geographical conditions. A relatively complicated parameterization *D*(*f*) may be developed, with knowledge of the local characteristics of a specific place, to better represent the dissipation effect of the demands. The physical model then becomes $df/dt = D(f) + P$, and this equation can be solved numerically. Modifications may also be made to the expression of the WAP with designing a new weight *W*(*n*), and the index is then expressed as $WAP = \sum_{n=0}^{N} W(n) P_n$.

The precipitation in cold season of some areas may be in form of snow, partially or totally, and later the snowmelt can contribute with rainfall to enhance the flood extent. It is better to replace the precipitation in the index with the combination of rainfall and snowmelt. This change may have some influence to the values of the index in snow and snowmelt seasons, but has little influence during the warm season. The physical model of this study can also be applied to the scale less than a day to study the diurnal variation of the flood extent, which may be large when there are thunderstorms.

The purpose of this note is to provide the physical basis of the daily flood and drought index and the general methodology of using the index to determine the start, duration, and strength of flood and drought from an example as well as their interannual variabilities and long-term changes. More calculations and detailed analyses will be presented in the follow-on studies for different regions of the globe.

References

Alley, W. M, 1984: The Palmer Drought Severity Index: Limitations and assumptions, *J. Clim. Appl. Meteorol., 23*, 1100–1109.

Changnon, S. A. (Ed.), 1996: *The Great Flood of 1993: Causes, Impacts, and Responses*, 321 pp., Westview, Boulder, Colo.

Dracup, J. A., K. S. Lee, and E. G. Paulson, 1980: On the definition of droughts, *Water Resour. Res., 16*, 297–302.

Glantz, M. H. (Ed.), 1988: *Societal Responses to Regional Climatic Change: Forecasting by Analogy*, 428 pp., Westview, Boulder, Colo.

Guttman, N. B., 1998: Comparing the Palmer Drought Index and the Standardized Precipitation Index, *J. Am. Water Resour. Assoc., 34*, 113-121.

Hayes, M. J., M. D. Svoboda, D. A. Wilhite, and O. V. Vanyarkho, 1999: Monitoring the 1996 drought using the Standardized Precipitation Index, *Bull. Am. Meteorol. Soc., 80*, 429–438.

Heddinghaus, T. R., and P. Sabol, 1991: A review of the Palmer Drought Severity Index and where do we go from here? Proceedings, 7th Conf. on Appl. Climatol., 10-13 September 1991, Boston, American Meteorological Society, 242-246.

Karl, T. R., and R. W. Knight, 1985: Atlas of monthly Palmer Hydrological Drought Indices (1931–1983) for the contiguous United States, *National Climatic Data Center Historical Climatology Series 3-7*, Asheville, NC, 319 pp. [Available from National Climatic Data Center, Federal Building, 151 Patton Ave., Asheville, NC 28801-5001.]

Keyantash, J., and J. A. Dracup, 2002: The quantification of drought: An evaluation of drought indices, *Bull. Am. Meteorol. Soc., 83*, 1167–1180.

Kogan, F. N., 1995: Droughts of the late 1980s in the United States as derived from NOAA polar-orbiting satellite data, *Bull. Am. Meteorol. Soc., 76*, 655–668.

Lu, E., E. S. Takle, and J. Manoj, 2009: The relationships between climatic and hydrological changes in the Upper Mississippi River Basin: A SWAT and multi-GCM study, *J. Hydrometeorol.*, accepted.

Mahanama, S. P. P., and R. D. Koster, 2003: Intercomparison of soil moisture memory in two land surface models, *J. Hydrometeorol., 4*, 1134–1146.

McKee, T. B., N. J. Doesken, and J. Kleist, 1993: The relationship of drought frequency and duration to time scales, Preprints, 8th Conference on Applied Climatology, pp. 179–184, January 17–22, Anaheim, California.

McKee, T. B., N. J. Doesken, and J. Kleist, 1995: Drought monitoring with multiple time scales, Preprints, 9th Conference on Applied Climatology, pp. 233–236, January 15–20, Dallas, Texas.

Mo, K. C., and M. Chelliah, 2006: The modified Palmer Drought Severity Index based on the NCEP North American Regional Reanalysis, *J. Appli. Meteorol. Climatol., 45*, 1362–1375.

Palmer, W. C., 1965: Meteorological drought, Research Paper No. 45, U. S. Department of Commerce Weather Bureau [NOAA Library and Information Services Division, Washington, D. C. 20852].

Redmond, K. T., 2002: The depiction of drought, *Bull. Am. Meteorol. Soc., 83*, 1143–1147.

Smith, D. I., M. F. Hutchinson, and R. J. McArthur, 1993: Australian climate and agricultural drought: Payments and policy, *Drought Network News, 5*, 11–12.

Svoboda, M. D., and Coauthors, 2002: The drought monitor, *Bull. Am. Meteorol. Soc., 83*, 1181–1190.

Wilhite, D. A., and M. H. Glantz, 1985: Understanding the drought phenomenon: The role of definitions, *Water International, 10*, 111-120.

Willeke, G., J. R. M. Hosking, J. R. Wallis, and N. B. Guttman, 1994: The National Drought Atlas, *Institute for Water Resources Report 94–NDS–4*, U. S. Army Corps of Engineers.

Zeng, X., and E. Lu, 2004: Globally unified monsoon onset and retreat indexes, *J. Clim., 17*, 2241-2248.

3. EXTREME WEATHER AND CLIMATE EVENTS

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

A Method for Identifying the Events That Can Best Become Extremes

Er Lu[1, 2], Wayne Higgins[1], Kingtse Mo[1], and Michael Halpert[1]

[1]*Climate Prediction Center, NCEP/NWS/NOAA, Camp Springs, MD*
[2] *Nanjing University of Information Science & Technology, Nanjing, China*

1. Introduction

Is climate going crazy under global warming? Are climate extremes more frequent in the warmed climate? Climate change and extremes have become one focus of climate research (*e.g.*, Karl and Knight 1997; Meehl *et al*. 2000; Easterling *et al*. 2000a; Dairaku *et al*. 2004; Allan and Soden 2008), and several international conferences have recently been organized to stress these hot topics. The workshop held in 2007 in Hawaii reviewed the understanding and prediction of extreme events and of changes in their frequency and intensity (Garrett and Müller 2008). The workshop on "metrics and methodologies of estimation of extreme climate events" held in 2010 in France sponsored by the World Climate Research Programme (WCRP) and the United Nations Educational, Scientific and Cultural Organization (UNESCO) pointed out that good statistical methods are essential for exploring, defining, and estimating weather and climate extremes (Zolina *et al.* 2010). Climate extremes include heavy precipitation, floods, heat waves, droughts, storm surges, and hurricanes, among many others. The Climate Extremes Index (CEI) proposed by Karl *et al*. (1996) has been used to indicate the overall extreme situations of the climate.

Suitable methods are required to find out the extremes from climate data (*e.g.*, Karl *et al*. 1996; Mudelsee 2006). The detection of extremes for a specific climatic quantity, *e.g.*, precipitation, has been investigated in many studies. One of the focuses is to analyze the extremes in daily precipitation (*e.g.*, Karl and Knight 1998; Klein Tank and Können 2003; Zhai *et al.* 2005; Wang *et al*. 2008). There are two problems with the previous daily precipitation analyses. One is that the extremes detected depend on the choice of the starting time of the day. The other is that the extremes are only for the 1-day (24-hour) duration, not including those over other durations (*e.g.*, the shorter durations of 12 or 6 hours and longer durations of 2 or 4 days). A precipitation event that cannot be an extreme over the 1-day duration may become extremes over other durations. These extremes with different durations can all bring, in one way or the other, serious economic and societal damages.

The extremes in multiday precipitation have been detected generally based on the lifespan of rainfall (Karl and Knight 1998; Dairaku *et al*. 2004; Junker *et al*. 2008). Junker *et al*. (2008) pointed out that the starting and ending times of multiday events can be difficult to determine when precipitation comes as a result of several consecutive storms with small breaks between them. It is possible that the intensity averaged over the entire multiday rainfall period is not sufficiently strong to be an extreme, but the intensity over part of the period is relatively strong enough to become an extreme.

Figure 1 shows conceptually some 12-day precipitation processes. In Figure 1a, although the precipitation in each of day 6 and day 7 may not be an extreme, the precipitation of these two days may possibly be an extreme. For the precipitation from day 5 to day 9 in Figure 1b, though the precipitation in each day (or every consecutive 2 days) of the period may not be an extreme, this 5-day precipitation event may be an extreme. The precipitation process in Figure 1c may be best detected as an extreme over a 9-day duration (it is also acceptable if treated as over an 8-day or 10-day duration), but relatively it should not be regarded as a 1-day or 3-day extreme.

The issue investigated in this study is, for a precipitation process, to identify the starting time and the duration so that the intensity of the event with the starting time and duration (the intensity averaged over the period with the starting time and duration) can best become an extreme, compared with those corresponding

Correspondence to: Er Lu, Nanjing University of Information Science & Technology, Nanjing, China; E-mail: lu_er@hotmail.com.

to other durations and starting times. In other words, extreme events should be "best described", in terms of starting time and duration, which can be determined based on intensity.

Extremes, along with drought and monsoons, are important objects of climate monitoring. Lu and Chan (1999) and Zeng and Lu (2004) proposed methods to determine the strength of monsoon and the onset and retreat dates of monsoon. Lu (2009) developed a methodology to monitor and predict drought. It is based on theoretical considerations and mathematical derivations, and the single parameter contained in the relation can be determined with data. The present study on detecting extremes is similar in style. The key is to establish a relation that prescribes how "extreme" intensity varies with duration.

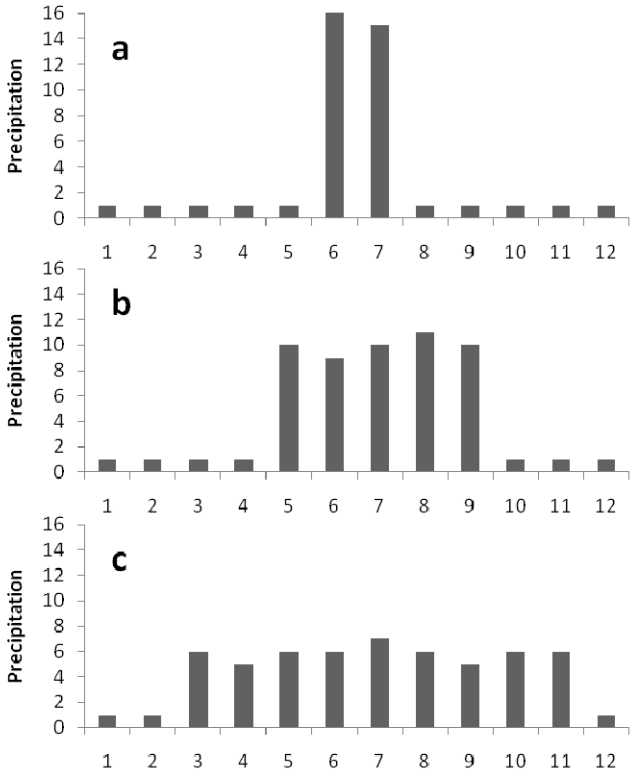


Fig. 1 Conceptual cases that may contain extremes over multi-day durations. An appropriate unit can be given to precipitation.

2. Best describing extremes and capturing them across a range of durations

a. Theoretical relation of "extreme" intensity with duration

Let's first consider the case of setting discount rates in stores for promotion. The general consideration for setting the rates of a product is that for the buyers to buy more, the unit price of the product should decrease with the number of the pieces to be bought in order to save money. However, for the sellers to gain more profit, the total money received from the sales should increase with the number of the pieces to be sold.

The same principle can be applied to establish the relation between "extreme" intensity and duration. The first constraint is that the "extreme" intensity I_e should decrease with the duration T, which can be expressed as

$$\frac{dI_e}{dT} < 0 . \tag{1}$$

The second constraint is that, though with a weaker intensity, the total amount over a longer duration should be larger. Or, the product of the intensity and duration should increase with duration, and this can be expressed as

$$\frac{d(I_e T)}{dT} > 0 . \tag{2}$$

Combining relations (1) and (2) yields

$$0 < a < 1, \tag{3}$$

where

$$a = -\frac{1}{I_e}\frac{dI_e}{d(\ln T)} \tag{4}$$

indicates the relative decrease rate of "extreme" intensity with respect to the logarithm of duration.

Based on the definition and meaning of the parameter a, it is reasonable to assume it as a constant, although the values of the parameter estimated from data over different duration ranges (based on certain definition of the extremes in the data) may change, as will be discussed in section c.

Denote $T = n\,\Delta T$, where ΔT is the increment in duration (*e.g.*, 1 day or 1 hour) based on the time resolution of the data, and n is the number of the increment, reflecting the length of the duration. A range of durations with n from 1 to N (*e.g.*, from 1 to 7) may be considered to capture extremes more completely.

Rewrite equation (4) as $d(\ln I_e) = -a\, d(\ln T)$. Integrating this equation, with T from ΔT to $n\Delta T$ and I_e correspondingly from $I_e(1)$ to $I_e(n)$, yields

$$I_e(n) = I_e(1)\, n^{-a}, \tag{5}$$

where $I_e(1)$ is the maximal "extreme" intensity that is over the shortest duration limited by the time resolution of the data.

Note that the duration here is not the lifespan of rainfall, but the time period to be found to make the precipitation intensity averaged over the period become an extreme. The duration can be part of the lifespan of rainfall, and may even contain a rainfall break.

It should also be noted that for the purpose of best describing extremes in this study, the intensity always decreases with duration, and the parameter a is always positive. Differently, in the intensity-duration-frequency (IDF) studies, the duration for examining the extremes (the events with long return periods) is the actual lifespan of the events. The Sherman's equation on IDF has the same form as equation (5), but the parameter a in the equation may take negative values. The reason is that the precipitation events that may become extremes should normally last for a certain period of time (*e.g.*, 30 minutes), and the events that last for shorter periods are generally weaker in intensity (*e.g.*, Hershfield 1972).

b. Best describing extremes with starting time and duration

Denote the intensity calculated from the data over duration $n\Delta T$ centering at time $m\Delta T$ as $I(n, m)$. The relative intensity of the event is defined as

$$R(n,m) \equiv \frac{I(n,m)}{n^{-a}} = n^{a} I(n,m). \tag{6}$$

The purpose of defining the relative intensity is to enlarge the data-calculated intensities of longer durations with the theoretically-derived "extreme" intensity-duration relation so that the intensities of the events over all other durations can be compared with the one over the shortest duration. Through the running of m and the comparison among different values of n, the possible extreme, which has the strongest relative intensity $R(n_0, m_0)$ in a process, can be identified. With the values of n_0 and m_0, the starting time and the duration of the extreme can be determined.

Figure 2 presents the relative intensities with input data from Figure 1. The parameter a takes the moderate values of 0.4, 0.5, and 0.6 to better obey the relations (1) and (2). The data before day 1 and after day 12 are all taken as zero in the calculation. Nine durations (1, 2, 3, 4, 5, 6, 7, 9, and 11 days) are included in the plot for comparing the relative intensities. For the convenience to make the plot, the relative intensity given to day m uses the data from day $m - (n - 1)/2$ to day $m + (n - 1)/2$ if the duration n is an odd number, but from day $m - (n - 2)/2$ to day $m - n/2$ if n is an even number.

Figure 2a shows that the precipitation event can be best described as an extreme (with strongest relative intensity) over the 2-day duration. It is also acceptable if regarded as over the 3-day duration. Both of these can be concluded from the plots with the three values of the parameter a. Whether the extreme can be regarded as over durations of 1-day or 4-days depends on the choice of the parameter. In Figure 2b, results from all the three values of the parameter show that the extreme is over the 5-day duration, and it is still acceptable if regarded as over the 6-day duration. However, whether it can be regarded as over durations of 4 days or 7 days depends on the value of the parameter. In Figure 2c, the different parameter values all indicate an extreme over duration of 9 days. It is acceptable if regarded as over durations of 11 or 7 days, or even 6 or

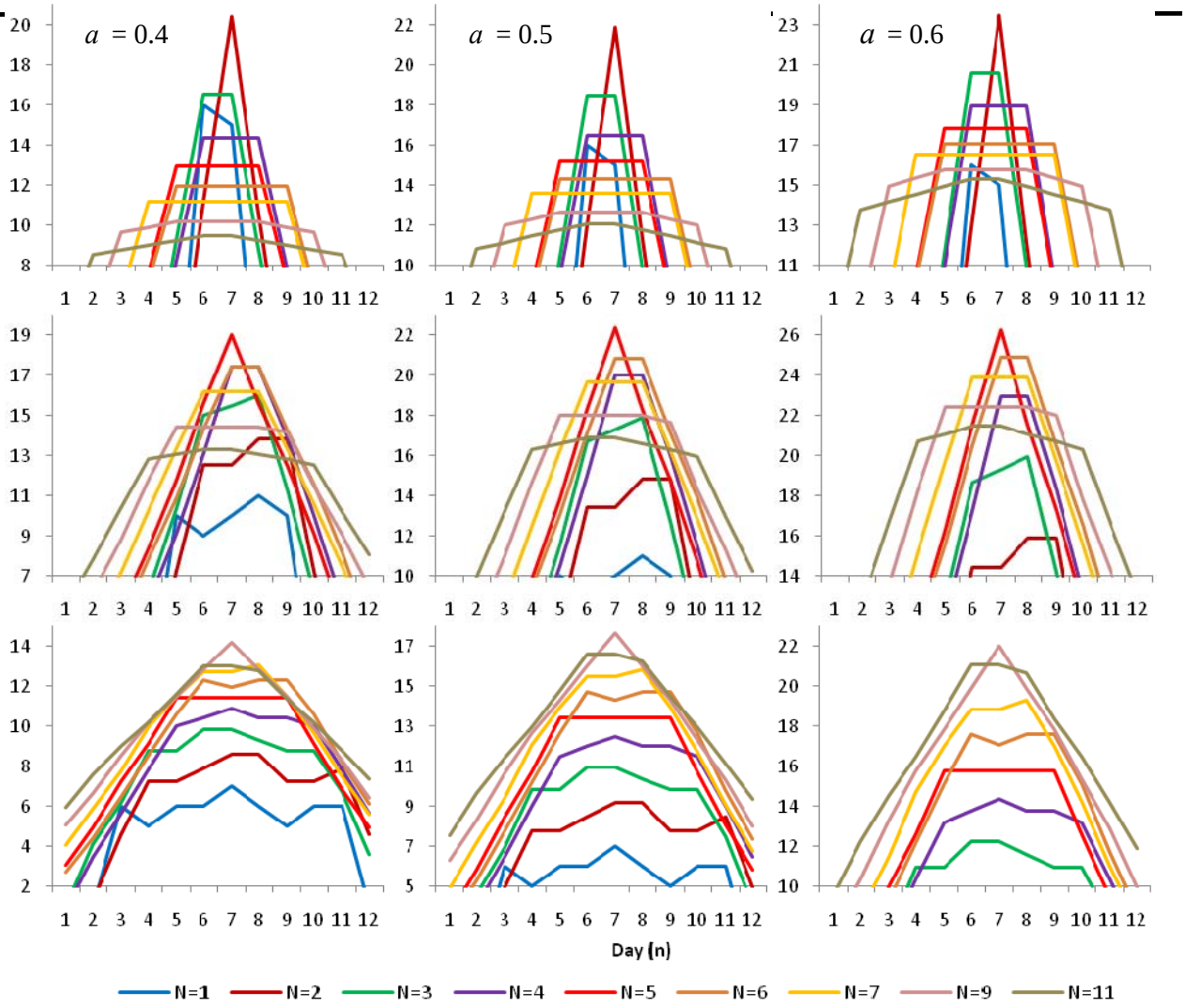


Fig. 2 Relative intensities over 9 durations (1, 2, 3, 4, 5, 6, 7, 9, and 11 days) with input data from Figure 1 and parameter a being 0.4, 0.5, and 0.6, respectively. Note that the curves with even numbers of duration have been made a half-day rightward shift for convenience of making the plots.

5 days, but definitely not 1 or 2 days. Along with the durations, the starting times for the best possible extremes can also be determined based on the maximal relative intensities in the plots and their corresponding values of n and m.

What are shown in Figure 1 are single precipitation processes with limited data. The results suggest that the present method can well detect these typical extremes with different durations through simply giving moderate values to the parameter a, and the detections are not very sensitive to the parameter.

c. Determining the parameter with regression from data

For climatic detection with multi-year data, the value of the parameter a can be determined from the data. Based on equation (5), a regression between the logarithms of duration n and "extreme" intensity I_e (n) can be established as

$$\ln I_e(n) = -a \ln n + c, \qquad (7)$$

where a is the parameter to be determined, and the constant c can also be determined from the data.

For a specific duration n, which varies from 1 to N, the "extreme" intensity I_e (n) over the duration can be defined as the intensity that is the strongest 5% of the intensities that are over all the n consecutive days within the multi-year data. With all these values of n and the such-defined I_e (n), the values of a and c can be

regressed from equation (7). Then, with the parameters a and c determined, the final value of each "extreme" intensity $I_e(n)$ can be calculated from the duration n with equation (7).

Note that due to the possible irregular structure in the data's spectrum of intensity over duration, the values of the parameter a determined from equation (5) with the data-defined $I_e(1)$ and $I_e(n)$ may be quite different. The advantage of using the above regression is that the determination of the parameter takes into account comprehensively the overall structure of the intensity spectrum. Also note that the constant c might have large deviation from the data-defined $I_e(1)$, depending on the structure of the intensity spectrum.

d. Capturing extremes across a range of durations

For assessing and understanding the changes in weather and climate extremes, it would be more complete to detect the extremes across a range of durations rather than just over a single (1-day) duration. With the running of the date, if the data-calculated intensity is greater than the "extreme" intensity obtained from regression (7), then an extreme is detected. If precipitation is particularly strong and persistent during a process so that the data-calculated intensities over several different durations are all greater than their corresponding "extreme" intensities, then the process can be detected as extremes over all these durations.

The annual (or seasonal) total of the extremes over each duration can be determined for each year (season), and the change over the multi-years can be analyzed. To be convenient, an alternative way can be used. That is, compare the data-calculated intensity and the regressed "extreme" intensity for each day, and take the number of the days with $I(n, m) > I_e(n)$ in the year (season) to examine the change of the extremes over the duration.

It would be interesting to investigate the relationships between the changes (trends) in the numbers of the extremes over different durations. If daily precipitation extreme has an increasing change, will the extremes over durations of 2 and 4 days have increasing changes either? In some areas, *e.g.*, in Japan, there was no increase in the seasonal total precipitation, but there was an increase in the frequency of the 1-day precipitation extremes (Easterling *et al.* 2000b). The total number of the extremes across a range of durations can also be analyzed to explore more completely the change of precipitation extremes.

In addition to the numbers of the extremes over different durations, the sum of the relative intensities, from equation (6), of all the extremes during the year (season) may also be utilized to assess the changes in the extremes.

3. Summary and discussions

The extremes detected by using daily precipitation data in the previous studies mainly include the following three types: the daily precipitation extremes; the extremes in monthly, seasonal, and annual precipitation totals; and the extremes in the continuous multiday rainfall processes. These studies had not made full use of the daily data. Many researchers examined the extremes in daily precipitation, and what they actually concerned about is not just the extremes that have the exact 24-hour duration; the major reason of finding the daily extremes is that they were using the daily precipitation data. Events with other durations (*e.g.*, 2 or 4 days) may also be important in bringing losses as long as they can be sufficiently strong (relative to their durations), and thus can become extremes.

Therefore, for the purpose of studying the changes of whether and climate extremes, it is far from complete to merely consider a single duration of a day, and it is more suitable to find out all the extremes that are over at least a range of durations (*e.g.*, from 1 to several days). The events that cannot become daily extremes might become extremes over other durations. The goal of this study is, for a precipitation process, to determine the duration and starting time of an event (within the process) that can best become an extreme.

The key of the approach in this study is to prescribe reasonably the intensities of the extremes that are over the different durations. Theoretically, the strongest intensity should be found from the instantaneous values of intensity. In the practical analysis with data, the strongest intensity can be obtained from the shortest duration considered. The value of the intensity averaged over a longer duration should go down. So, the "extreme" intensity always decreases with duration.

The other constraint of the "extreme" intensity-duration relation is that, in spite of the decrease of the intensity with duration, the total accumulation of precipitation should increase with duration. These are reasonable considerations, and might be applied to many other problems, such as determining how the intensity of flood or drought that make human or crops unendurable varies with the lasting time of the flood or drought. These constraints ensure that "extreme" intensity does not decrease linearly with duration.

A theoretical "extreme" intensity-duration relation is thus derived. The relation contains only a single parameter, and it can be treated as a constant. The conceptual examples given in this study are just rainfall episodes, but the extremes over the different durations can be well detected with simply giving moderate values to the parameter, and the detections are not very sensitive to the parameter.

For detecting extremes with multi-year data, the value of the parameter can be determined from the data by using the regression between the logarithms of the duration and the corresponding initial "extreme" intensity defined with the data. The final values of the "extreme" intensities can be computed with the parameter determined from the regression equation. It is noticed that the regression relation obtained from the data may not be statistically significant, but this does not matter, since the purpose here is just to prescribe the "extreme" intensities with considering the overall structure of the intensity spectrum.

Through capturing the extremes over different durations with the method of this study, relationships between daily extremes and the extremes over other durations will be analyzed. The changes of the total extremes across a range of durations will also be investigated. The Climate Extremes Index (CEI) even combines the extremes of different quantities.

Although daily precipitation is used in this study as an example, the method can be applied to detect extremes over durations at hourly scales if hourly data are available. The method can also be used to detect extremes of other climate quantities such as heat waves. All these applications and relationship analyses will be carried out in the next-step work by using real observations and data from climate models.

The definition of weather and climate extremes is still an issue in debate, as indicated from the workshops of the recent years, which aimed to summarize, compare, and assess the various definitions and to develop a common framework. Based on the understanding from this research, a definition can be proposed as follows. To a specific weather and climate quantity of interest, an extreme is the event whose intensity corresponding to its starting time and duration is relatively the strongest compared with those with other durations and starting times.

References

Allan, R. P., and B. J. Soden, 2008: Atmospheric warming and the amplification of precipitation extremes. *Science,* **321**, 1481-1484, doi:10.1126/science.

Dairaku, K., S. Emori, and T. Oki, 2004: Rainfall amount, intensity, duration, and frequency relationships in the Mae Chaem watershed in Southeast Asia. *J. Hydrometeorol.*, **5**, 458-470.

Easterling, D. R., J. L. Evans, P. Ya Groisman, T. R. Karl, K. E. Kunkel, and P. Ambenje, 2000a: Observed variability and trends in extreme climate events: A brief review. *Bull. Am. Meteorol. Soc.*, **81**, 417-425.

Easterling, D. R., G. A. Meehl, C. Parmesan, S. A. Changnon, T. R. Karl, and L. O. Mearns, 2000b: Climate extremes: Observations, modeling, and impacts. *Science*, **289**, 2068-2074, doi: 10.1126/science.289.5487.2068.

Garrett, C., and P. Müller, 2008: Extreme events. *Bull. Am. Meteorol. Soc.*, **89**, 1733-1733, doi: 10.1175/2008BAMS2566.1.

Hershfield, D. M., 1972: Estimating the extreme-value 1-minute rainfall. *J. Appli. Meteorol.*, **11**, 936-940.

Junker, N. W., R. H. Grumm, R. Hart, L. F. Bosart, K. M. Bell, and F. J. Pereira, 2008: Use of normalized anomaly fields to anticipate extreme rainfall in the mountains of Northern California. *Wea. Forecasting*, **23**, 336-356.

Karl, T. R., and R. W. Knight, 1997: The 1995 Chicago heat wave: How likely is a recurrence? *Bull. Am. Meteorol. Soc.*, **78**, 1107-1119.

Karl, T. R., and R. W. Knight, 1998: Secular trends of precipitation amount, frequency, and intensity in the United States. *Bull. Am. Meteorol. Soc.*, **79**, 231-241.

Karl, T. R., R. W. Knight, D. R. Easterling, and R. G. Quayle, 1996: Indices of climate change for the United States. *Bull. Am. Meteorol. Soc.*, **77**, 279-292.

Klein Tank, A. M. G., and G. P. Können, 2003: Trends in indices of daily temperature and precipitation extremes in Europe, 1946-99. *J. Clim*., **16**, 3665-3680.

Lu, E., 2009: Determining the start, duration, and strength of flood and drought with daily precipitation: Rationale. *Geophys. Res. Lett.*, **36**, L12707, doi:10.1029/2009GL038817.

Lu, E., and J. C. L. Chan (1999): A unified monsoon index for South China, *J. Clim*., **12**, 2375-2385.

Meehl, G. A., F. Zwiers, J. Evans, T. Knutson, L. Mearns, and P. Whetton, 2000: Trends in extreme weather and climate events: Issues related to modeling extremes in projections of future climate change. *Bull. Am. Meteorol. Soc.*, **81**, 427-436.

Mudelsee, M., 2006: CLIM-X-DETECT: A Fortran 90 program for robust detection of extremes against a time-dependent background in climate records. *Computers & Geosciences*, **32**, 141-144.

Wang, W., X. Chen, P. Shi, and P. H. A. J. M. van Gelder, 2008: Detecting changes in extreme precipitation and extreme streamflow in the Dongjiang River basin in southern China. *Hydrol. Earth Syst. Sci.*, **12**, 207-221.

Zeng, X., and E. Lu, 2004: Globally unified monsoon onset and retreat indexes. *J. Clim.*, **17**, 2241-2248.

Zhai, P., X. Zhang, H. Wan, and X. Pan, 2005: Trends in total precipitation and frequency of daily precipitation extremes over China. *J. Clim.*, **18**, 1096-1108.

Zolina, O., V. Detemmerman, and K. E. Trenberth, 2010: Improving the accuracy of estimation of climate extremes. *Eos Trans. AGU,* **91**, doi:10.1029/2010EO510013.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Primary Factors Contributing to Japan's Extremely Hot Summer of 2010

Nobuyuki Kayaba, Shotaro Tanaka, Shingo Ushida

Climate Prediction Division, Japan Meteorological Agency, Tokyo, Japan

1. Introduction

Many areas in the world experienced high temperatures for boreal summer (June – August) 2010, especially from western Russia to the Middle East, in northern China, from southern Indonesia to southern Polynesia, and in the eastern USA (Figure 1.1). Japan also experienced nationwide record-breaking high temperature. The seasonal mean temperature in Japan for summer 2010 (*i.e.*, the three-month period from June to August) ousted that of summer 1994 from the top spot as the highest since JMA's records began in 1898, with a deviation of +1.64°C from the 1971 – 2000 average (Figure 1.2). Shortly after this period on 3 September, the Japan Meteorological Agency (JMA) organized an extraordinary meeting of the Advisory Panel on Extreme Climate Events (see Chapter 2 for details). This report summarizes the atmospheric characteristics and possible influences identified by the Advisory Panel as main background factors to Japan's hottest summer on record.

2. The Advisory Panel on Extreme Climate Events

The Advisory Panel on Extremely Climate Events was established in June 2007 by JMA to investigate extreme climate events based on the latest knowledge and findings and provide JMA with advice on the causes of extreme climate events. Its members are prominent experts on climate science from universities and research institutes. On a regular basis, JMA provides the panel with observational data and monitoring information on climate system. The panel provides JMA with scientific advice on

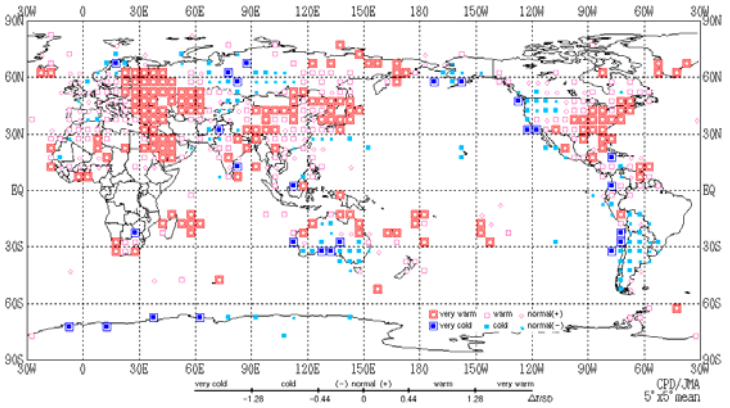


Fig. 1.1 Three-monthly mean temperature anomalies for summer (June – August) 2010. Categories are defined by three-month mean temperature anomaly against the normal (*i.e.*, the 1971 – 2000 average) divided by its standard deviation and averaged in 5° × 5° grid boxes. The thresholds of each category are -1.28, -0.44, 0, +0.44 and +1.28.

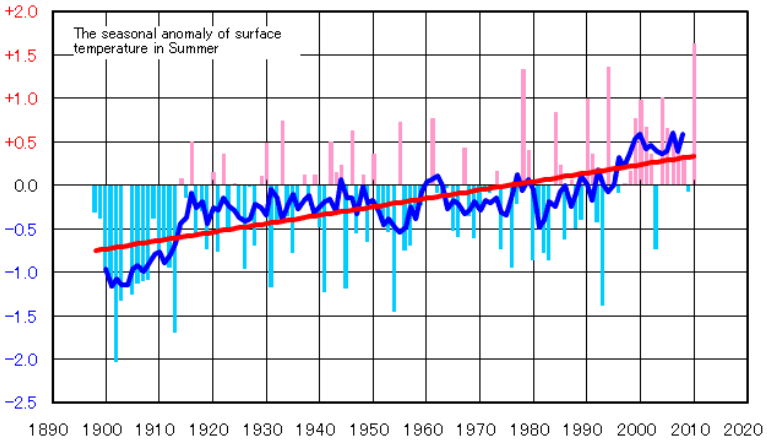


Fig. 1.2 Long-term change in seasonal temperature anomalies for summer (June – August) in Japan. Anomalies are calculated as the average of temperature deviations from the 1971 – 2000 normal at the 17 observation stations. The observatory stations that represent the average temperature of Japan are selected from those deemed to be least influenced by the urban heat island phenomenon. The bars indicate temperature anomalies for each summer. The blue line indicates the five-year running mean, and the red line shows the long-term linear trend.

Correspondence to: Nobuyuki Kayaba, Climate Prediction Division, Japan Meteorological Agency, Tokyo, Japan; E-mail: n-kayaba@met.kishou.go.jp.

climate system monitoring and suggestions on analysis methods. When an extreme climate event occurs or is likely to occur, an extraordinary meeting of the panel is organized to analyze the causes of the event and JMA provides timely statement on the event based on advice from the panel (Figure 2.1).

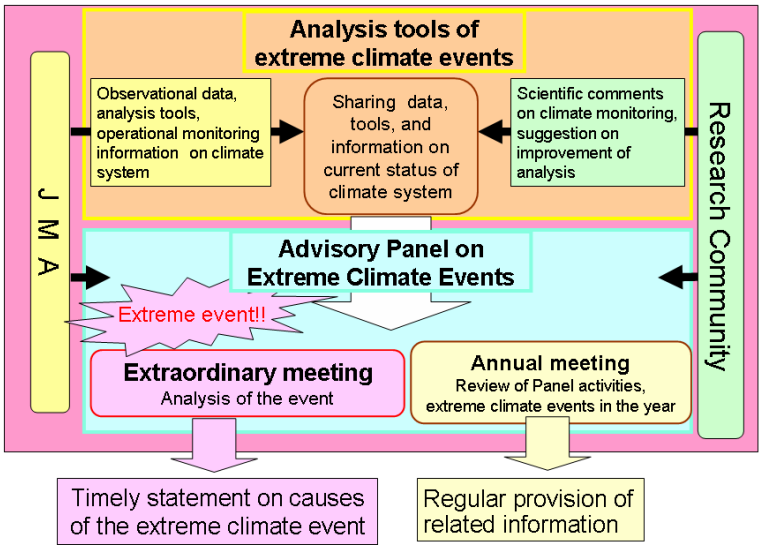


Fig. 2.1 Schematic chart of the framework at the Advisory Panel on Extreme Climate Events.

3. Data source

The datasets used in this analysis are as follows:

1) Surface observational data: JMA's in-site observation.
2) Sea surface temperature (SST) data: the JMA's SST analysis for climate monitoring (COBE-SST) (Ishii *et al.* 2005).
3) Atmospheric circulation data: the Japanese reanalysis (JRA-25/JCDAS) (Onogi *et al.* 2007).
4) Outgoing longwave radiation (OLR) data: original data provided by the National Oceanic and Atmospheric Administration (NOAA).

The base period for normal is 1971 – 2000 for 1) and 2) datasets, and it is 1979 – 2004 for 3) and 4) datasets.

4. Ocean conditions and convective activity in the tropic

In summer 2010, a La Niña event followed the El Niño period that started in summer 2009 and ended in spring (March – May) 2010 (Figure 4.1). In association with this, sea surface temperatures (SSTs) were above normal over the western Pacific, the Indian Ocean and the tropical

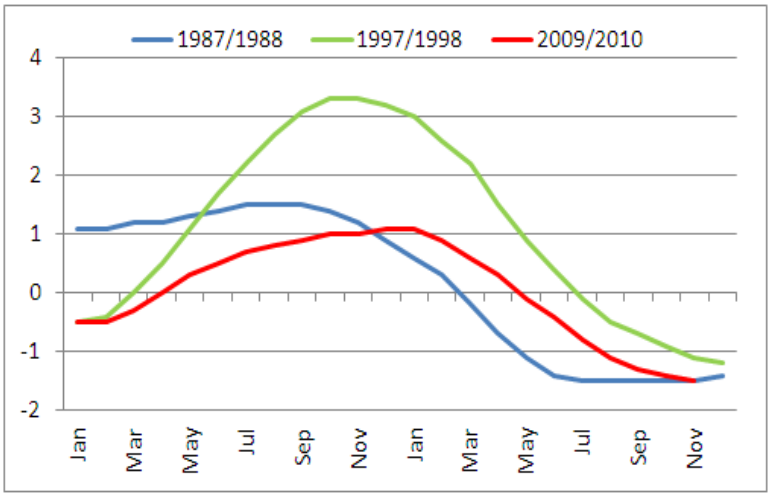


Fig. 4.1 Time series of five-month running mean sea surface temperature deviations (unit: °C) from the climatological mean based on a sliding 30-year period for NINO.3 (5°S – 5°N, 150°W – 90°W).

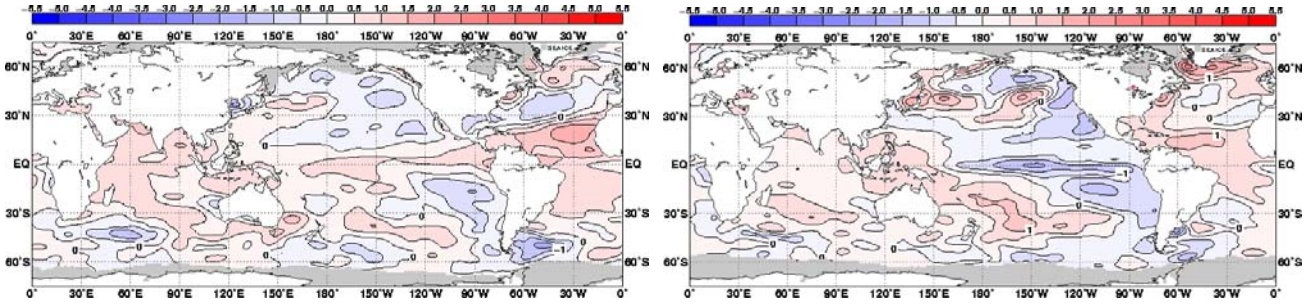


Fig. 4.2 Three-month mean sea surface temperature anomaly (unit: °C) for March – May (left) and June – August (right) 2010.

Atlantic, and were below normal over the central and eastern Pacific (Figure 4.2). Convective activity (inferred from OLR) associated with the Asian summer monsoon was suppressed across a broad area from northern India to the Indochina Peninsula and to the east of the Philippines in the first half of summer 2010, and was enhanced over the Arabian Sea and from the Indochina Peninsula to the north of the Philippines in the second half (Figure 4.3). Over the eastern Indian Ocean and on the Maritime Continent, convective activity was enhanced throughout the summer.

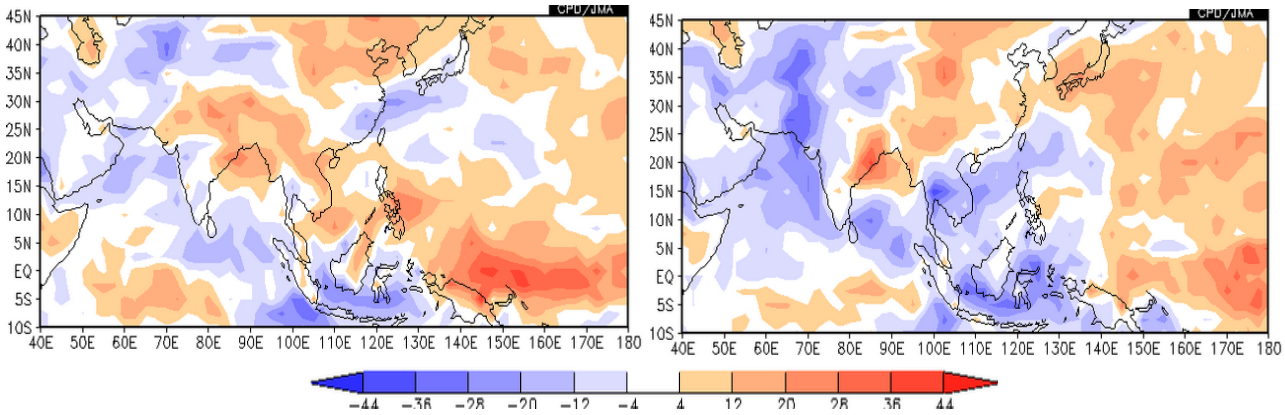


Fig. 4.3 Outgoing longwave radiation (OLR) anomaly (unit: W/m^2) in the first half (left) and the second half (right) of the summer 2010. Cold and warm shading indicates enhanced and suppressed convective activity, respectively, in relation to the normal.

5. Characteristic atmospheric circulation

5.1 Tropospheric air temperature

The zonally averaged tropospheric air temperature in the mid-latitudes of the Northern Hemisphere was the highest for summer (June – August) since 1979 (Figure 5.1). The zonally averaged tropospheric thickness in the tropics reached an above-normal level in July 2009 and matured in the first half of 2010 (Figure 5.2). In June 2010, the thickness in the tropics remained above normal but decreased, while that in the mid-latitudes of the Northern Hemisphere rapidly increased and remained significantly above normal throughout the summer. The evolution of a thickness anomaly similar to that seen in 2010 was identified in 1988 and 1998 (not shown). As research so far (*e.g.*, Angell 2000) has indicated, tropospheric air temperatures increase on a global scale after an El Niño event. Examination of past La Niña events shows a tendency for higher-than-normal tropospheric air temperatures in the mid-latitudes of the Northern Hemisphere. This is consistent with the results of previous research (*e.g.*, Seager *et al.* 2003). It is therefore possible that zonally averaged tropospheric air temperatures in the mid-latitudes of the Northern Hemisphere were extremely high in summer 2010 from the influence of the El Niño event and partly due to the effects of the La Niña event. A warming trend can be identified in the zonally averaged tropospheric air temperature in the mid-latitudes of the Northern Hemisphere (Figure 5.1). This trend may be associated with global warming due to the buildup of anthropogenic greenhouse gases.

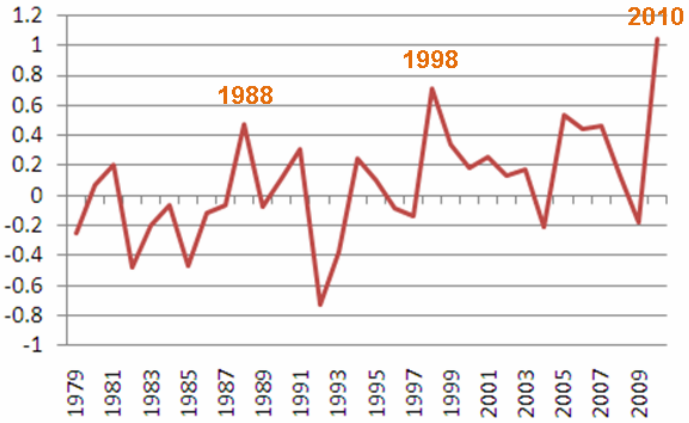


Fig. 5.1 Time series of three-month (June – August) zonally averaged temperature anomalies (unit: K) in the mid-latitudes (30°N – 60°N) of the Northern Hemisphere calculated from thickness (850 – 300 hPa).

5.2 Remarkably strong anticyclone over Japan

The subtropical jet stream in the vicinity of Japan showed a southward-shifted tendency from its normal position in the first half of

summer 2010, while the jet stream was shifted northward of its normal position with a frequent northward meander (a ridge of high pressure) in the second half (Figure 5.3). In line with the characteristics of the subtropical jet stream, the extension of the Tibetan High to Japan was weaker than normal in the first half of summer 2010, while it was stronger than normal (Figure 5.4) and equivalent-barotropic highs developed and persisted over Japan in the second half (Figure 5.5). As detailed in Chapter 4, convective activity was broadly enhanced in and around the Indian Ocean in the second half of summer 2010 (Figure 4.3). Considering that research so far (*e.g.*, Krishnan and Sugi 2001; Enomoto 2004) indicates a link between the Asian summer monsoon and Japan's climate, statistical analysis was implemented to examine the relationship between convective activity linked to the Asian summer monsoon and the subtropical jet stream in the vicinity of Japan. The results indicated that when convective activity is enhanced (or suppressed) over the region from the northern Indian Ocean to the northeast of the Philippines (10°N – 20°N, 60°E – 140°E), the subtropical jet stream near Japan tends to shift northward (or southward) of its normal position (Figure 5.6). It therefore seems that the northward shift of the subtropical jet in the vicinity of Japan was associated with active convections across the broad region over and around the Indian Ocean. In addition, considering previous studies (*e.g.*, Nitta 1987) indicating the influence of convective activity around the Philippines on Japan, active convections from the northern South China Sea to the area northeast of the Philippines may have been partly responsible for the strength of anticyclones around Japan, especially from the second half of August to early September.

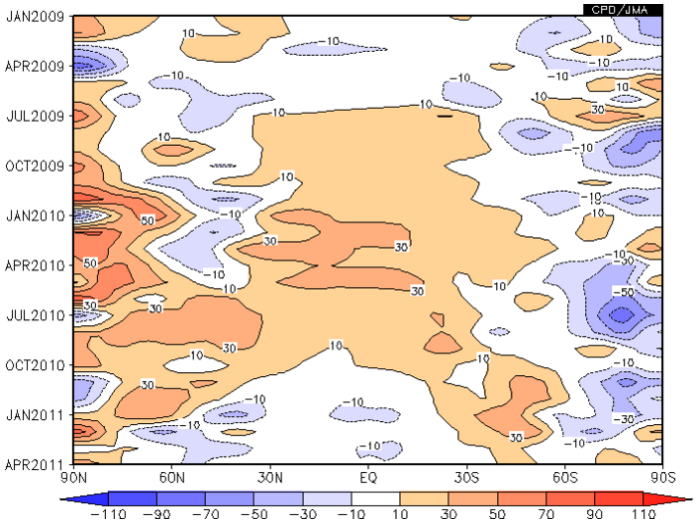


Fig. 5.2 Time-latitude cross section of monthly zonally averaged thickness (300 – 850 hPa) anomaly (unit: m) for January 2009 – March 2011.

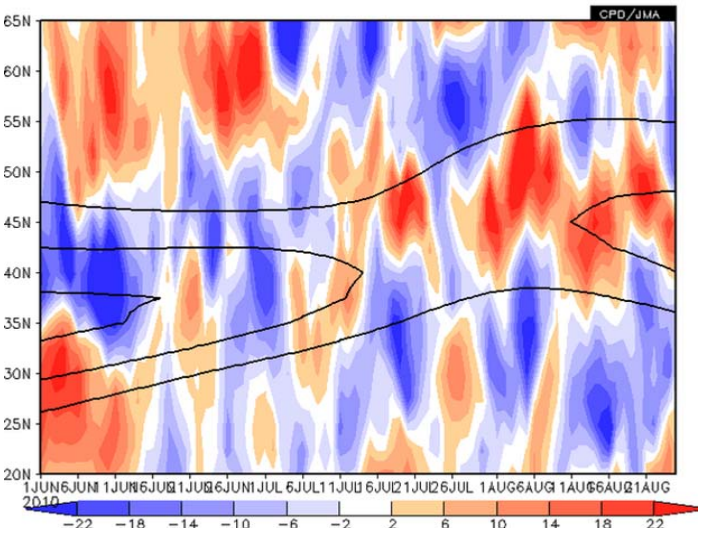


Fig. 5.3 Latitude-time cross section of normal (black lines) and anomalies (colored shading) of the 200-hPa zonal wind speed averaged in the vicinity of Japan (125˚E – 145˚E).

5.3 Okhotsk High

The Okhotsk High (a cool semi-stationary anticyclone) often develops around the Sea of Okhotsk from spring to autumn and occasionally brings cool air to the Pacific side of northern and eastern Japan, resulting in cool summers. In the period from June to the first half of July 2010, the Okhotsk High was less developed than in past years (Figure 5.7). In addition, warm anticyclones frequently covered Japan (particularly its northern parts), and significant anticyclones to the east of the country brought warm air from the south, leading to significantly higher-than-normal temperatures in the first half of summer 2010 (Figure 5.8). A blocking high in the upper troposphere over the Sea of Okhotsk plays an integral role in the formation of the Okhotsk High (Nakamura and Fukamachi 2004). In the first half of July, a blocking high developed near the Sea of Okhotsk, but its position was not suitable for the formation of the Okhotsk High. In the second half of

July, the phenomenon temporarily appeared but influenced Japan little due to the northward shift of the subtropical jet near the country and the strong Pacific High to its east.

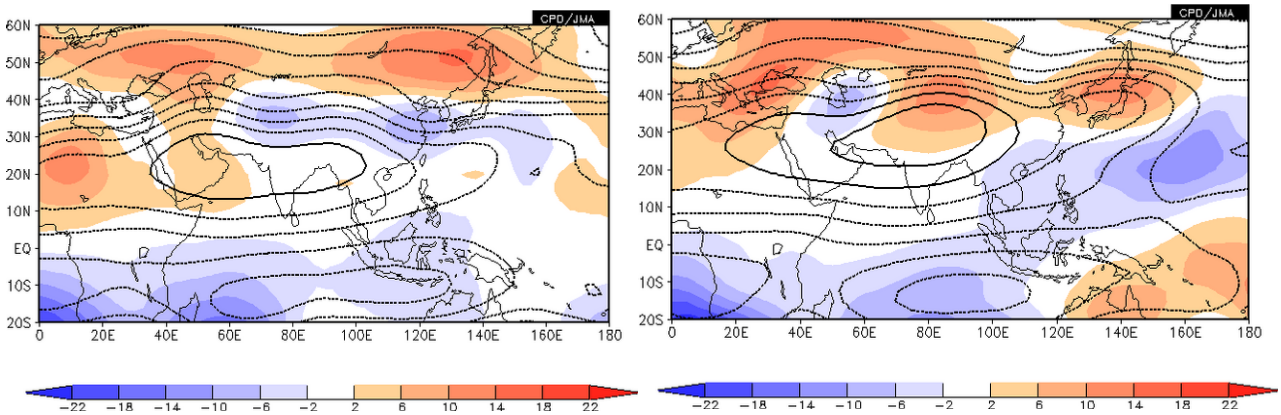


Fig. 5.4 Monthly-mean 200-hPa stream function and anomaly for June (left) and August (right) 2010. The contours show the stream function at intervals of 1×10^7 m^2/s, and the shading indicates stream function anomalies. In the Northern (Southern) Hemisphere, warm (cold) shading denotes anticyclonic (cyclonic) circulation anomalies.

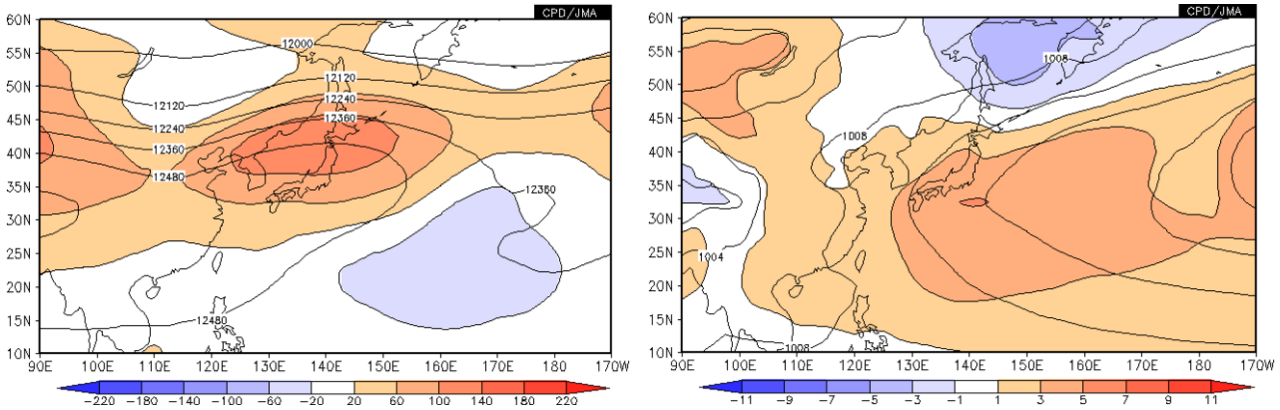


Fig. 5.5 Monthly-mean atmospheric circulation around Japan for August 2010. Left panel: The contours indicate 200-hPa height at intervals of 120 m. Right panel: The contours indicate sea level pressure at intervals of 4 hPa. The shading shows their anomalies.

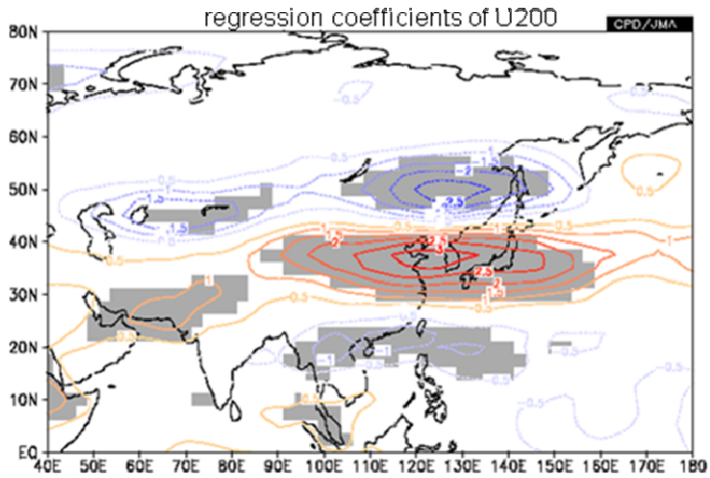


Fig. 5.6 Linear regression coefficients of 200-hPa zonal wind speed onto area-averaged convective activity (OLR) over the region from the northern Indian Ocean to the area northeast of the Philippines (10°N – 20°N, 60°E –140°E) for July and August. When convective activity was enhanced (suppressed) over the region, the subtropical jet stream to the north of Japan was stronger (weaker) than normal at the cold (warm) colored contours. The shading shows a 95% confidence level based on F-testing. The base period for the statistical analysis is 1979 – 2009.

6. Conclusion

The Advisory Panel on Extreme Climate Events summarized the characteristics of the atmospheric circulations leading to the extremely hot summer in Japan and identified the major factors contributing to them as follows:

(1) The zonally averaged tropospheric air temperature in the mid-latitudes of the Northern Hemisphere reached its highest summer level since 1979. This can be attributed to the delayed effect of the El Niño event and partly to the effects of the La Niña event.

(2) Japan was significantly influenced by the pronounced Pacific High caused by enhanced convection in the broad area over the Indian Ocean and the surrounding seas. It is possible that the broadly active convection was associated with above-normal SSTs over the Indian Ocean and with the La Niña event.

(3) Japan was influenced less than usual by the Okhotsk High (a cool semi-stationary anticyclone) because the occurrence frequency of the Okhotsk High itself was lower than normal. In addition, a predominant anticyclone to the east of Japan intercepted the influence of the Okhotsk High.

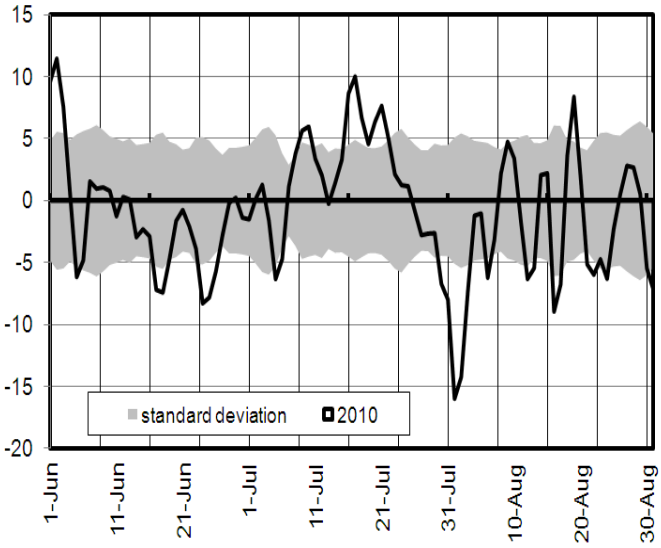


Fig. 5.7 Time series of average sea level pressure anomaly (unit: hPa) over the Sea of Okhotsk (45°N – 60°N, 140°E – 155°E) from June to August 2010. The thick line indicates daily mean values of the area-averaged sea level pressure anomaly over the region. The gray shading denotes the range of one standard deviation.

These atmospheric characteristics and the factors that may have contributed to them are illustrated in Figure 6.1 The primary factors outlined above are supported by statistical analysis and research performed to date, but the results of this work do not wholly explain the extreme conditions seen. In order to further understand the event and clarify the dynamic mechanism behind it, it is necessary to investigate other possible

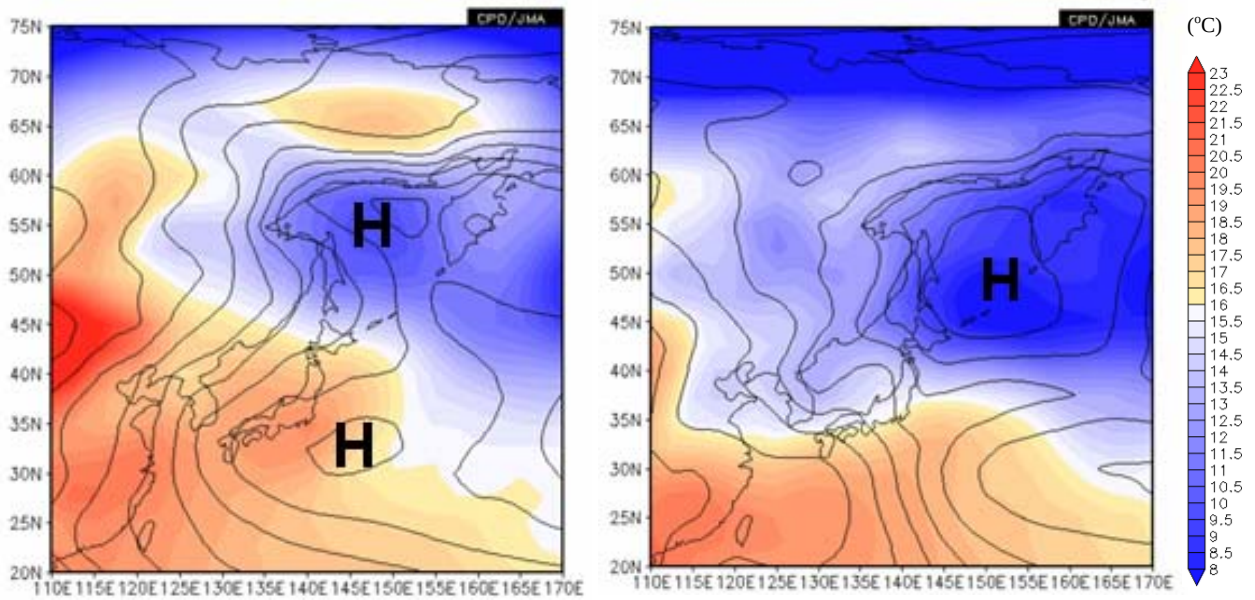


Fig. 5.8 The cool semi-stationary anti-cyclone (Okhotsk high). The contours show sea level pressure, and the shading indicates temperature at 850 hPa. The left panel and right panel are for 15 -24 July 2010 and for 19 - 28 July 2093, respectively.

factors and perform numerical model experiments.

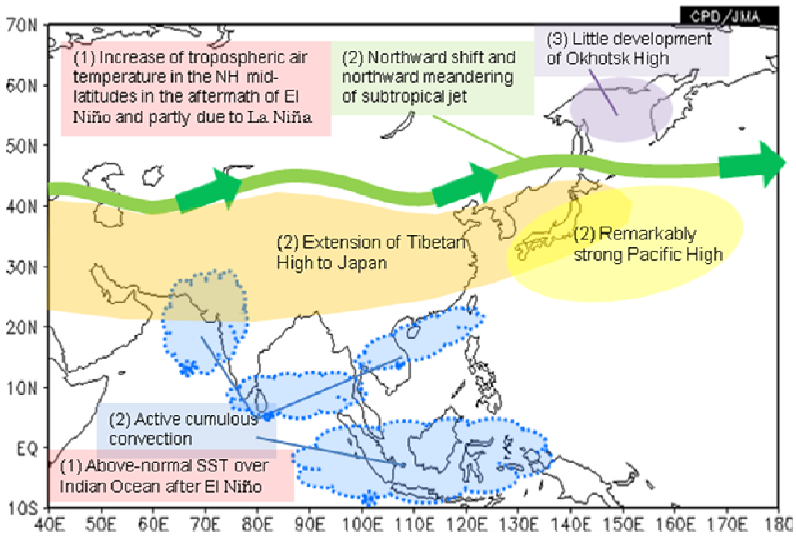


Fig. 6.1 Primary factors contributing to Japan's extremely hot summer (June – August) of 2010. Here, (1), (2) and (3) correspond to the numbers in Chapter 6 of the main text.

References

Angell, J. K., 2000: Tropospheric temperature variations adjusted for the El Niño influence, 1958 – 1998. *J.Geophys. Res.*, **105**, 11841 – 11849.

Enomoto, T., 2004: Interannual variability of the Bonin high associated with the propagation of Rossby waves along the Asian jet. *J. Meteor. Soc. Japan*, **82**, 1019 – 1034.

Krishnan, R. and M. Sugi, 2001: Baiu rainfall variability and associated monsoon teleconnections. *J. Meteor. Soc. Japan*, **79**, 851 – 860.

Ishii, M., A. Shouji, S. Sugimoto and T. Matsumoto, 2005: Objective Analyses of Sea-Surface Temperature and Marine Meteorological Variables for the 20th Century using ICOADS and the Kobe Collection. *Int. J. Climatol.*, **25**, 865-879.

Nakamura, H. and T. Fukamachi, 2004: Evolution and dynamics of summertime blocking over the Far East and the associated surface Okhotsk high. *Quart. J. Roy. Meteor. Soc.*, **130**, 1213 – 1233.

Nitta, T., 1987: Convective activities in the tropical western Pacific and their impact on the Northern Hemisphere summer circulation. *J. Meteor. Soc. Japan*, **65**, 373 –390.

Onogi, K., J. Tsutsui, H. Koide, M. Sakamoto, S. Kobayashi, H. Hatsushika, T. Matsumoto, N. Yamazaki, H. Kamahori, K. Takahashi, S. Kadokura, K. Wada, K. Kato, R. Oyama, T. Ose, N. Mannoji and R. Taira, 2007: The JRA-25 Reanalysis. *J. Meteorol. Soc. Japan*, **85**, 369 - 432.

Seager, R., N. Harnik, Y. Kushnir, W. Robinson and J. Miller, 2003: Mechanisms of hemispherically symmetric climate variability. *J. Climate*, **16**, 2960 – 2978.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Extreme Precipitation and Its Long-Term Changes over China and USA

Fang Wang[1, 2], Song Yang[2], and R. Wayne Higgins[2]

[1]*National Climate Center, China Meteorological Administration, Beijing, China*

[2]*Climate Prediction Center, NOAA/NWS/NCEP, Camp Springs, MD, USA*

1. Introduction

Accompanied by global warming, both precipitation and precipitation extremes in the world has changed greatly in historical records. Significantly increased precipitation has been observed in eastern parts of North and South America, northern Europe and northern and central Asia. Meanwhile, the frequency of heavy precipitation events has increased over most land areas, even in those where there has been a reduction in total precipitation amount, consistent with warming and observed increases of atmospheric water vapor (IPCC, 2007).

China and USA are located across the Pacific Ocean, and are characterized by different climate features due to their differences in land-sea thermal contrast and ocean-atmosphere interaction between the western and eastern Pacific. Many previous studies have demonstrated the link of climate signals across the North Pacific, especially between Asia and North America (Lau *et al.* 2002; Wang *et al.* 2001; Zhang *et al.* 2005; Ding *et al.* 2005; Li *et al.* 2005; Zhao *et al.* 2011). This study will focus on the long-term changing features of extreme precipitation and their links to SST/Ts and atmospheric patterns from annual and seasonal timescale, mainly in the country-wide regard for both China and USA.

2.Data and methodology

2.1 Data

The following data sets are used in this study: (1) Station-based daily precipitation over China from 1961 to 2009, (2) USA daily precipitation from US Daily Precipitation Gridded Analysis from 1948 to 2009, (3) NOAA extended reconstructed SSTs V3b datasets from 1854 to the present (Smith *et al.*, 2008), (4) Surface air temperature from the station observation-based global land monthly mean surface air temperature dataset from 1948 to the present (Fan *et al.*, 2008), (5) Monthly wind and vertical velocity data from NCEP/NCAR Reanalysis1 (Kalnay *et al*., 1996).

2.2 Methodology

Three precipitation indices are used to portrait the characteristics of the total and extreme precipitation. They are defined as follows:

P_{95} - Extreme precipitation based on 95% percentile. Let R_{wj} be the daily precipitation amount on a wet day w ($R \geq 1.0$mm) in period j (j can be year or season, the same below) and let $R_{wn}95$ be the 95th percentile of precipitation on wet days in the 1971-2000 period. If W represents the number of wet days in the period, then:

$$P_{95\,j} = \sum_{w=1}^{W} R_{wj}, \; where \; R_{wj} > R_{wn}95$$

P_{TOT} - Annual or seasonal total precipitation in wet days. Let R_{ij} be the daily precipitation amount on day i in period j. If I represents the number of days in j, then

$$P_{TOT\,j} = \sum_{i=1}^{I} R_{ij}$$

Correspondence to: Fang Wang, National Climate Center, China Meteorological Administration, Beijing, 100081, P.R.China; E-mail: fangwang@cma.gov.cn.

R_{95p} - Ratio of extreme precipitation to total precipitation. It measures the relative importance of extreme precipitation.

$$R_{95p} = P_{95} / P_{TOT} * 100$$

In this study, we will focus on R_{95p} and P_{TOT}, mainly due to the high correlation between P_{TOT} and P_{95}.

The angular distance weighting (ADW) algorithm is used to interpolate station indices to regular grids over China (New *et al*. 2000). The long-term trends of the indices are estimated by the non-parametric Mann-Kendall test and Sen's method (Kendall 1955; Sen 1968).

3. Results

3.1 Annual features

The annual variations and long-term trends of total precipitation and the ratio are given in Fig. 1. Obvious differences can be seen between China and USA from the country-wide perspective. Over China, total precipitation nearly has no trend, and the ratio shows a slight positive trend of about 0.22% per decade, but not statistically significant. In contrast, over USA, total precipitation increases significantly, with a trend about 13.2mm per decade. Meanwhile, the ratio shows a moderate decrease by -0.15% per decade.

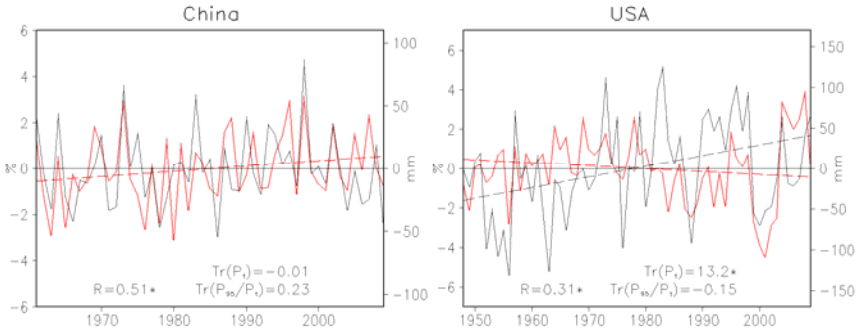


Fig. 1 Annual variations of total precipitation (black line) and the ratio (red line) for China (left) and USA (right). Dashed lines represent the linear trend.

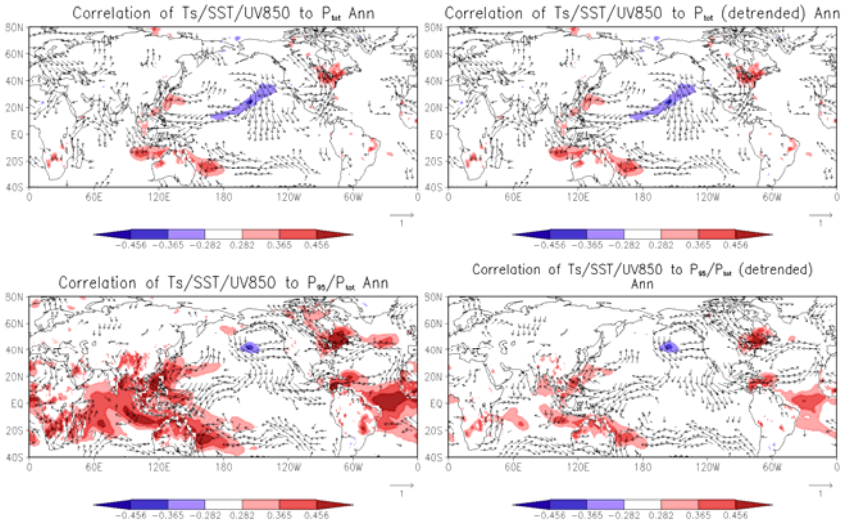


Fig. 2 Correlation patterns of total precipitation (upper) and the ratio (bottom) over China to SST/Ts and V850 wind. Left plots: original data; right plots: detrended data.

Fig. 2 gives the correlation patterns between precipitation and SST/Ts and 850hPa wind. We compare these patterns between original data and detrended data to distinguish the long-term features from the features on interannual timescales. For China, only few significant features can be seen for total precipitation. However, the ratio has strong relations to the SSTs in the Indian Ocean to western Pacific, the South and East China Sea, ocean east of Australia and the tropical Atlantic Ocean. Corresponding to SST pattern, a strong easterly from central Pacific passes through Philippine Sea and South China Sea, then turns to the northeast, forming a significant anti-cyclonic pattern in the western Pacific. When we removed the trend of the ratio, these SST relationships become much weaker especially in Indian Ocean, indicating an apparent link of SST and the ratio for their long-term changes. That is, the positive trend of China ratio is related to the long-term warming in these oceans.

As for USA, total precipitation has more significant features related to SST/Ts and circulation than the ratio (Fig. 3). It can be seen that total precipitation has significant relationships with the SSTs in the Indian Ocean and the eastern Pacific. The total precipitation also has a negative correlation with the North Pacific

SST. Especially, the correlation pattern in Pacific may be related to the long-term change of Pacific decadal Oscillation (PDO), which has experienced an obvious shift from cool phase to warm phase about in 1976. In warm phase, strong warm SST anomalies occurred in central and eastern Pacific and distinct cooling is located in North Pacific around 40°N, corresponding to a relatively-abundant rainy era of USA although strong inter-annual variation exists. Meanwhile, the trade wind in central and eastern Pacific weakens strongly, leading to higher SST in tropical eastern Pacific and near seas to the west of North America, and subsequently abundant water vapor transferring from Pacific to east of USA, passing through Central America and Gulf of Mexico, favoring USA precipitation. It also can be seen that USA total precipitation has a high correlation to SST in Indian Ocean. After the trend is removed, these SST relationships become apparently weaker in Pacific, and entirely disappear in Indian Ocean, again suggesting that the relationships are for the long-term changes of SST and precipitation. Namely, the increasing trend of PDO index in recent decades may play important role in the positive trend in total precipitation over USA. Meanwhile, Indian Ocean warming may also play an important role in its long-term trend.

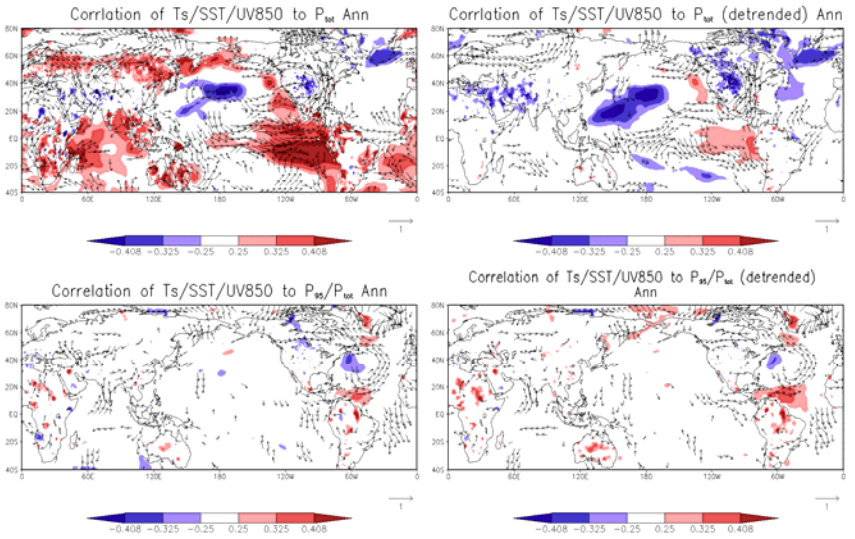


Fig. 3 Same as Fig. 2, but for USA.

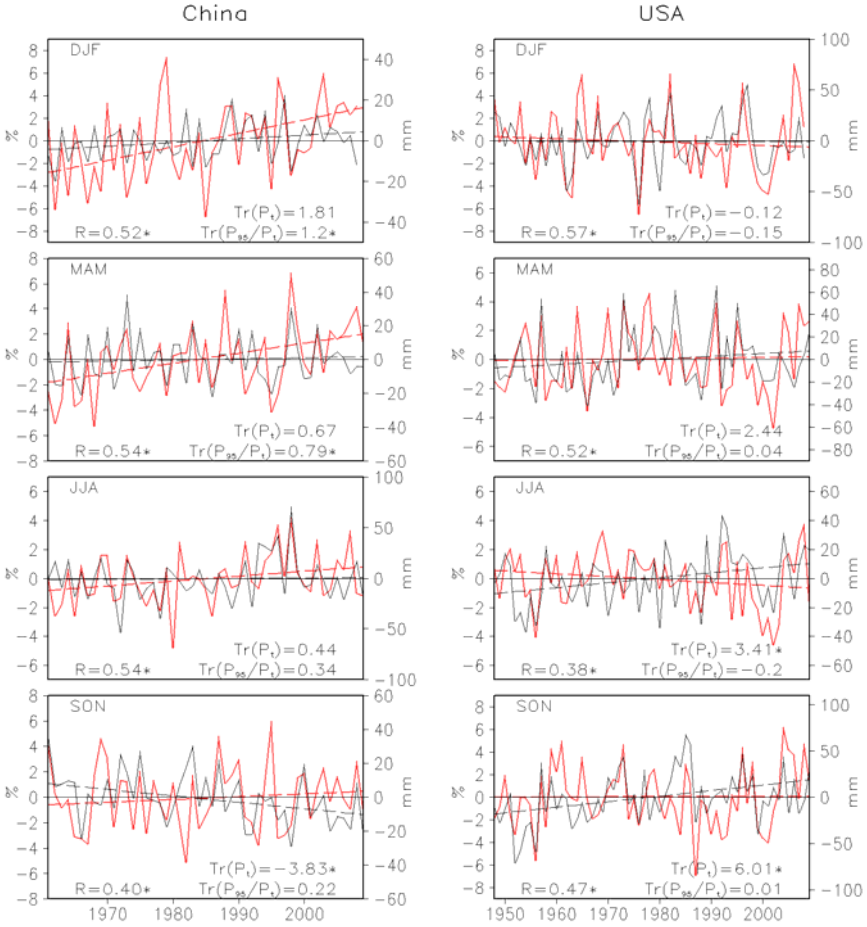


Fig. 4 Same as Fig. 1, but for different seasons.

3.2 Seasonal features

Precipitations show distinct features in different seasons especially over China. In this section, the seasonal features of precipitation indices are addressed as well as their links to oceanic and atmospheric conditions. The methods used here are the same as the annual mean.

Fig. 4 provides the variations of seasonal-mean ratio and total precipitation, averaged for China and USA. Over China, total precipitation changes insignificantly in winter, spring and summer, only slight positive trends being detected. While in fall, it shows significant negative trend. However, the ratio shows consistent positive trend in all seasons, especially significant in winter and spring. In contrast, over USA, total precipitation shows positive trend in seasons other than winter, especially significant in summer and fall. However, the ratio has little long-term change except the moderate negative trend in winter and summer.

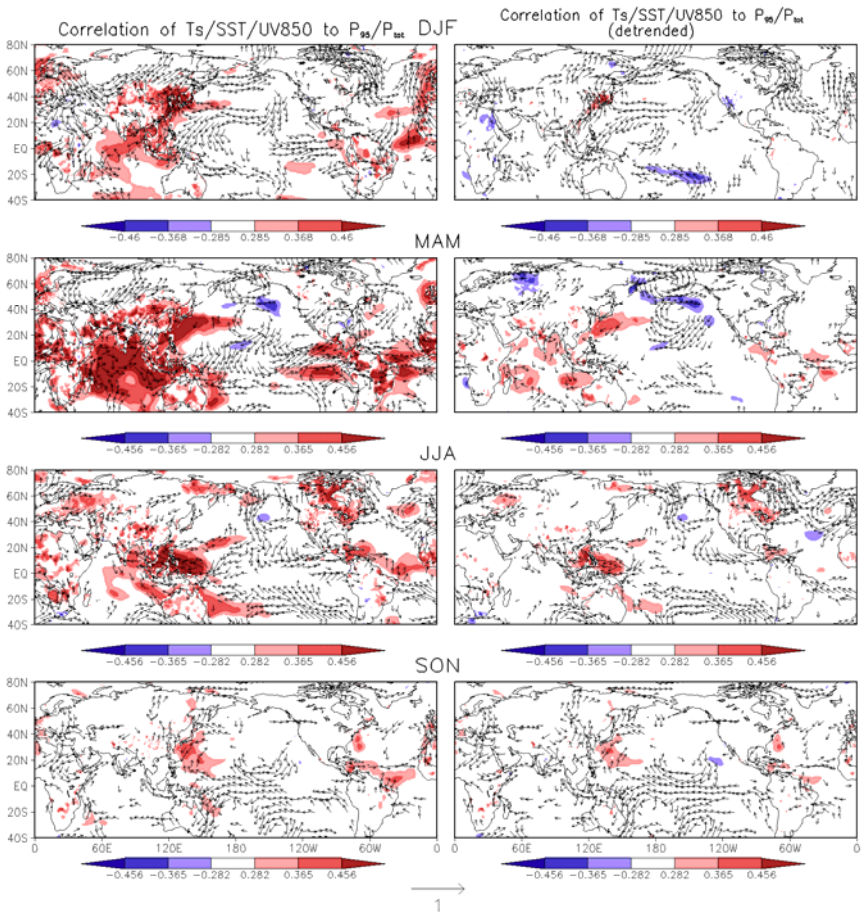


Fig. 5 Correlation patterns of the ratio over China to SST/Ts and V850 winds. Left plots: original data; right plots: detrended data.

Over China, only few SST and atmospheric circulation features are linked to total precipitation (Figure nor shown), so we focus on the ratio. The ratio over China has consistent long-term positive trend in all seasons, particularly significant in winter and spring. Here we extract the SST and circulation features associated with the long-term changes of the ratio by comparing the original and detrended correlations (Fig. 5). In winter, the ratio is significantly related to SSTs in Indian Ocean to Maritime Continent, South China Sea and seas to the east of China, concurrent with an intensification of northeast wind and an anomalous cyclonic pattern in north Pacific. In summer, the relationships enhance greatly as compared to spring. Furthermore, the ratio also has a significant relation to SSTs in the eastern tropical Pacific and tropical Atlantic Ocean. In fall, the high correlation is mainly located in South China Sea and Philippine Sea, which are key areas to East Asia Summer Monsoon (EASM). After the ratio is detrended, its relationships to SST become much weaker in most areas, indicating the apparent link of the long-term changes between the ratio and SST in these areas. In particular, tropical Oceans warming especially in Indian Ocean, eastern Pacific and Atlantic Ocean, and their corresponding atmospheric circulations, may play an important role in the long-term change in the ratio over China, particularly for spring.

Over USA, total precipitation increases obviously except in winter, especially significant in summer and fall, However the ratio has little trend. Therefore, we mainly focus on the SST and circulation features related to total precipitation (Fig. 6). It can be seen that total precipitation over USA is greatly affected by the SST and atmospheric circulation anomalies across North Pacific to North Atlantic Ocean and these features are mainly on interannual timescale. For all seasons a cyclonic pattern can be seen located in USA continent, although the center varies with the season, which is beneficial to precipitation over USA. By comparing the original and detrended features, we find SSTs in tropical eastern Pacific has significant positive correlation with total precipitation over USA, accompanied by strong weakening of trade wind in central and eastern Pacific, which is apparently due to their long-term changes. In addition, SST warming trend in Indian Ocean and tropical Atlantic Ocean also has significant teleconnection with USA summer precipitation. As for the

ratio, few features can be seen linked to SST and circulation both on interannual and long-term timescales (figure not shown). Similar with annual mean, it can be inferred that the seasonal long-term change of PDO exerts an important effect on the precipitation trends, especially in summer and fall. Meanwhile, the long-term warming trend in Indian Ocean SST may also play an important role in change in USA precipitation in fall.

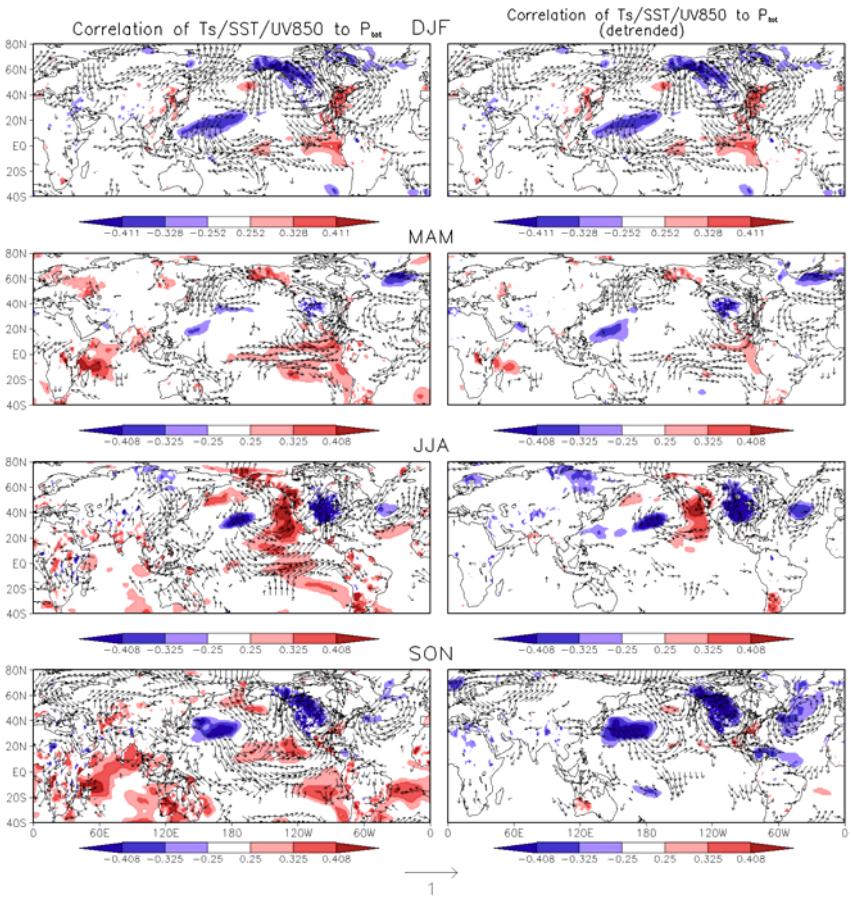


Fig. 6 Same as Fig. 5, but for total precipitation over USA.

4. Summary

Using station-based daily precipitation over China and high-resolution daily precipitation analysis over USA, we analyzed the temporal and spatial patterns of total precipitation and extreme precipitation with the focus on long-term changes.

Features of both total precipitation amount and the extreme precipitation ratio are apparently different between China and the USA, and differences occur not only in the annual means but also in seasonally-averaged values. Annually, the total precipitation over China changes insignificantly and the ratio shows only a slight positive trend. However, the annual total precipitation over USA increases significantly although the ratio decreases moderately. In China, the ratio exhibits positive trends in all seasons, especially in winter and spring, and the total precipitation shows small positive trends except a significant negative trend in fall. In the USA, the total precipitation increases remarkably in all seasons, except a slight decrease in winter, and the ratio decreases in winter and summer but increases in spring and fall.

The change in extreme precipitation ratio over China has a strong link to the SSTs in the Indian Ocean to western Pacific, the South and East China Sea, ocean east of Australia and the tropical Atlantic Ocean. The change in total precipitation over USA is associated with the change in SSTs over both the Indian Ocean and the eastern Pacific Ocean. These relationships become much weaker when the trend of total precipitation or ratio is removed, indicating an apparent impact of SST on the long-term change in precipitation. The trends of precipitation are also linked to the long-term changes in atmospheric circulation including the trade wind, the North Pacific anticyclone, the Asian cyclone, and others.

References

Ding, Q.H., and B. Wang, 2005: Circumglobal teleconnection in the Northern Hemisphere summer. *J. Climate*, **18**, 3484–3505.

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.

IPCC, 2007: Climate change 2007: The physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. Available from http://www.ipcc.ch .

Kalnay, E., and Coauthors, 1996: The NMC/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437-471.

Kendall, M.R., 1955: Rank correlation methods. *Charles Griffin and Company*, London.

Lau, K.M., and H.Y. Weng, 2002: Recurrent teleconnection patterns linking summertime precipitation variability over east Asia and North America. *J. Meteor. Soc. Japan*, **80**, 1309–1324.

Li, Q.Q., S. Yang, V.E. Kousky, R.W. Higgins, K.M. Lau, and P. Xie, 2005: Features of cross-Pacific climate shown in the variability of China and U.S. precipitation. *Int. J. Climatol.*, **25**, 1675–1696.

New, M., M. Hulme, and P.D. Jones, 2000: Representing twentieth century space-time climate variability. Part 2: development of 1901-96 monthly grids of terrestrial surface climate. *J. Climate*, **13**, 2217-2238.

Sen, P.K., 1968: Estimates of regression coefficient based on Kendall's tau. *J. Am. Stat. Assoc*., **63**, 1379–1389.

Smith, T.M., R.W. Reynolds, T.C. Peterson, and J. Lawrimore: 2008: Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880-2006). *J. Climate*, **21**, 2283-2296.

Wang., B, R.G. Wu, K.M. Lau, 2001: Interannual variability of the Asian summer monsoon: contrasts between the Indian and the Western North Pacific–East Asian monsoons. *J. Climate*, **14**, 4073–4090.

Zhang, P.Q., S. Yang, V.E. Kousky, 2005: South Asian high and Asian–Pacific–American climate teleconnection. *Adv. Atmos. Sci.*, **22**, 915–923.

Zhao, P., S. Yang, H.J. Wang, Q. Zhang, 2011: Interdecadal Relationships between the Asian–Pacific Oscillation and Summer Climate Anomalies over Asia, North Pacific, and North America during a Recent 100 Years. *J. Climate*, **24**, 4793–4799.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Changing Monsoon Extremes and Dynamics: Example in Pakistan

S.-Y. (Simon) Wang[1, 2], Robert Davies[1], Robert Gillies[1, 2], and Jiming Jin[2]

[1]*Utah Climate Center, Utah State University, Logan, UT*

[2]*Department of Plants, Soils, and Climate, Utah State University, Logan, UT*

1. Introduction

The Pakistan floods of 2010 constituted an extreme hydrologic event with substantial sociological consequences. At their height they submerged roughly 20% of Pakistan's land area, killing and injuring nearly 5,000 people directly, and displacing as many as 20 million. Though influenced by additional factors (*e.g.*, snowmelt and water management practices), the floods were closely tied to an unusually heavy rainfall event. The countrywide precipitation total during July–September 2010 was the highest since 1994 and the sixth highest in the last 50 years (PMD 2010). The weather event that led to flash floods in 2010 began with a series of local, intense rainstorms in July, followed immediately by monsoon rain throughout the first half of August. Synoptic analysis by Houze *et al.* (2011) pointed out that both the intensity and structure of the July rainstorms were abnormal for northern Pakistan. Due in part to myriad other extreme events in the summer of 2010, such as the Russian heat wave, questions have arisen as to whether the unusual 2010 Pakistan rainfall was simply a manifestation of climate variability or was connected in some measure with global climate change (Marshall 2010). The role of climate change in causing heavy monsoon rains needs to be addressed.

Here we summarize a recent paper by Wang *et al.* (2011) that examined the 2010 monsoon conditions, climatology, and trends. We further investigate the extent to which the trends are associated with internal variability vs. external forcing of the climate system. Additional results from a dynamical downscaling approach (*i.e.*, nesting regional climate models into global reanalyses) are also presented.

2. Data sources

We used the Asian Precipitation-Highly Resolved Observational Data Integration Towards Evaluation (APHRODITE; Xie *et al.* 2007), a gridded daily precipitation dataset derived from rain gauges with a 0.5° resolution for the period 1971–2007. The APHRODITE dataset allowed for an investigation of long-term daily precipitation over Pakistan, where stations with adequate records are sparse. Other data included the 3-hr Climate Prediction Center (CPC) Morphing technique precipitation (CMORPH; Joyce *et al.* 2004) for the analysis of the 2010 event, and the Tropical Rainfall Measuring Mission (TRMM)-based Lightning Imaging Sensor (LIS) 7-year climatology. For meteorological variables we used the ERA40 reanalyses (1971–2002) and the ERA-Interim reanalyses (1989–present; Uppala *et al.* 2005, 2008). These two reanalyses assimilated 2-meter air temperatures (T_{2m}). For data uniformity, we reduced the resolution of the ERA-Interim to be consistent with the ERA40 using bilinear interpolation. Finally, we performed downscaling simulations with the Weather Research and Forecasting (WRF) model, Advanced WRF (ARW) version 3.1, using the the National Centers for Environmental Prediction Reanalysis (NCEP2; Kanamitsu *et al.* 2002) as boundary conditions.

3. Results

a. Case analysis

The movement of the monsoon rainband and its association with the 2010 convective storms are displayed in Fig. 1a as the latitude-time diagram of CMORPH precipitation, and in Fig. 1b as the ERA-Interim's 850 mb relative vorticity and column-integrated precipitable water – all averaged between 70°E-75°E, which encompasses northern Pakistan. Three northward-moving monsoon troughs are readily visible in these meteorological fields. These features are typical of the 30–60 day mode of the Indian monsoon.

Correspondence to: Shih-Yu (Simon) Wang, Utah Climate Center & Department of Plants, Soils, and Climate, Utah State University, Logan, UT; E-mail: simon.wang@usu.edu.

However, only the second reached northern Pakistan (beyond 30°N), elevating moisture and producing near-continuous rainfall there during the first half of August. Arrival of the second monsoon trough also increased the convective available potential energy (CAPE; Fig. 1c contours) in northern Pakistan. In contrast, the three mesoscale convective systems (MCSs) responsible for the flash floods occurred during a drier and warmer period in July (arrow indicated), prior to the second monsoon trough's arrival; these features are discernible in the precipitable water and T_{2m} (Figs. 1b and 1c). Moisture surges accompany each MCS (Fig. 1b), apparently associated with passing synoptic disturbances as noted in Houze *et al.* (2011); these moisture surges are also evident in CAPE. Moreover, each of the MCSs occurred during the period when a series of midlatitude troughs intruded south, as is revealed in the 250 mb vorticity (not shown). Such a coupling between upper-level troughs and monsoon depressions exemplifies the tropical-midlatitude interaction that is known to take place over this region (*e.g.*, Chen *et al.* 2005). Such a feature, together with warm temperatures over sloping terrain and strong moisture gradients, suggests a thermodynamically unstable atmosphere prior to the monsoon trough's arrival in August.

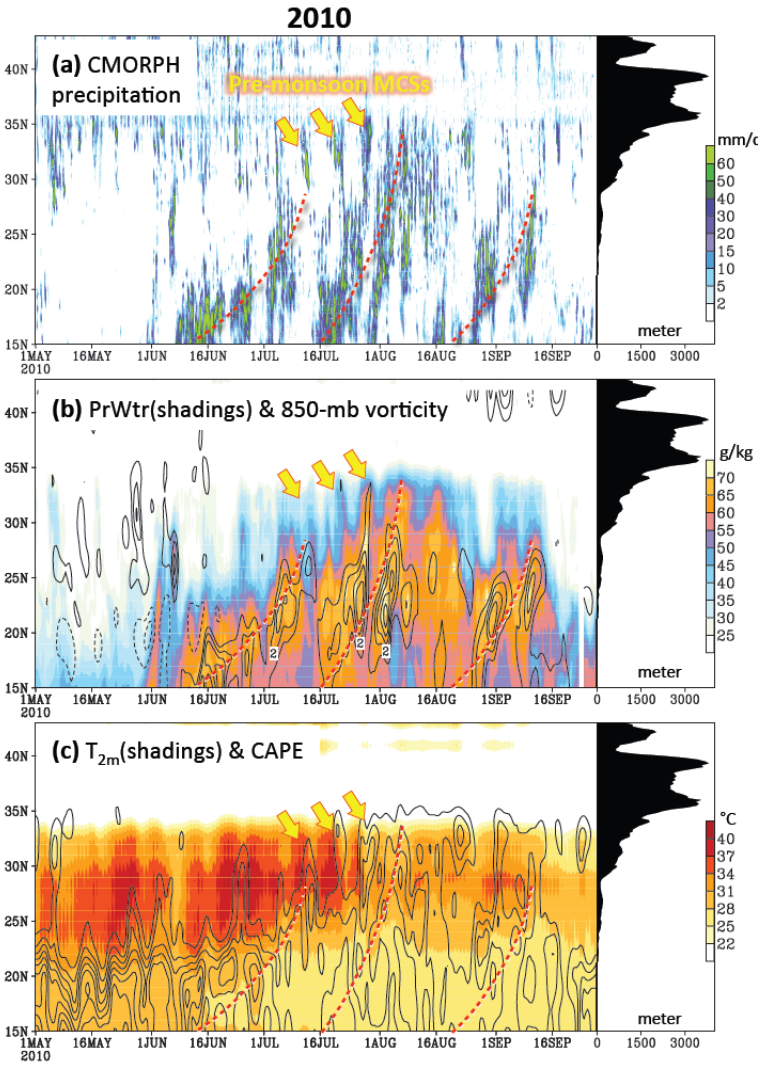


Fig. 1 Latitude-time diagrams for (a) CMORPH 3-hr precipitation, (b) precipitable water (shadings) and 850mb relative vorticity (contour interval 10^{-5} s^{-1} omitting zero and -10^{-5}), and (c) T_{2m} (shadings) and CAPE (contour interval 300 J Kg^{-1} beginning 300) averaged between 70°E and 75°E from May 1 to September 30, 2010. The three monsoon troughs are indicated by dashed lines. The three MCSs in July are indicated by yellow arrows. Mean elevation across the analysis area is shown to the right of each panel.

b. Climatology

The analysis in Fig. 1 suggests that the common description of the Pakistan monsoon as a single entity (*i.e.*, July as the monsoon onset and August as the mature monsoon) may not be adequate. Instead, these summer rains comprise two distinct phases of development: (i) a pre-monsoon trough phase with drier, warmer air and episodic strong convection; and (ii) a monsoon trough phase influenced by northward migrating monsoon troughs and associated humid, cooler air. Examination of climatological T_{2m}, dew point temperature (T_d), and total precipitation in northern Pakistan from 1971 to 2010 (Fig. 2) shows that the warmest T_{2m} occurs in June, with a rapid increase in T_d in response to the intrusion of moist air from the south. Then in July, T_{2m} begins to decrease; T_d reaches the seasonal maximum; and precipitation intensity peaks and is characterized by episodic features. In August, T_{2m} cools gradually while T_d peaks and then decreases in association with declining rainfall. These climatological features are in good agreement with the 2010 analysis in Fig. 1c, in which the July pre-monsoon trough phase comprises stronger instability and more intense convective systems. This can also be inferred from the increased lightning frequency accompanying

the high values of both T_{2m} and T_d in Fig. 2—strong lightning activity being common prior to the monsoon onset (*e.g.*, Kodama *et al.* 2005). By contrast, the August monsoon trough phase is characterized by less episodic rainfall of decreasing intensity and declining lightning activity.

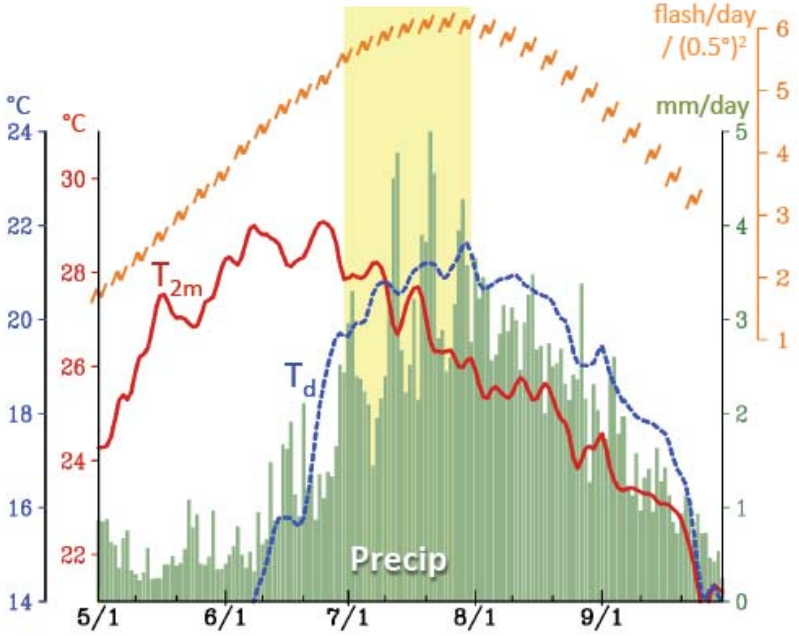


Fig. 2 Climatological means (1971-2009) of T_{2m} (red), Td (blue), precipitation (green bars), and lightning frequency (orange lightning symbols; plotted every 5 days) over northern Pakistan.

Such differences in the monsoon phases are also evident in the 10-day mean patterns of the 250 mb streamlines and precipitation in 2010 (Fig. 3). During July 10–20 (*i.e.*, the rainstorm episode), northern Pakistan is situated north of the monsoonal anticyclone, under a deepened synoptic trough to the immediate west. This feature echoes the upper-level "short waves" influence on rainfall (Khan 1993). Low-level circulations (850 mb) show that southwesterly flows prevail deeply inland over northern Pakistan, reflecting the moisture source of the MCSs. During August 1–11 (*i.e.*, the beginning of the monsoon trough episode), the monsoonal anticyclone matures and migrates north, completely covering northern Pakistan, while at 850 mb an enhanced monsoon trough migrates toward northern Pakistan, as outlined by the red dotted line in Fig. 3d.

c. Trends

The identification of two distinct phases in the northern Pakistan monsoon, attributable to separate environmental conditions, suggests separate analysis for each phase. Mean precipitation of both the July pre-monsoon trough phase (Fig. 4a) and the August monsoon trough phase (Fig. 4b), shows virtually no trend over the past 40 years (orange bars and line). This is consistent with the fact that the summer rainfall in 2010, albeit extreme, ranks only the sixth highest in observational history (PMD 2010). However, the convective nature of the pre-monsoon trough phase suggests the possibility that the unstable environment may be enhanced owing to post-1970 global warming—warming also evident in northern Pakistan (Figs. 4a and 4b). We analyzed the frequency of intense precipitation by computing the occurrences of grid-scale daily precipitation exceeding 5 mm (*i.e.*, one standard deviation of the monthly mean) and then averaging the occurrences in northern Pakistan. The frequency reveals a substantial upward trend in July (Fig. 4a; blue bars and line) but shows only a weak trend in August (Fig. 4b). A significance test of the trends is given in the caption, while trends excluding 2010 are overlaid as a cyan dashed line (apparently 2010 did not alter the trend). Moreover, the contrast between trends in the mean precipitation vs. the frequency of intense precipitation signifies an overall increase in rainfall intensity, which is linked to enhanced instability. This feature is particularly pronounced in the pre-monsoon trough phase of July.

d. Role of climate change

- *Diagnostics:* Since mean precipitation amounts in northern Pakistan reveal no trends, we focus the following analysis on the observed increase in convective activity. Atmospheric instability, especially during the pre-monsoon trough phase, is examined by the lapse rate of equivalent potential temperature (θ_e). The lapse rate, denoted as $d\theta_e/dp$, is shown in Fig. 5a with the climatological mean removed (shadings). A marked increase in $d\theta_e/dp$ is observed in the lower troposphere associated with a corresponding decrease in the upper troposphere. This suggests an overall increase in conditional instability, as well as potential instability for convective storms. In other words, additional moisture as seen in the increased T_d in Fig. 4a, combined with surface warming over sloping terrain, may have

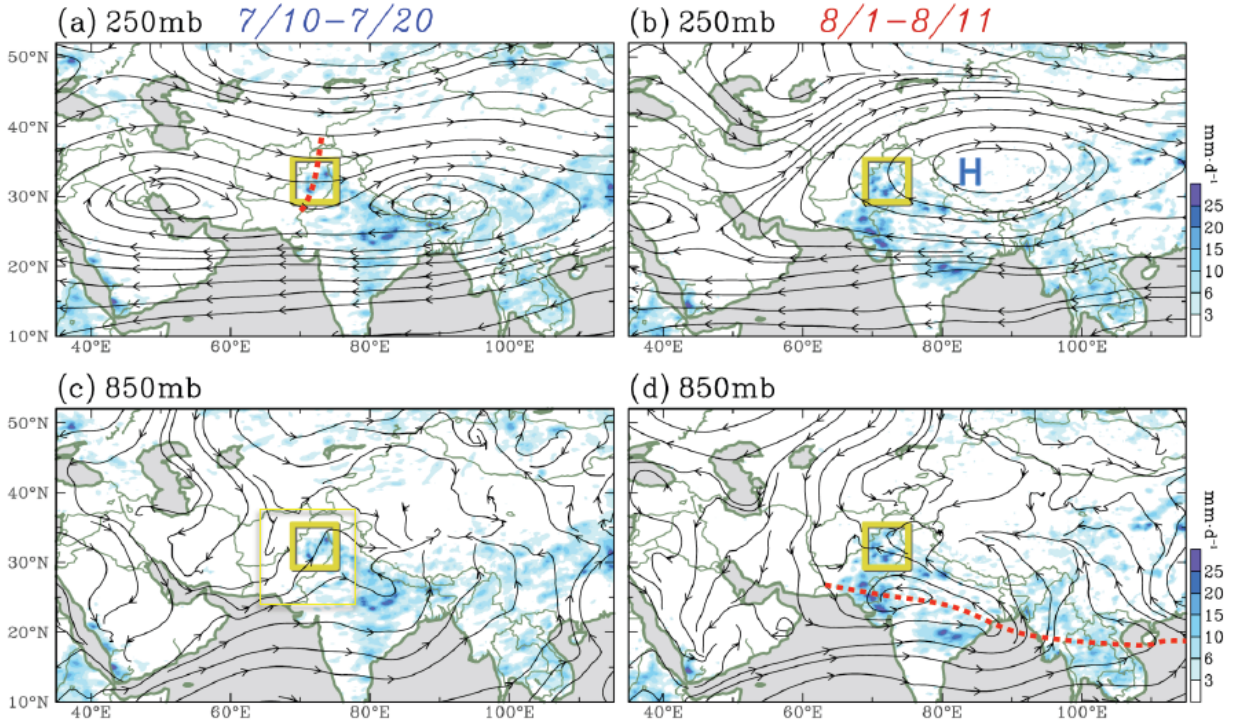


Fig. 3 Streamlines at 250 mb and the CMORPH precipitation (shadings) averaged in (a) 7/10-20 and (b) 8/1-11, 2010. The corresponding 850-mb streamlines are given in (c) and (d). Northern Pakistan is outlined by the yellow box. Red dashed lines in (a) and (d) indicate the synoptic-scale trough and the monsoon trough, respectively. The letter "H" in (b) indicates the monsoon anticyclone.

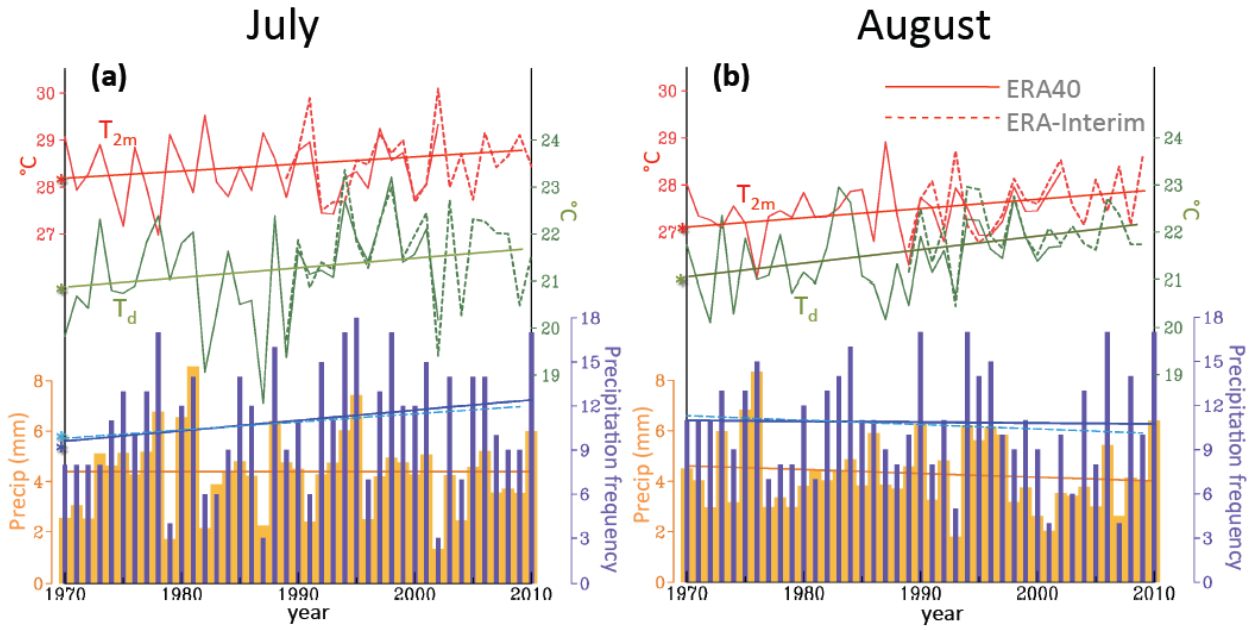


Fig. 4 Monthly mean T_{2m} (red line), T_d (green line), precipitation (orange bar), and frequency of intense precipitation (blue bar) over northern Pakistan in (a) July and (b) August, overlaid with linear trends. Dashed T_{2m} and T_d lines indicate data derived from ERAInterim (1989-2010). Trends that are significant at the 99% confidence interval (CI) are indicated by a star to the left, based on Student's *t*-test. Trends without a star are considered insignificant (*i.e.* < 99% CI). Trends without the inclusion of 2010 are shown as cyan dashed lines.

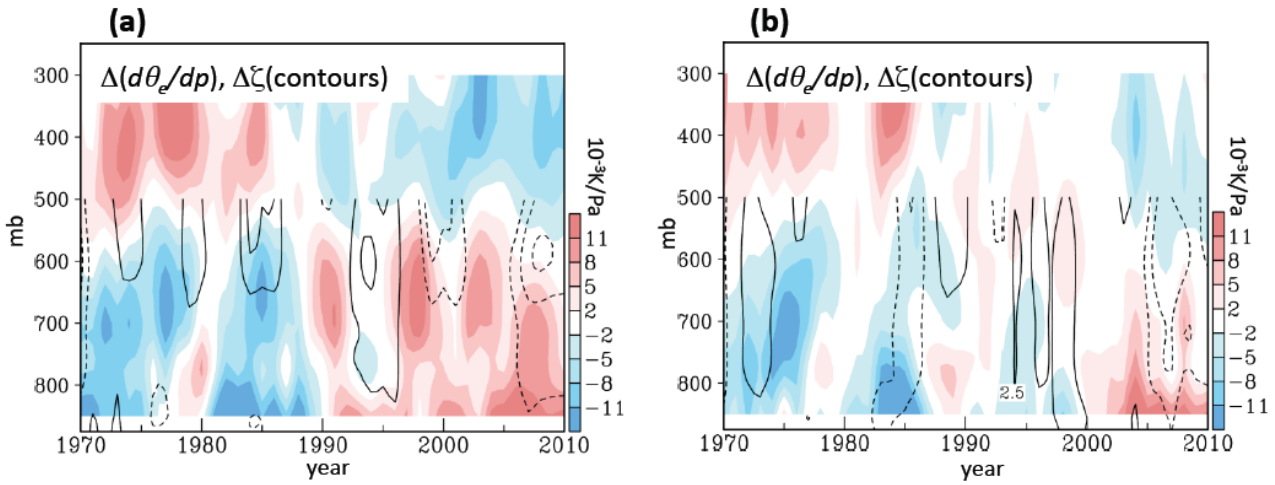


Fig. 5 Time-height cross-sections of the θ_e lapse rate (shadings) and relative vorticity (contour interval 10^{-5} s^{-1} omitting zero) averaged over northern Pakistan for (a) July and (b) August, with the long-term means removed, and divergence of water vapor flux (*i.e.* negative means convergence) during (c) July and (d) August. Linear trends in (c) and (d) are insignificant (blue broken lines).

enhanced conditional instability and, in turn, increased the chance for stronger convective precipitation events. Such a situation is also observed in the August monsoon trough phase (Fig. 5b), though the contrast of $d\theta_e$ /dp anomalies between the lower and upper troposphere is not as apparent as in July, except in the last decade. This recent enhancement of $d\theta_e$ /dp nonetheless coincides with the post-2000 increase in the frequency of intense precipitation (Fig. 5b, after 2000). To investigate changes in ambient circulation that may control variations in $d\theta_e$ /dp, relative vorticity anomalies over northern Pakistan are overlaid on Fig. 5 as contours. There is no clear indication of increased or decreased low-level vorticity (*i.e.*, the monsoon trough); in fact, the negative vorticity anomalies prevailing in the last decade suggest a weakened monsoon trough, which would seem inconsistent with the increased frequency of intense precipitation.

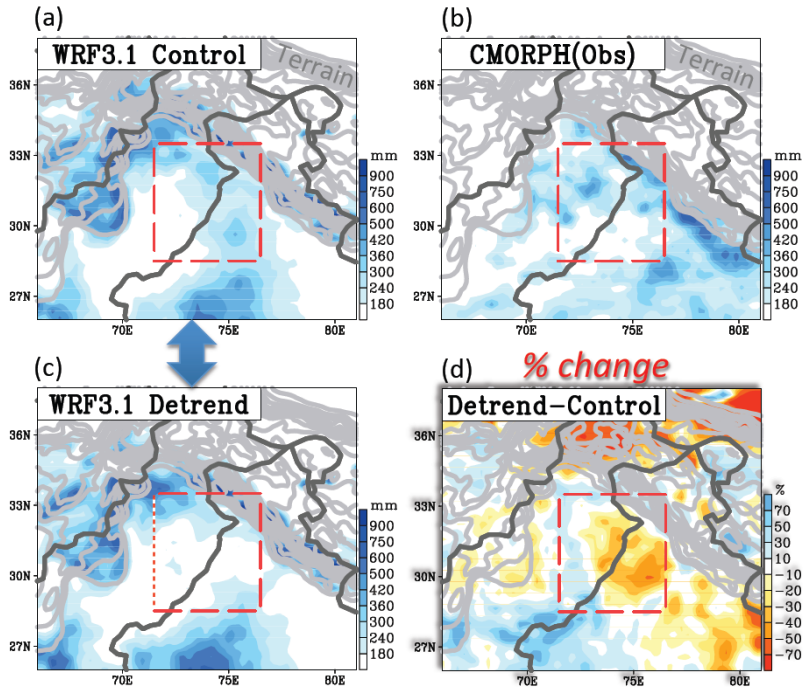


Fig. 6 WRF downscaling of July-August precipitation in 2010 from (a) control run, (c) "detrended run" (see text), and (d) their percentage change. Observed precipitation from CMORPH is shown in (b). Terrain is outlined by gray contours.

- *Modeling:* A WRF downscaling approach was adopted to quantify the impact of climate trends on the extreme nature of the 2010 precipitation event. We performed a control run using NCEP2 as boundary conditions during June–August at the 32 km resolution and compared it with another NCEP2 run that was "detrended" over the 1979–2010 period (on all input variables). The control simulation of July–August precipitation (Fig. 6a) reasonably depicts the observed features (Fig. 6b). The "detrended" simulation (Fig.

6c) shows an apparent reduction in precipitation over northern Pakistan and part of northern India. The reduction in rainfall over the rainstorm region is about 30%; all types of climate trends, including the post-1970 global warming signal, contribute to this reduction.

e. Dynamical implications

It is known that strong monsoons over India and Pakistan occur in association with enhanced upper-level anticyclones and an easterly jet (Webster 2006). To compare with the 2010 anomalies, we used the domain-averaged monthly precipitation over northern Pakistan (domain as in Fig. 6) to regress with the 200 mb geopotential height through 1974–2010. The regression pattern (Fig. 7a) depicts an anticyclone over and to the west of northern Pakistan, consistent with the literature. However, the July 2010 circulation anomalies reveal a marked cyclonic cell over northern Pakistan (Fig. 7b) rather than an anticyclone, as has been the case for strong monsoons. Noteworthy is the robust anticyclone over Eurasia that is linked to the Russian heat wave, embedded in the so-called circumglobal teleconnection (Branstator 2002). Such a discrepancy in the subtropical circulations is intriguing. As illustrated in Ding and Wang (2005), the upper-level anticyclone coupled with strong monsoons has a tropical "baroclinic" structure (*i.e.*, vertically reversed), whereas the rest of the Eurasian wave train features a barotropic structure (*i.e.*, vertically uniform). The 2010 cyclonic anomaly near northern Pakistan resembles the latter, which subsequently enhances southerly water vapor flux. This feature corresponds to the abnormal moisture supply and lifting mechanism found in the July rainstorms (Houze *et al.* 2011). The coupling of the monsoonal southerlies and the upper-level cyclonic flows also suggests a strong tropical-midlatitude interaction. Such interaction can further enhance conditional instability

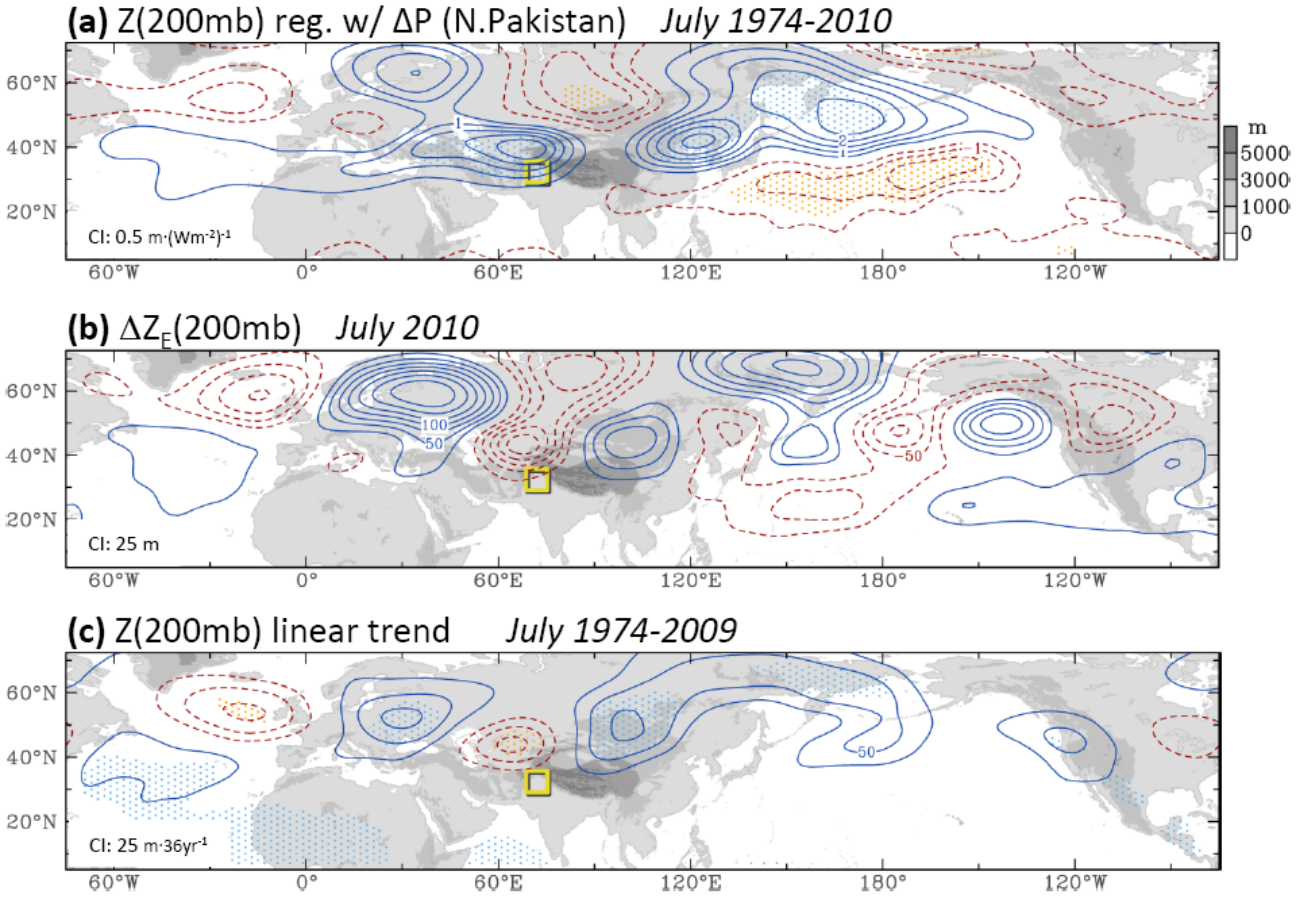


Fig. 7 (a) 200 mb geopotential height regressed upon ΔP averaged in northern Pakistan (yellow box) for July of 1974-2010, (b) departure of the July 2010 geopotential height from the climatology with the zonal mean removed, and (c) linear trends of the July geopotential height during 1974-2009. Dotted areas in (a) and (c) outline significant regions with the 99% confidence interval. Background shadings indicate the topography.

owing to the combination of high temperature and increased moisture (*cf.* Fig. 1). Regardless, any empirical model similar to Fig. 7a would suggest the cyclonic anomaly in 2010 be associated with a weak monsoon over Pakistan, rather than a strong one. This is alarming because such a departure from the empirical model would have led to false predictions of the monsoon.

Is the 2010 circulation pattern sporadic, or could it be systematic? Figure 7c shows the linear trends of geopotential height over the 1974–2009 period; the year 2010 was removed in order to conduct an independent assessment of the trends. The result reveals a zonally oriented wave train that is in-phase with the 2010 circulation pattern across the North Atlantic, Eurasia, and East Asia. Such long-term circulation changes may be associated with the Eurasian warming of recent decades, a phenomenon thought to enhance a land-ocean thermal gradient conducive to a strong monsoon in India (Kumar *et al.* 1999). However, the cause of the transcontinental wave train and its changing relation with the Indian monsoon is complex and requires further analysis.

4. Discussion and conclusion

We summarized the paper of Wang et al. (2011), which examined the 2010 summer precipitation in Pakistan and its link with climate change, and introduced a new dynamical downscaling analysis. We suggest that summer precipitation in northern Pakistan comprises two distinct phases: a pre-monsoon trough phase (July) with more-episodic and intense rainfall, occurring without the monsoon trough; and a monsoon trough phase (August) with large but less-episodic rainfall, driven by northward advancement of the monsoon trough coupled with the 30–60 day mode. It is demonstrated that separating the two monsoonal phases aids in the detection and attribution of climate change signals. Long-term trends linked to the 2010 summer rainfall saw an intensification of the pre-monsoon trough phase only. Diagnostic analyses support such intensification as part of a long-term (and ongoing) process—an observation consistent with expectations from a warming and moistening lower troposphere. Conversely, evidence does not support long-term intensification of the August monsoon trough phase in northern Pakistan, nor any trend in the northward extension of the monsoon rainband (other than that accounted for via internal climate variability such as the La Niña).

Large-scale circulation analysis suggests that the typical linkage between increased monsoon rains in northern Pakistan and enhanced anticyclones in the upper troposphere has changed, such that cyclonic anomalies combined with enhanced conditional instability (*i.e.*, due to a warming surface and increasing moisture supply) could also trigger intense precipitation—a process substantiated through the unambiguous increase in convective activity. Various global reanalyses agree that the circulation pattern in July 2010, while abnormal, is not sporadic. Instead, it is part of a long-term trend of the larger-scale circulation that defies the typical monsoon dynamics expected in northern Pakistan (*i.e.*, strong monsoons associated with upper-level anticyclones). At this point it is reasonable to conclude that the increased convective activity in northern Pakistan is not only a result of unusual circulation anomalies (*i.e.*, internal variability), but rather is a combined process of circulation changes acting on local destabilization due to warming and moistening of the lower troposphere (*i.e.*, internal + external variabilities). This process may, at least partially, explain the observation by Houze *et al.* (2011) that "the rainstorms responsible for the floods were of a type that does not normally occur in this region."

References

Branstator, G., 2002: Circumglobal teleconnections, the jet stream waveguide, and the North Atlantic Oscillation. *J. Climate*, **15**, 1893–1910.

Chen, T.-C., J.-H. Yoon, and S.-Y. Wang, 2005: Westward propagation of the Indian monsoon depression. *Tellus*, **57A**, 758–769.

Ding, Q., and B. Wang, 2005: Circumglobal teleconnection in the Northern Hemisphere summer. *J. Climate*, **18**, 3483–3505.

Houze, R. A., K. L. Rasmussen, S. Medina, S. R. Brodzik, and U. Romatschke, 2011: Anomalous atmospheric events leading to the summer 2010 floods in Pakistan. *Bull. Amer. Meteor. Soc.*, **92**, 291–298.

Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydromet.*, **5**, 487–503.

Kanamitsu, M., and Coauthors, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.,* **83**,

1631–1643.

Khan, J. A., 1993: *The Climate of Pakistan*, Rehber Publishers, Karachi. 79 pp.

Kodama, Y.-M., A. Ohta, M. Katsumata, S. Mori, S. Satoh, and H. Ueda, 2005: Seasonal transition of predominant precipitation type and lightning activity over tropical monsoon areas derived from TRMM observations. *Geophys. Res. Lett.*, **32**, L14710, doi:10.1029/2005GL022986

Kumar, K. K., B. Rajagopalan, and M. A. Cane, 1999: On the weakening relationship between the Indian monsoon and ENSO. *Science*, **284**, 2156–2159.

Marshall, M., 2010: Frozen jet stream links Pakistan floods, Russian fires. *New Scientist*, Iss. 2773, 14-15.

PMD, 2010: Pakistan's Monsoon 2010 update. Issued by Climate Data Processing Centre, Pakistan Meteorological Department (PMD), Karachi-75720. 4 pp.
http://www.pakmet.com.pk/cdpc/prg/monsoon2010/monsoon2010progress.pdf

Uppala, S. M., and Coauthors, 2005: The ERA-40 re-analysis. *Q. J. Royal Meteor.Soc.*, **131**, 2961–3012.

Uppala, S. M., D. Dee, S. Kobayashi, P. Berrisford, and A. Simmons, 2008: Towards a climate data assimilation system: Status update of ERA-Interim, *ECMWF Newsletter*, **115**, 12–18.

Wang, S.-Y., R. E. Davies, W. R. Huang, and R. R. Gillies, 2011: Pakistan's two-stage monsoon and links with the recent climate change. *J. Geophys. Res.*, **116**, D16114, doi:10.1029/2011JD015760.

Webster, P. J, 2006: The coupled monsoon system. *The Asian Monsoon*. Bin Wang, ed., 3–66.

Xie, P., A. Yatagai, M. Chen, T. Hayasaka, Y. Fukushima, C. Liu, and S. Yang, 2007: A gauge-based analysis of daily precipitation over East Asia. *J. Hydrometeor.*, **8**, 607–627.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Association of U.S. Tornado Counts with the Large-scale Environment on Monthly Time-scales

Michael K. Tippett [1], Adam H. Sobel [2, 3] and Suzana J. Camargo[3]

[1]*International Research Institute for Climate and Society, Columbia University, Palisades, NY*
[2]*Department of Applied Physics and Applied Mathematics and*
Department of Earth and Environmental Sciences, Columbia University, New York, NY
[3]*Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY*

1. Introduction

The record-breaking tornado activity of spring 2011 highlighted both the large impact of tornadoes on human life and property (an estimated 336 tornadoes during the April 2011 "Super Outbreak" killed 346 people and caused total damages estimated as high as $10 billion) and the difficulty of attributing increased tornado activity to large-scale climate phenomena such as La Niña or climate change.

Many studies have successfully associated tornado activity with immediate observed environmental parameters (Brooks *et al.*, 1994; Brooks *et al.*, 2003). Observations and short-term predictions of these environmental parameters provide guidance for forecasters who currently issue severe thunderstorm outlooks up to one week in advance and tornado warnings when tornadoes appear imminent (Shafer *et al.*, 2010). Relatively few studies have related seasonal or monthly tornado activity with time-averaged climate phenomena, though the roles of ENSO and the Intra-Americas Sea low-level jet have investigated on such time-scales (Cook and Schaefer, 2008; Muñoz and Enfield 2011).

The identification of associations between tornado activity and environmental parameters on seasonal or monthly time-scales is highly desirable since such associations would help to explain observed variability and could provide a path toward extended-range prediction of tornado activity. Here we investigate the connection between monthly tornado activity and atmospheric parameters using an empirical "tornado index" constructed to represent the expected monthly tornado activity conditional on the local environment. A similar index-based approach has been used in the study of tropical cyclone (TC) formation, identifying favorable environmental factors and capturing features of climatological and interannual TC genesis variability (Camargo *et al.*, 2007; Tippett *et al.*, 2011). We find that the tornado index constructed here is able to reproduce aspects of climatological and interannual variability.

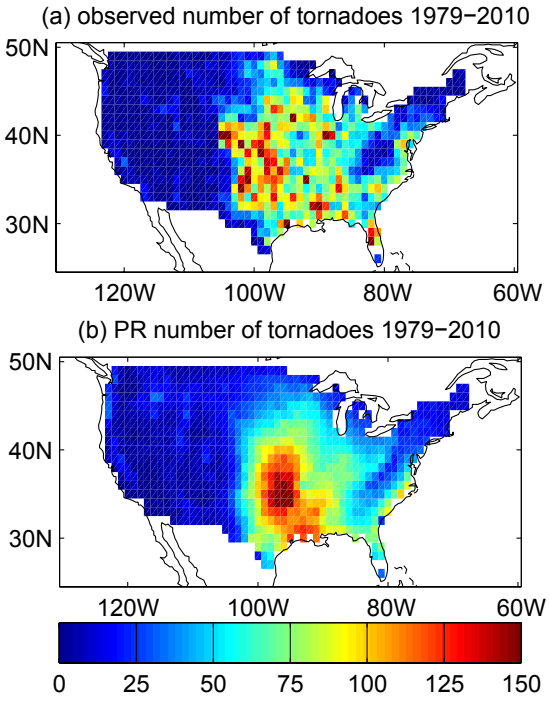


Fig. 1 Colors indicate the (a) observed and (b) Poisson regression (PR)-fit number of tornadoes for the period 1979-2010.

2. Data and methodology

2.1 Tornado data

U.S. tornado data covering the period 1979-2010 are taken from the Storm Prediction Center Tornado, Hail and Wind Database (Schaefer and Edwards, 1999). There are substantial inhomogeneities in the reported data. An

Correspondence to: Michael K. Tippett, International Research Institute for Climate & Society, Columbia University, Palisades, NY; E-mail: tippett@iri.columbia.edu.

adjusted[1] annual number of U.S. tornadoes can be computed using a trend line for the period 1954-2007 and taking 2007 as a baseline. We compute a monthly tornado climatology by counting for each calendar month the number of reported tornadoes in each grid box of a 1 x 1 degree latitude-longitude grid. There is no adjustment for trends in the computation of the gridded monthly tornado climatology, likely leading to an underestimate of the climatological number of tornadoes.

2.2 NARR data

Monthly averaged environmental parameters are taken from the North American Regional Reanalysis (NARR; Mesinger and Coauthors, 2006). NARR data are provided on a 32-km Lambert conformal grid which we interpolate to a 1 x 1 degree latitude-longitude grid. We primarily select environmental parameters that have been recognized as being relevant to tornado formation (Brooks *et al.*, 1994; Brooks *et al.*, 2003): surface convective available potential energy, surface convective inhibition, best (4-layer) lifted index, the difference in temperature at the 700 hPa and 500 hPa levels divided by the corresponding difference in geopotential height (lapse rate), the average specific humidity between 1000 hPa and 900 hPa (mixing ratio), 3000-0m storm relative helicity, the magnitude of the vector difference of the 500 hPa and 1000 hPa winds (vertical shear), precipitation, convective precipitation and elevation. Lapse rate and vertical shear are computed using monthly averages of the constituent variables.

2.3 CFSv2 reforecasts

Forecast data comes from reforecasts of the Climate Forecast System version 2 (CFSv2), the successor of the CFS version 1 (Saha *et al.*, 2006) with improved physics, increased resolution and overall improved skill (Yuan *et al.*, 2011). CFSv2 reforecasts are initialized using the Climate Forecast System Reanalysis, a coupled data assimilation system (Saha *et al.*, 2010). Over the 29-year reforecast period 1982-2010, single member ensemble forecasts are started every 5 days (without accounting for leap year days) at 0, 6, 12 and 18Z and integrated for 9 full months. Zero-lead forecasts are initialized in the month prior to the month being predicted or in the first pentad of the month being predicted. Here we use the most recent 16 ensemble members.

2.4 Poisson regression

We relate the *climatological* monthly number of U.S. tornadoes with *climatological* monthly averages of NARR atmospheric parameters using a Poisson regression (PR), the standard statistical method for modeling of count data. A similar method was used to develop a TC genesis index (Tippett *et al.*, 2011). Fitting climatological data avoids many of the issues related to inhomogeneity in the observational record. Moreover, fitting the PR with climatological data means that the yearly varying data provide an independent test of the PR fit.

To select the parameters to include in the PR, we perform a forward selection procedure in which one variable is added at a time to the PR, and the variable whose addition most reduces the deviance is identified. The deviance is computed using 10-fold cross-validation in which the data is randomly separated in 10 subsets, 9 of which are used to estimate the regression coefficients, and one is used to compute the deviance. Using 10 partitions, we obtain 100 estimates of the deviance. There is a substantial decrease in the deviance as the number of environmental parameters is increased from 1 to 2, but further increases in the number of environmental parameters do not result in significant (95% level) decreases in deviance.

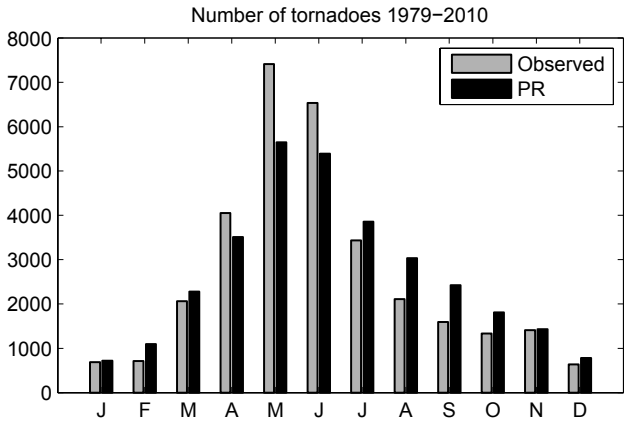


Fig. 2 Observed (gray) and Poisson regression (PR)-fit (black) number of tornadoes per month during the period 1979-2010.

[1] http://www.spc.noaa.gov/wcm/adj.html

3. Results

The spatial distribution of the total number of reported tornadoes 1979-2010 and the corresponding PR-fit values are similar (Fig. 1). The dominant feature is the so-called ``Tornado Alley'' running north-south in the Central U.S. Observations and PR-fit values show few tornadoes west of the Rocky Mountains and over the Appalachian Mountains. Relatively high tornado activity is observed in northeastern Colorado and Florida but is not seen in the PR-fit values. This difference may be due to non-supercell tornadoes being common in both of these areas (Brooks and Doswell, 2001) while the PR focuses on quantities related to supercell dynamics. Tornado activity along the coasts of the Atlantic seaboard states are seen in both observations and PR-fit values. The observed and PR-fit seasonal cycles of tornado occurrence have similar phasing and amplitude (Fig. 2) with maximum values occurring in May, followed closely by those of June. There is no explicit accounting for seasonality in the PR; seasonality comes from the environmental parameters. PR-fit values are too small during May and June and too large in late summer and autumn.

Although the PR fits well the climatological tornado data on which it was developed, there is no assurance that the same relations are relevant to year-to-year tornado activity variability. However, when the PR developed with climatological data is applied to yearly varying monthly values (1979-2010), the PR-estimated values correlate well with the observed monthly number of tornadoes (Table 1); the seasonal cycle is removed and does not contribute to the monthly correlations. The correlation between the observed annual number of tornadoes and that given by the PR is 0.51, and this value increases to 0.64 when observations are adjusted to account for changes in observing system (Fig. 3a). Notable also is the PR value of 501 for April 2011, the most active U.S. tornado month on record (Fig. 3b).

The demonstrated relation between monthly averaged environmental parameters and monthly tornado numbers provides an approach for extended-range forecasts. One first predicts the environmental parameters in the PR and then uses the PR to predict the impact of those parameters on tornado activity. We assess the feasibility of such a forecast procedure using the CFSv2 reforecasts of monthly averaged environmental parameters. Correlations between observed numbers of tornadoes and PR-estimated values are computed by month. For the most part the correlations are modest. However, the June correlation is particularly strong (0.79), though the CFSv2 predicted values have low amplitudes compared to observations (not shown).

4. Conclusions

We have shown that monthly U.S. tornado activity can be related to observed monthly-averaged environmental parameters and that Poisson regression can be used to construct an index that captures aspects of the climatological and year-to-year variability of tornado activity. To the extent that the index can be forecast skillfully, it provides a basis for monthly tornado activity forecasts. The same procedure based on an operational seasonal forecast model shows significant skill in forecasting June tornado activity.

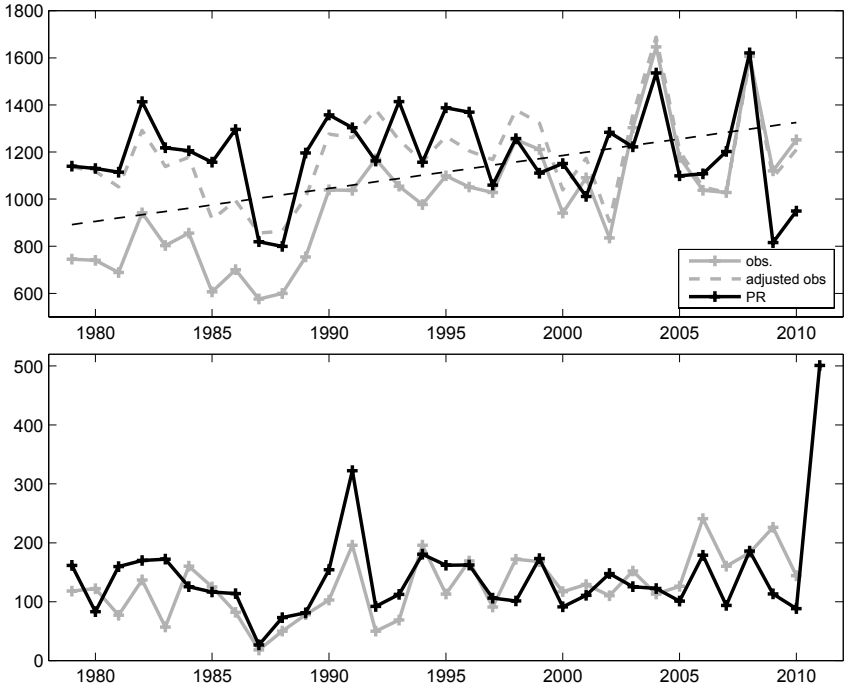


Fig. 3 Observed (gray), trend adjusted observations (dashed gray) and Poisson regression (black) (a) annual and (b) April numbers of tornadoes during the period 1979-2010. The dashed black line in panel (a) is the 1954-2007 trend-line used in the adjustment. Panel (b) includes the April 2011 Poisson regression value.

References

Brooks, H. E., J. W. Lee, and J. P. Craven, 2003: The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmos. Res*., **67-68**, 73–94.

Brooks, H. E. and C. A. Doswell, III, 2001: Some aspects of the international climatology of tornadoes by damage classification. *Atmos. Res*., **56 (1-4)**, 191 – 201.

Brooks, H. E., C. A. Doswell, III, and J. Cooper, 1994: On the environments of tornadic and nontornadic mesocyclones. *Wea. Forecasting*, **9**, 606–618.

Camargo, S. J., A. H. Sobel, A. G. Barnston, and K. A. Emanuel, 2007: Tropical cyclone genesis potential index in climate models. *Tellus*, **59A**, 428–443.

Cook, A. R. and J. T. Schaefer, 2008: The Relation of El Niño–Southern Oscillation (ENSO) to Winter Tornado Outbreaks. *Mon. Wea. Rev*., **136**, 3121–3137, doi:10.1175/2007MWR2171.1.

Mesinger, F. and Coauthors, 2006: North American Regional Reanalysis. *Bull. Am. Meteor. Soc*., **87**, 343–360.

Muñoz, E. and D. Enfield, 2011: The boreal spring variability of the Intra-Americas low-level jet and its relation with precipitation and tornadoes in the eastern United States. *Clim. Dynam*., **36**, 247–259, doi:10.1007/s00382-009-0688-3.

Saha, S., et al., 2006: The NCEP Climate Forecast System. *J. Climate*, **19**, 3483–3517.

Saha, S., et al., 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc*., **91**, 1015–1057, doi:10.1175/2010BAMS3001.1.

Schaefer, J. T. and R. Edwards, 1999: The SPC Tornado/Severe Thunderstorm Database. *Preprints, 11th Conf. Applied Climatology*, Dallas, TX, Amer. Meteor. Soc., 215–220.

Shafer, C. M., A. E. Mercer, L. M. Leslie, M. B. Richman, and C. A. Doswell, III, 2010: Evaluation of WRF model simulations of tornadic and nontornadic outbreaks occurring in the spring and fall. *Mon. Wea. Rev*., **138**, 4098–4119, doi:10.1175/2010MWR3269.1.

Tippett, M. K., S. J. Camargo, and A. H. Sobel, 2011: A Poisson regression index for tropical cyclone genesis and the role of large-scale vorticity in genesis. *J. Climate*, **24**, 2335–2357.

Yuan, X., E. F. Wood, L. Luo, and M. Pan, 2011: A first look at Climate Forecast System version 2 (CFSv2) for hydrological seasonal prediction. *Geophys. Res. Lett*., **38**, L13 402.

4. CLIMATE MODELS

AND FORECAST TECHNIQUES

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

An Iterative Projection Method to Calculate EOFs Successively without Use of the Covariance Matrix

Huug van den Dool

Climate Prediction Center, NCEP/NWS/NOAA, Camp Springs, MD

1. Basics

Given is a data set X(s,t) at discrete positions in time (equidistant $1 \le t \le n_t$) and space($1 \le s \le n_s$). The concept of EOFs is such that the data set X(s,t) can be written as

$$X(s, t) = \sum_m \alpha_m(t)\, e_m(s) \qquad (1)$$

where both

$$\sum_t \alpha_m(t)\, \alpha_n(t) = 0 \qquad (1a)$$

and

$$\sum_s w(s)\, e_m(s)\, e_n(s) = 0 \qquad (1b)$$

for any pair $n \neq m$. One EOF, mode m, consists of a time series $\alpha_m(t)$ and a space pattern $e_m(s)$. w(s) is a weight that varies in space (depending on the type of grid or the station distribution used). The summation over m has to be over a sufficient number of modes (<= the smaller of n_t and n_s) to achieve the equal sign in Eq(1). Orthogonality of the set $\alpha_m(t)$ alone (1b), or the set $e_m(s)$ alone (1a), is sufficient to satisfy (1) – the bi-orthogonality (both (1a) and (1b) valid) is one of the features that makes EOFs unique compared to other orthogonal functions.

2. Orthogonality relations

Multiply the lhs and rhs of (1) by $\alpha_k(t)$ and sum over all times t (k is a specific mode number). Using (1a) the result is:

$$e_m(s) = \sum_t \alpha_m(t)\, X(s, t) \,/\, \sum_t \alpha^2{}_m(t) \qquad (2a)$$

Likewise, multiply the lhs and rhs side of (1) by $e_k(s)$ and sum over all space s with weight w(s). Using (1b) the result is:

$$\alpha_m(t) = \sum_s w(s)\, e_m(s)\, X(s, t) \,/\, (\sum_s w(s)\, e^2{}_m(s)) \qquad (2b)$$

The above are the orthogonality relationships we will employ in the iteration. They state, (2a), that the space pattern of a particular mode can be calculated by projecting the data X onto the time series of that mode, and (2b) that by projecting the data X onto a space pattern the associated time series can be retrieved. These calculations are quite simple.

3. The iteration method

The above is abstract since we don't know yet the EOFs. We only know the data X(s,t). Now randomly pick (or make) a guess time series $\alpha^0(t)$, and stick into (2a). This yields a pattern, the starting guess for the space pattern, denoted as $e^0(s)$. Stick $e^0(t)$ into (2b). This yields $\alpha^1(t)$. This two-step process is counted as one iteration. With only a few more iterations one generally converges quickly to the first EOF $\alpha_1(t)$, $e_1(s)$. We now form a reduced data set as

$$X^{reduced}(s,t) = X(s,t) - \alpha_1(t)\, e_1(s) \qquad (3)$$

and repeat the procedure. One now finds mode#2, reduces further etc. It does not matter whether one starts the iteration with a guess time series or a guess spatial map.

Correspondence to: Huug van den Dool, Climate Prediction Center, NCEP/NWS/NOAA, Camp Springs, MD 20746; E-mail: Huug.Vandendool@noaa.gov.

The above describes tersely how a number of (leading) EOFs can be calculated rather quickly. Some examples are given in Appendix II. Please note that EOFs can thus be calculated without explicitly evaluating the covariance matrix.

4. Discussion

The iterative method of calculating EOFs is not widely known in meteorology and climatology. The above was described once before in an appendix in an earlier paper (Van den Dool *et al.* 2000) and possibly elsewhere even earlier. A paper by Baldwin *et al.* (2009) drew attention to iterative projection methods as being particularly useful for calculating EOFs from VERY LARGE data sets. Why? Commonly known and widely used methods determine EOFs as eigenvectors of the covariance matrix, but to determine the covariance matrix is a huge computational task when the dimensions of the data matrix X are very large. This situation arises for any modern Reanalysis, even for a single variable.

Some in the audience of CDPW#36 thought of the above projection method (a ping-pong between (2a) and (2b)) as simple and instructive, and therefore useful regardless of the size of the problem. Advantages in terms of CPU are none (to write home about) for a small problem, but at least the code is short. The iteration is a factor 30 faster (than the route via the covariance matrix) for a problem where n_s and n_t are order 4000. When n_s and n_t go higher (10,000 or 100,000) the covariance matrix method quickly becomes 'impossible' (depending on computer and smarts of programmer), while iteration is still possible (although not inexpensive).

Why does the iteration method work? Apparently, the reason is the same as the reason why the "power method" works. Suppose we do know the (real, square and symmetric) covariance matrix Q = either X^TX or XX^T. A guess $e^0(s)$ can always be written as a linear combination of projections onto the unknown EOFs:

$$e^0(s) = \sum_m \alpha_m e_m(s) \qquad (4)$$

Keep in mind $Q e_m = \lambda_m e_m$ by definition. Execution of $Q e^0$ (which is done for the power method) thus yields $\sum_m \lambda_m \alpha_m e_m(s)$, *i.e.* the projection onto $e_m(s)$ gets multiplied by the eigenvalue of that mode during each iteration. Consequently the projection onto the gravest modes is amplified relative to the lesser modes, and the iteration converges to mode#1. In the classic power method this process may require renormalizations, but with (2a)-(2b) this is done automatically. Curiously, iteration works for the same reason as the power method even though Q is not calculated.

Theoretically the iteration method may fail, if, for instance, we have two EOF modes with EXACTLY identical eigenvalue. This never happens. Likewise, if we came up with a guess that has no projection onto any mode the method will fail. No worry about that either on finite precision machines. What is a problem is that after calculating ten or a hundred EOF modes, the remaining spectrum is very flat, and convergence of the iteration becomes slower and slower. One would have to iterate longer and longer for less and less added variance. This circumstance kills the method as an improvement over the covariance matrix based approach when too many modes are required. Truncation is in order somewhere.

In Toumazou and Cretaux (2001) some sort of attempt is made to systematically list EOF calculation methods. According to these authors there are three groups of approaches. The first is to (i) calculate singular vectors from the rectangular asymmetric X. The second and third group of methods use directly or indirectly (properties of) the square symmetric and real covariance matrix to apply either (ii) the eigenmode method called QR method, or (iii) a Lanczos strategy. It is not 100% clear where the iterative projection described in the above fits in this list, and whether Toumazou and Cretaux's list is even complete. Iterative projection appears closest to the Lanczos method. Incidentally, the connection between eigen vectors of the covariance matrix and singular vectors of the data matrix is given in Appendix I.

Acknowledgement. Discussions with Gaby Hegerl and Chris Bretherton (in about 2001) clarified why the iteration method works. The paper by Mark Baldwin in 2009 gave impetus to apply the iteration method to a huge data set like Reanalysis data. I do acknowledge heroic attempts by Bob Kistler, using all the computer tricks known to mankind, to calculate Q. Suranjana Saha made the CFSR data available. Input by Michael

Tippett on a systematic list of EOF methods is much appreciated. The appendix was copied (and adjusted) from Wikipedia.

References

Baldwin, Mark P., David B. Stephenson, Ian T. Jolliffe, 2009: Spatial weighting and iterative projection methods for EOFs. *J. Climate*, **22**, 234–243.

Saha, S., and Co-authors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015-1057. doi: 10.1175/2010BAMS3001.1

Van den Dool, H. M., S. Saha, Å. Johansson, 2000: Empirical orthogonal teleconnections. *J. Climate*, **13**, 1421–1435.

Toumazou, Vincent, Jean-Francois Cretaux, 2001: Using a Lanczos eigensolver in the computation of empirical orthogonal functions. *Mon. Wea. Rev.*, **129**, 1243–1250.

Appendix I:

Formally, the singular value decomposition of an $n_t \times n_s$ real matrix X is a factorization of the form $X = U \Sigma V^T$,where U is an $n_s \times n_s$ real matrix, Σ is an $n_s \times n_t$ diagonal matrix with nonnegative real numbers on the diagonal, and V^T (the transpose of V) is an $n_t \times n_t$ matrix. The diagonal entries $\Sigma_{i,i}$ of Σ are known as the singular values of X. The m columns of U and the n columns of V are called the left singular vectors and right singular vectors of X, respectively.

Singular value decomposition and eigen-decomposition are closely related. Namely: The left singular vectors of X are eigenvectors of XX^T . The right singular vectors of X are eigenvectors of X^TX. The non-zero singular values of Σ are the square roots of the non-zero eigenvalues of X^TX or XX^T

Appendix II: Some examples

First is a low resolution example when the covariance matrix can comfortably be calculated. We use JFM seasonal Z500 mean over the period 1948-2011 at 2.5 degree grid north of 20ºN. There are 64 time levels and 144*29=4176 gridpoints, a 'small' problem nowadays. Fig.1 shows that the first four EOFs are the same for the covariance matrix method (top) and iteration method (bottom). This result establishes that the iteration methods (a completely different code) works. The reader will also recognize renditions of the NAO and PNA as the first two modes.

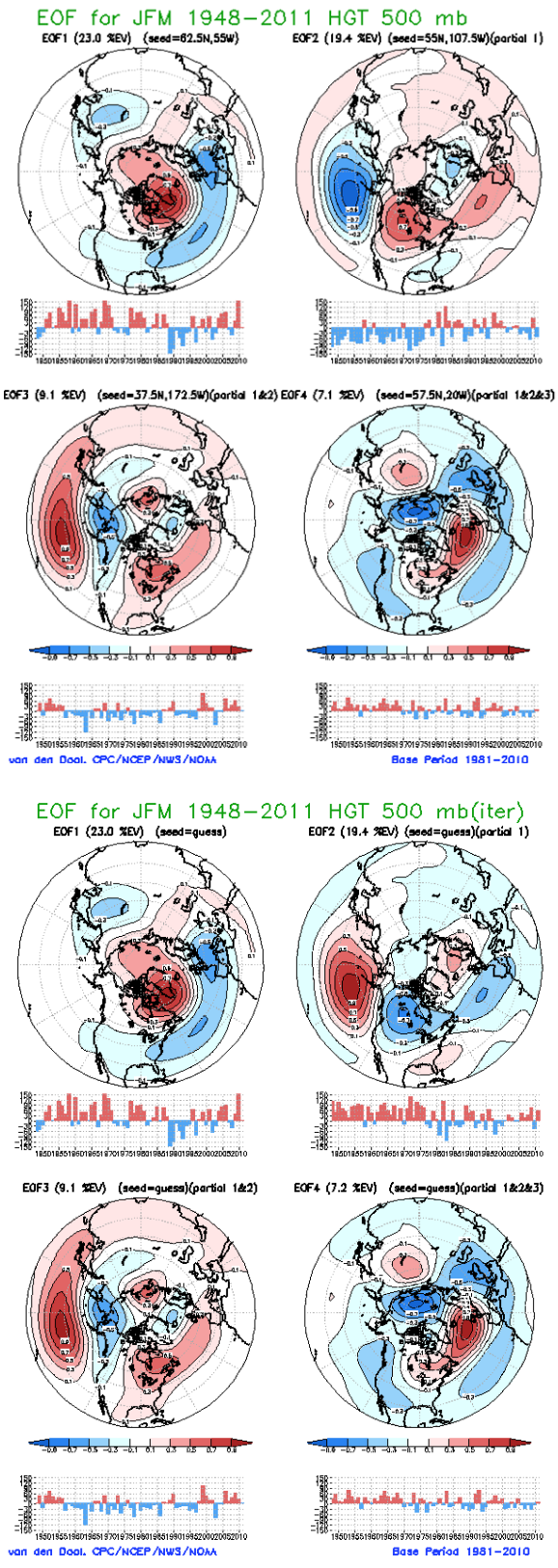


Fig.1 The first four EOFs for three month mean JFM Z500 for the area 20ºN-pole, and 1948-2011. The four panels on the top (bottom) are calculated as eigenvectors of the covariance matrix (based on iteration). Blue is negative, red positive. The space patterns are normalized, so the decreasing variance can be seen in the smaller amplitude of the time series.

Figure 2 shows the successive states obtained by the first four iterations. We start from an absurd first guess (upper left), a wheel, that after just one iteration already takes on the familiar features of the first EOF looking like the NAO. Both in terms of space pattern and time series we quickly arrive at the first EOF, which was given in upper left in Fig.1. Regardless of the first guess we always converge to the first EOF, although it may take more iterations if the first guess happens to be close to the 2nd EOF (the PNA); example given in CDPW ppt, but not shown.

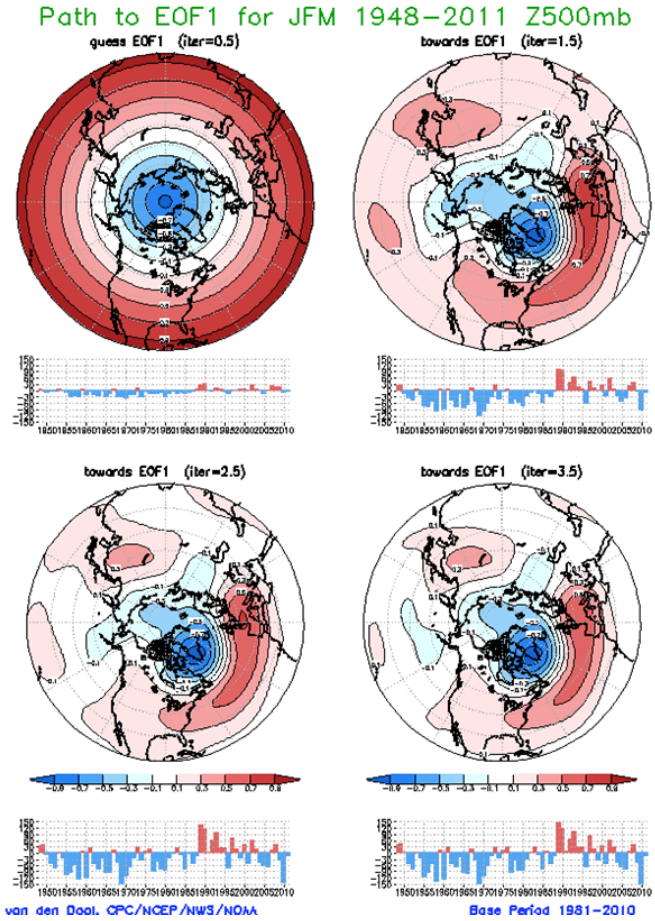


Fig. 2 Successive states obtained by iteration starting with an arbitrary (even non-meteorological) field in the upper left (a 'wheel'). The time series in the upper left shows small projections onto the first guess space patterns – but this is enough to make just a single iteration very efficient. The path to EOF#1 appears very quick.

Figure 3 is testimony that we achieved something rather unlikely, namely we calculated the first 200 EOFs of a data set of year-round daily (1979-2010; n_t = 11688) global full resolution 2 meter temperature data (at n_s = 663552 Gaussian gridpoints) produced by CFSR (Saha *et al.* 2010). To be frank: we did the iteration at lower spatial resolution for speed, and made only the final iteration on full resolution. We are not sure that these are the true EOFs beyond about mode#25, because convergence appears to be an issue. This qualified success is nevertheless noteworthy and may potentially be considered a big plus for data compression and data transmission issues.

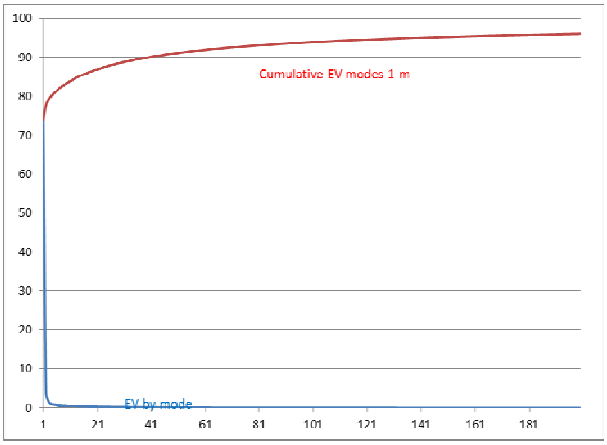


Fig. 3 The variance explained (EV) of daily year-round global 2 meter temperature (given on a T384 grid, 1979-2010) by EOF modes calculated one by one by the iteration method. The blue line is EV by mode, the red line cumulative. The high EV of the first mode (>70%) is reflective of the role of the enormous annual cycle in T2m on earth.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Introduction to the KMA-Met Office Joint Seasonal Forecasting System and Evaluation of its Hindcast Ensemble Simulations

Hyun-Suk Kang, Kyung-On Boo, and ChunHo Cho

National Institute of Meteorological Research, Korea Meteorological Administration, Seoul, Korea

1. Introduction

Dynamical seasonal forecasts using coupled models, *i.e.*, 1-tier approach, are now routinely made at many operational centers in the world. For example, among twelve Global Producing Centers (GPCs) that have contributed for providing their own real-time seasonal forecast to WMO Lead Center for Long-Range Forecast Multi Model Ensemble (LC-LRFMME), seven centers are using coupled models while only five centers are still based on two-tier approach. The rapid transition from two-tier to one-tier approach in seasonal forecast are mainly caused by recent progresses in development of coupled climate models and enlargement of understanding air-sea interactions obtained from international collaborative efforts such as TOCA program (Wang *et al.*, 2009). In this context, Korea Meteorological Administration (KMA), as the WMO LC-LRFMME jointly with NOAA and one of the GPCs, is also trying to replace its operational seasonal forecast model with a coupled model by the collaboration with U. K. Met Office.

Recently, the GloSea4 (Global Seasonal Forecasting System version 4) of the Met Office based the HadGEM3-AO was implemented and hindcast ensemble simulations for 14 years from 1996 to 2009 have been accomplished. The purpose of this article is to introduce the KMA-Met Office Joint Seasonal Forecasting system and to evaluate overall performance of its retrospective seasonal forecast particularly in terms of predictability and skill scores. Section 2 briefly describes the joint forecast system, the model, and design of hindcast simulations. Results of predictability and skill scores on sea surface temperature, precipitation and surface air temperature are shown in Section 3. Finally, Section 4 summarizes the results and further works for the operation of joint system.

2. GloSea4 and its Hindcast Simulations

2.1 GloSea4 and Joint Forecasting System

GloSea4 is the fourth version of the Met Office seasonal ensemble prediction system based on the latest version of HadGEM3 (Hewitt *et al.*, 2010). It consists of the UM (Met Office Unified Model) for atmosphere, NEMO (Nucleus for European Modeling of the Ocean) for ocean, CICE (Los Alamos sea ice model) for sea ice, and MOSES (Met Office Surface Exchange Scheme) for land surface components with OASIS flux coupler. The spatial resolution in the current configuration (GA 2.0) is N96L85 for atmosphere, which is approximately 135 km in the horizontal with 85 vertical levels, and tri-polar ORCA1L75 for ocean, in which the horizontal grid distance are 1 degree with 1/3 of a degree between 20^{o}S and 20^{o}N with 75 vertical levels from the sea surface to the bottom. Details of the GloSea4 description are given in Arribas *et al.* (2011).

One of the distinctive features of the GloSea4 compared to other typical seasonal forecasting system including the current LRF system at KMA, *i.e.*, Global Data Assimilation and Prediction System (GDAPS), is that both the hindcast and forecast suites are run simultaneously, which allows preventing quite a burden of resources for producing model climatology a prior to make seasonal forecast if any modification of the system and/or bud fix is necessary. Hindcast and forecast suites are initialized with the weekly-based time cycle so that they can update initial conditions nearly real-time, which is quite valuable to maintain consistency from short-to-long-range forecasts. Eventually, the major benefit of KMA-Met Office joint forecasting system is to reducing uncertainties of seasonal forecast by share ensemble members as many as possible from two centers for both the hindcast and forecast suites. The only differences will be the initial condition for atmosphere that

Correspondence to: Hyun-Suk Kang, National Institute of Meteorological Research and Korea Meteorological Administration, Seoul, Korea; E-mail: hyunsuk306.kang@gmail.com.

comes from each center's own 4DVAR system. At this moment, since the KMA does not have its own ocean and sea-ice data assimilation system, initial conditions for ocean and sea-ice will be obtained from the Met Office.

2.2 Hindcast Simulations

The ECMWF-interim reanalysis (ERA-interim) is used to initialize the atmosphere and land surface because there is no atmospheric reanalysis available for the HadGEM3-AO. In the case of land surface variables, an anomaly initialization approach, in which ERA-interim anomalies are calculated and then added to the HadGEM3 model climatology, is followed to avoid the inconsistency from the very different land surface model used in HadGEM3 and ERA-interim reanalysis. The ocean field is initialized in the same way as in the forecast suite, *i.e.*, the GloSea4 Ocean Data Assimilation scheme which consists of a parallel version of the Met Office optimal interpolation scheme used for short-range ocean forecasting, except for the fact that atmospheric fluxes to force the ODA scheme are obtained from the ERA-interim rather than from the operational NWP system. The hindcast period is 14 years from 1996 to 2009. Initial dates are 1st, 9th, 17th, and 25th of each month. In order to generate ensemble members by considering model uncertainties, 3 members per each initial date are generated using the stochastic kinetic energy backscatter scheme version 2 (SKEB2) (Shutts, 2005). Therefore, in total, number of 7-month-long integration of the GloSea4 system is 2,016.

3. Results

The observation dataset used for evaluation of the GloSea4 in this study are Hadley Center's sea surface temperature, CMAP precipitation, and ERA-interim reanalysis for the surface air temperature.

3.1 Sea Surface Temperature

Figure 1 shows the bias of seasonal mean SST from 1-month and 3-month lead forecasts. As forecast lead-time increases, in general, the SST bias also increases. In the results of 1-month lead forecasts, the

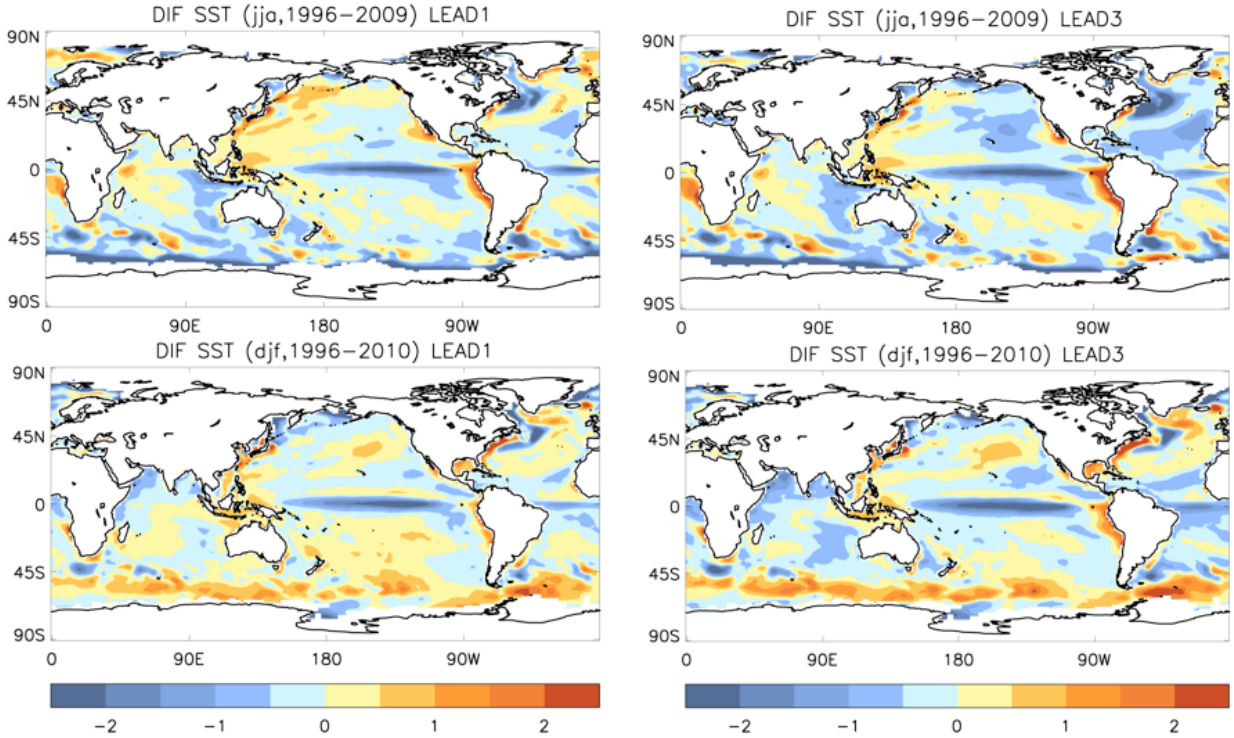


Fig. 1 Biases of seasonal mean SST for boreal summer (upper) and winter (lower) seasons. Left and right panels are obtained from the results one and three months' forecasting lead-time simulations.

strongest cold bias appears along the equatorial Pacific both for the boreal summer and winter seasons. The spatial pattern of SST bias seems to be somewhat systematic, particularly in the southern Hemisphere, which shows overall cold bias in JJA but warm bias in DJF season. Despite the general increase in the amplitude of SST bias according to the forecast leading time, its spatial structure is fairly similar and persistent to each other. On the contrary to the skill scores (*e.g.*, bias or anomaly correlation) that indicate the performance of the model against the observation, signal-to-noise ratio is a measure of predictability that implies that how much the each ensemble member spreads compared to ensemble mean variance. The results of signal-to-noise ratio show also persistent spatial patterns with decreasing values according to the forecast leading months (not shown). As expected, predictability in boreal winter season is higher than in summer season regardless to the forecast leading time.

Anomaly correlations of the NINO3.4 SST anomalies for the four different initial months are shown in Figure 2. Each month has 12 ensemble members with time-lagged initial dates and SKEB2 physics. Overall, skill scores for NINO3.4 index are higher in cold season than in warm season. The skill score drops rapidly from April to July from the simulations initialized on February and November, which is associated limitation of predictability of SST during the spring time, called "spring barrier". The spring barrier issue is one of the common problematic features in coupled GCM, and suspected to be associated with failure of surface wind stress over the equatorial Pacific. It is interesting to note skill score for the JJA forecast is relatively lower than other seasons in the beginning of the forecast, however; the score remains with persistent and relatively higher values for the longer forecast lead-time. The red lines in Figure 2 denote the score calculated from the ensemble mean, and black solid, dashed and dotted lines are average, maximum and minimum values from each individual

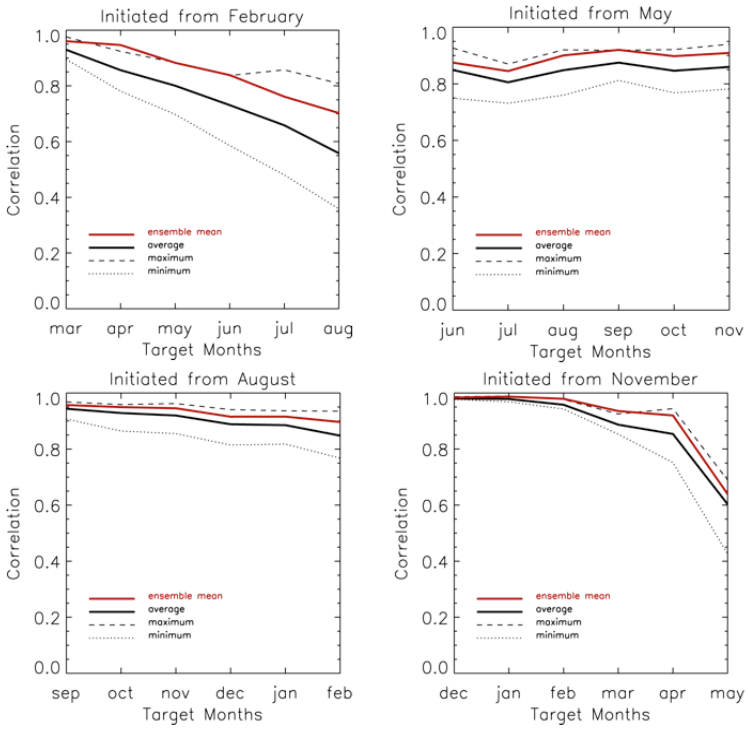


Fig. 2 Correlation skill for the SST anomaly averaged over the NINO3.4 area from the simulations initialized in February, May, August, and November.

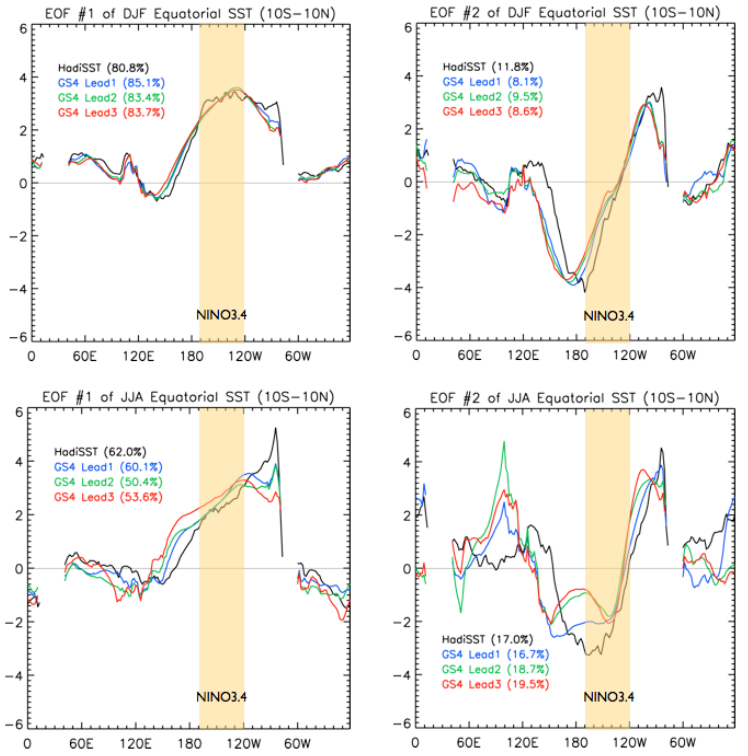


Fig. 3 The first (left panel) and second (right panel) leading EOF modes for the SSTA averaged over 10S-10N. Black, blue, green, and red lines indicate observation, GS4 results with one, two, and three month lead time, respectively.

ensemble members. It is clearly recognized that scores from the ensemble mean are quite close to the maximum scores of individual member, or in some cases, it is superior to maximum of individual ensemble members.

In order to investigate SST variability, EOF analysis was conducted for the SST anomaly against the latitudinal mean between 10°S~10°N (Fig. 3). During the boreal summer (JJA), the observed leading mode represents a peak SSTA in the Nino 3 region rather than 3.4 region (lower left in Fig. 3). Meanwhile, that of boreal winter (DJF) is apparent in somewhat wide areas including both the Nino3.4 and Nino 3 areas. Those patterns of leading mode of SSTA along the equator are captured pretty well by the GloSea4. In JJA, the variability of SST over central Pacific tends to be overestimated by the GloSea4, which are getting stronger to the longer forecast lead-time. From the second leading mode during DJF season, the area of strong variability extends westward in results from the GS4 compared to the observation.

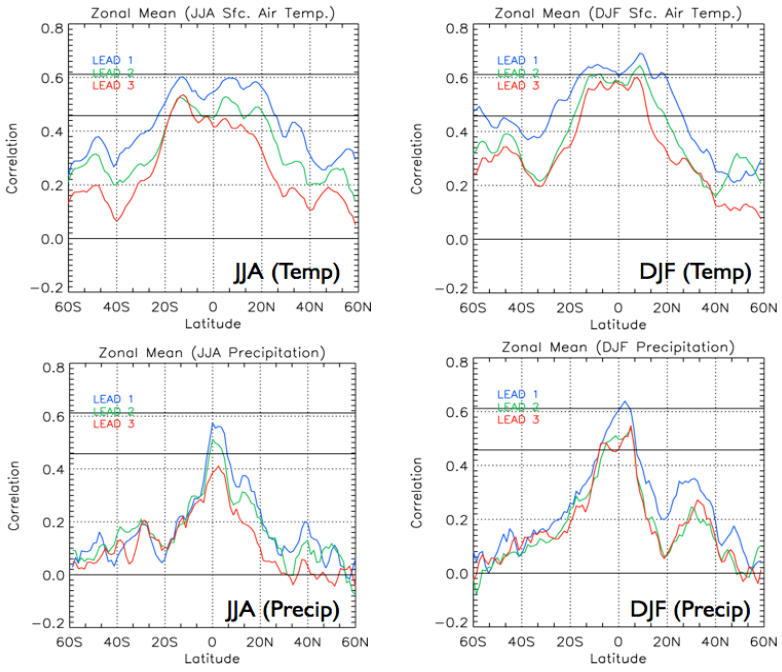


Fig. 4 Anomaly correlation of surface air temperature (upper) and precipitation (lower) for JJA (left panel) and DJF (right panel). Blue, green, and red lines indicate one, two and three months' forecast leads.

3.2 Precipitation and Surface Air Temperature

Since the hindcast period of the GloSea4 is somewhat short (only 14 years), the corresponding correlation value with 0.05 and 0.01 significance levels are somewhat higher which are about 0.45 and 0.61, respectively. As like in other coupled seasonal forecasting system, significant anomaly correlation scores for the surface air temperature and precipitation are concentrated mainly over tropical regional about between 20°S-20°N (Fig. 4). It is clear that anomaly correlation scores are decreasing rapidly in accordance with the forecast leading month. East Asia region, in which the skill scores are quite low as in other extra-tropical areas, meaningful scores with 0.05 significance levels are limited only spring and autumn seasons surface air temperature in cases of less than three months' forecast leading time (not shown). Nevertheless, in terms of practical sense of seasonal forecast, it is promising to note that biases of surface air temperature and precipitation over East Asia are quite systematic and persistent as a function of forecast leading months.

4. Summary and Further Works

In this study, overall skill of the GloSea4 system, which will be operated as an operational seasonal forecasting system at KMA and joint system between KMA and Met Office, have been examined. The skill scores obtained from hindcast ensemble simulations seem to be comparable against with other coupled climate models. However, it should be carefully investigated within intercomparison framework to find out strength and weakness of the GloSea4. Robust evaluation of hindcast ensemble runs including the Asian monsoon, sub-seasonal variability such as MJO and their impacts over Asia should be further investigated. In addition, horizontal resolution both for the atmosphere and ocean will be increased a prior to the operation up to N216 (~ 60 km) and quarter degrees (in extra-tropical region), respectively.

Acknowledgements. This study was supported by the Grant NIMR-2011-B-2. The authors are grateful for the Met Office colleagues, Drs. A. Arribas and C. MacLachlan for their providing successful implementation of the GloSea4 to the KMA.

References

Arribas, A., M. Glover, A. Maidens, K. Peterson, M. Gordon, C. MacLachlan, R. Graham, D. Fereday, J. Camp, A. A. Scaife, P. Xavier, P. McLean, A Colman, and S. Cusack, 2011, The GloSea4 Ensemble Prediction system fore seasonal forecasting, *Mon. Wea. Rev.,* **139**, 1891-1910.

Hewitt, H. T., D. Copsey, I. D. Culverwell, C. M. Harris, R. S. R. Hill, A. B. Keen, A. J. McLaren, and E. C. Hunke, 2010, Design and implementation of the infrastructure of HadGEM3: The next generation Met Office climate modeling system. *Geosci. Model Dev. Discuss.,* **3**, 1861-1937.

Shutts, G., 2005: A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quart. J. Roy. Meteor. Soc.,* **131**, 3079-3102.

Wang, B. and co-authors, 2009: Advance and prospectus of seasonal prediction: Assessment of the APCC/CliPAS 14-model ensemble retrospective seasonal prediction (1980-2004), *Clim. Dyn.,* **33**, 93-117.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

A Simplified Early August Atlantic Basin Seasonal Hurricane Prediction Scheme

Philip Klotzbach

Department of Atmospheric Science
Colorado State University, Fort Collins, Colorado

1. Introduction

The Tropical Meteorology Project (TMP) at Colorado State University (CSU) has been issuing Atlantic basin seasonal tropical cyclone (TC) forecasts in early June with an update in early August since 1984 (Gray 1984). While these forecasts have shown moderate skill in real-time (Klotzbach and Gray 2009) (Figure 1), this paper investigates the potential to improve this skill through the development of a new, simplified early August seasonal prediction scheme that uses newer reanalysis data along with forecasts of El Niño – Southern Oscillation (ENSO) from a dynamical model. This paper briefly discusses the results of this new, primarily statistically-based, forecast scheme. Full documentation is available in Klotzbach (2011).

2. Data

All tropical cyclone (TC) data for this project were taken from the National Hurricane Center's "best track" dataset (Jarvinen *et al.* 1984). The target forecast metric is Net Tropical Cyclone (NTC) activity, which is defined to be the sums of the following six parameters: named storms, named storm days, hurricanes, hurricane days, major hurricanes and major hurricanes days, normalized by their 1950-2000 average values. Consequently, 100 NTC units is an average season by definition.

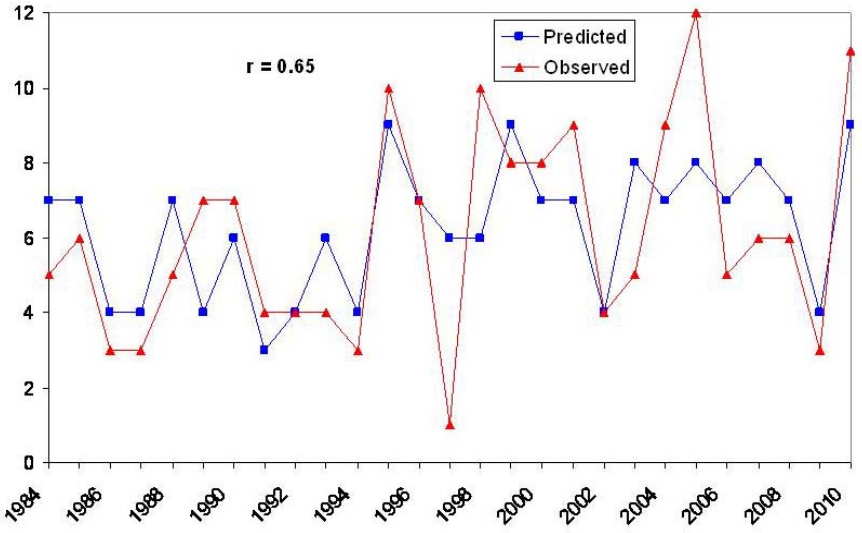


Figure 1 Real-time predicted vs. observed post-31 July Atlantic Basin hurricanes issued by the TMP from 1984-2010.

Large-scale data for the period from 1982-2009 were calculated from the newly-developed Climate Forecast System Reanalysis (CFSR) product (Saha *et al.* 2010). Improved coupling, vertical resolution and data assimilation are generally considered to make the CFSR a more accurate product than its predecessor, the NCEP/NCAR Reanalysis I (Kistler *et al.* 2001). However, the CFSR is currently not available in real-time, so consequently, NCEP/NCAR Reanalysis I products are used to estimate predictor values in real-time forecasts. Testing of atmospheric predictor values prior to 1982 was done using the 20th Century Reanalysis (Compo *et al.* 2011).

Sea surface temperatures (SSTs) from the NOAA Optimum Interpolation SST (OI SST) version 2 are utilized from 1982-present (Reynolds *et al.* 2002). Prior to 1982, SST measurements are calculated from the NOAA Extended Reconstructed SST v3b dataset (Smith *et al.* 2008).

ENSO hindcasts from the European Centre for Medium-Range Weather Forecasts (ECMWF) seasonal forecast system 3 model (Stockdale *et al.* 2011) were provided by Frederic Vitart.

Correspondence to: Philip Klotzbach, Department of Atmospheric Science, Colorado State University, Fort Collins, CO 80523; E-mail: philk@atmos.colostate.edu.

3. Forecast model development

Predictors were selected from the CFSR and NOAA OI SST datasets. Precursor signals were investigated during the June-July time period, to find areas that had the strongest correlation with NTC activity over the period from 1982-2009 (the overlapping period for both datasets). Only low-level fields were investigated (*e.g.*, sea level pressure, surface zonal wind), as these predictors were deemed to be more reliable during the earlier part of the 20th century, and the intention was to be able to test the skill of these predictors on earlier-period data. In addition, the ECMWF model's forecasts were examined to determine if they showed significant skill in predicting ENSO from 1 July issue date. The ECMWF model's September forecast for the Nino 3 region (5°S-5°N, 150-90°W) was quite impressive, correlating with observations at 0.90 over the period from 1982-2010. Predictors were only added if they explained an additional three percent of the variance from 1982-2009, and strong physical linkages between each predictor and TC activity were required to have been demonstrated. Each predictor was also required to significantly correlate with NTC over the period from 1982-2009, using a one-tailed Student's t-test. When this predictor qualification procedure was employed, a total of three predictors were selected. These three predictors are displayed in Figure 2.

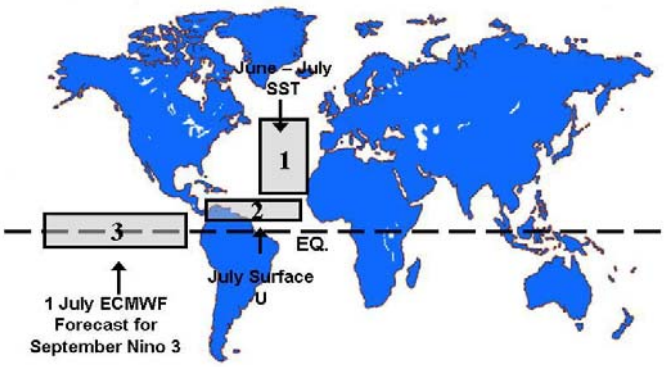


Figure 2 Predictors selected for the new early August statistical forecast model for post-31 July NTC in the Atlantic basin.

Predictor Type	Predictor Location	Linear Correlation with NTC (1982-2010)
June-July SST	20-50°N, 35-15°W	0.67
July 10 meter U	10-17.5°N, 80-40°W	0.83
ECMWF September SST Forecast (Model Initialized 1 July)	5°S-5°N, 150-90°W	-0.49

Table 1 Predictors selected for the post-1 August NTC forecast. Also presented are the linear correlations between each individual predictor and post-1 August NTC.

Table 1 displays each predictor's individual correlation with NTC over the period from 1982-2009. All correlations are statistically significant at the 99% level using a one-tailed Student's t-test and assuming that each year represents an individual degree of freedom.

A full discussion of each predictor's individual relationship with NTC is discussed in Klotzbach (2011). All predictors are closely related to either the Atlantic Warm Pool (AWP) or ENSO, which have both been documented in many previous papers (*e.g.*, Gray 1984, Wang and Lee 2007) to have a significant impact on Atlantic TC activity levels, through alterations in large-scale

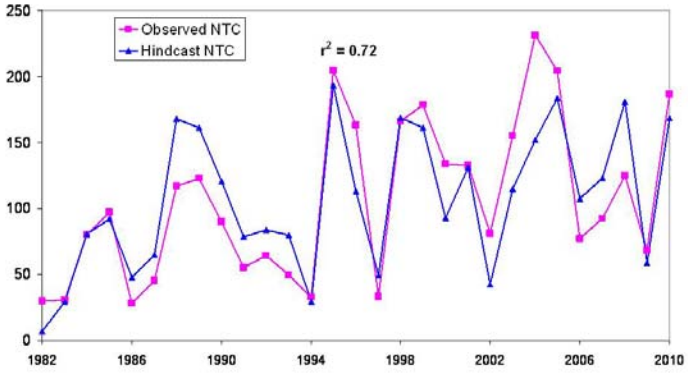


Figure 3 Observed versus post-31 July model jackknifed NTC hindcast over the period from 1982-2010. The three-predictor model explains 72 percent of the variance in post-31 July NTC.

dynamic (*e.g.*, wind shear, low-level vorticity) and thermodynamic (*e.g.*, SST, mid-level humidity) properties.

These three predictors were then combined using linear regression. When they were combined and a drop-one cross validation technique was applied, the linear regression model explained 72 percent of the variability in post-31 July NTC (Figure 3).

4. Earlier period (1900-1981) model verification

The forecast model outlined in the previous section was then examined for similar levels of skill during the earlier part of the 20th century (from 1900-1981). Since the ECMWF forecast model hindcasts are not available prior to around 1980, observed values of the Nino 3 index were used for verification, effectively assuming a perfect ENSO forecast. The ECMWF model correlated with observations at 0.90 over the 1982-2010 period, so assuming a perfect forecast during the earlier part of the 20th century is not too much of a stretch.

Table 2 displays the correlations between each predictor and post-31 July NTC for the 1900-1981 period, as well as the 1900-1947 and 1948-1981 sub-periods, respectively. Correlations are lower than for the 1982-2010 period; however, they remain significant at the 90% level using a one-tailed Student's t-test. In addition, one would expect some degradation in correlation, since both observed large-scale fields (*e.g.*, SLP, SST, low-level wind) as well as TC activity have greater uncertainties associated with them as one goes back further in time.

1900-1981	
Predictor Number (Name)	NTC
1 (Subtropical Atlantic SST)	0.31
2 (Tropical Atlantic U)	0.41
3 (Observed September Nino 3)	-0.32
1900-1947	
Predictor Number (Name)	NTC
1 (Subtropical Atlantic SST)	0.34
2 (Tropical Atlantic U)	0.50
3 (Observed September Nino 3)	-0.46
1948-1981	
Predictor Number (Name)	NTC
1 (Subtropical Atlantic SST)	0.25
2 (Tropical Atlantic U)	0.48
3 (Observed September Nino 3)	-0.25

Table 2 Correlation between predictors and post-1 August NTC over the period from 1900-1981, 1900-1947, and 1948-1981, respectively.

5. New Forecast Model's Improvement upon Klotzbach (2007) Model

This newly-developed forecast model shows modest improvement upon the earlier model developed by Klotzbach (2007). While both models improve significantly upon climatology over the period from 1982-2010, the new model has a smaller mean absolute error than the Klotzbach (2007) model 66% of the time, while also explaining an additional 20% of the variability from climatology (Figure 4).

6. Conclusions and future work

A newly-developed early August statistical forecast model for post-31 July NTC prediction in the Atlantic basin shows significant levels of skill compared against a climatological forecast. The new model utilizes a total of three predictors, which are all closely related to either ENSO or the AWP. The combination of these predictors explains 72% of the variance in cross-validated post-31 July NTC

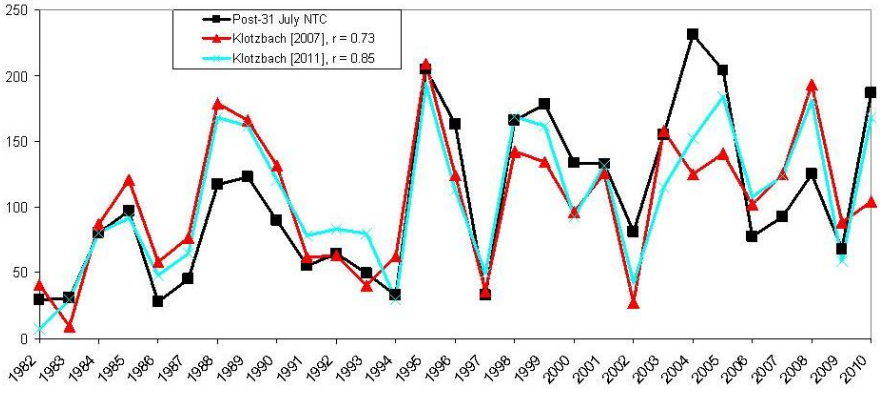


Figure 4 Observed post-31 July NTC (black line), Klotzbach (2007) statistical model forecasts of post-31 July NTC (red line) and Klotzbach (2011) statistical model forecasts of post-31 July NTC (blue line). The Klotzbach (2011) shows improved forecast skill when compared with Klotzbach (2007).

activity.

In the future, additional predictors will be considered including mid-level moisture predictors (such as 500-mb relative humidity). Also, since the ERA-Interim Reanalysis has recently been extended backward to 1979 (Dee *et al.* 2011), this reanalysis product will also be evaluated for forecast development potential.

References

Compo, G. P., and Co-Authors, 2011: The twentieth century reanalysis project. *Q. Jour. Roy. Meteor. Soc.*, **137,** 1-28, doi:10.1002/qj.776.

Dee, D. P., and Co-Authors, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. Jour. Roy. Meteor. Soc.*, **137,** 553-597, doi:10.1002/qj.828.

Gray, W. M., 1984: Atlantic seasonal hurricane frequency. Part II: Forecasting its variability. *Mon. Wea. Rev.*, **112,** 1669-1683.

Jarvinen, B. R., C. J. Neumann, and M. A. S. Davis, 1984: A tropical cyclone data tape for the North Atlantic basin, 1886-1983: Contents, limitations, and uses, 21 pp. *NOAA Tech. Memo.* NWS NHC 22, Miami, FL.

Kistler, R., and Co-Authors, 2001: The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.*, **82,** 247-267.

Klotzbach, P. J., 2011: A simplified Atlantic basin seasonal hurricane prediction scheme from 1 August. *Geophys. Res. Lett.*, **38**, L16710, doi:10.1029/2011GL048603.

___, and W. M. Gray, 2009: Twenty-five years of Atlantic basin seasonal hurricane forecasts. *Geophys. Res. Lett.*, **36**, L09711, doi:10.1029/2009GL037580.

___, 2007: Revised prediction of seasonal Atlantic basin tropical cyclone activity from 1 August. *Wea. Forecasting*, **22**, 937-949.

Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes, and W. Wang, 2002: An improved in situ and satellite SST analysis for climate, *J. Climate*, **15,** 1609-1625.

Saha, S., and Co-Authors, 2010: The NCEP Climate Forecast System Reanalysis, *Bull. Amer. Meteor. Soc.*, **91,** 1015-1057.

Smith, T. M., R. W. Reynolds, T. C. Peterson, and J. Lawrimore, 2008: Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880-2006), *J. Climate*, **21,** 2283-2296.

Stockdale, T. N., D. L. T. Anderson, M. A. Balmaseda, F. Doblas-Reyes, L. Ferranti, K. Mogensen, T. N. Palmer, F. Molteni and F. Vitart, 2011: ECMWF Seasonal Forecast System 3 and its prediction of sea surface temperature, *Clim. Dyn.*, **137,** 553-597, doi: 10.1007/s00382-010-0947-3.

Wang, C., and S-K. Lee, 2007: Atlantic warm pool, Caribbean low-level jet, and their potential impact on Atlantic hurricanes, *Geophys. Res. Lett.*, **34,** L02703, doi:10.1029/2006GL028579.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

CWRF Ready for Climate Service

Xin-Zhong Liang[1] and Julian X.L. Wang[2]

[1]*Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD*
[2]*NOAA Air Resources Laboratory, Silver Spring, MD*

The CWRF has been developed as the Climate *extension* of the Weather Research and Forecasting model (WRF, Skamarock *et al.* 2008) by incorporating numerous improvements in representation of physical processes and integration of external (top, surface, lateral) forcings that are crucial to climate scales, including interactions between land–atmosphere–ocean, convection–microphysics and cloud–aerosol–radiation, and system consistency throughout all process modules (Liang *et al.* 2011). This extension inherits all WRF functionalities for numerical weather prediction while enhancing the capability for climate modeling. As such, it can be applied for seamless weather forecast and climate prediction. The CWRF has been built with an unprecedentedly comprehensive *ensemble* of alternative parameterization schemes for each of the key physical processes, including surface (land, ocean), planetary boundary layer, cumulus (deep, shallow), microphysics, cloud, aerosol, and radiation, and their interactions. This facilitates the use of an optimized physics ensemble approach to improve weather or climate prediction along with a reliable uncertainty estimate. The CWRF also emphasizes the societal *service* capability to provide credible information for climate impacts analyses. For that, it has been coupled with detailed models of terrestrial hydrology, coastal ocean, crop growth, air quality, and recently expanding interactive water quality and ecosystem. Their outputs will form a scientific basis for decision makers to select optimal pathways to achieve economic, societal and environmental goals.

The CWRF improvements have been accomplished through iterative, extensive model refinements, sensitivity experiments, and rigorous evaluations over the past 9 years under close collaborations between the Illinois State Water Survey in the University of Illinois at Urbana-Champaign (2003-2010), the Earth System Science Interdisciplinary Center (ESSIC) in the University of Maryland at College Park (2011 onward), the NOAA Air Resource Laboratory (ARL), and the NOAA Center for Atmospheric Sciences (NCAS). As a result, the CWRF has demonstrated greater capability and better performance in simulating the U.S. regional climate than the existing CMM5 (Liang *et al.* 2004b) and the original WRF. The present study provides an introduction of the CWRF for its application over the U.S., elaborating a few unique features that are relevant to providing credible model results for climate service.

CWRF physical process representations

Figure 1 illustrates the current CWRF physics options and executing structure (see all the abbreviations and acronyms listed after the References). There are seven major drivers that each controls multiple alternative schemes for the physical processes of cloud, aerosol, radiation, surface, PBL, cumulus, and microphysics, in the sequential order of computation. The first three drivers (*cloud, aerosol, radiation*) form the Cloud-Aerosol-Radiation (CAR) ensemble modeling system that incorporates over 10^{18} different ways to simulate interactions among cloud, aerosol and radiation, developed from seven packages available in the leading global and regional models around the world. This replaces the original WRF single *radiation* driver that consists of the CAM and AER packages, and the MISC schemes now obsolete. The *surface* driver manages all schemes handling surface and subsurface processes over land and oceans, as well as surface-atmosphere flux exchanges. In particular, the CWRF adds the advanced CSSP and CROP for terrestrial hydrology and crop growth over land, and SOM and UOM for mixed-layer and upper ocean effects. The two urban schemes are separated from the NOAH and now work with all land surface schemes. All 7 surface layer schemes, originally tied to specific options, are now interchangeable for all *surface* and *PBL* schemes. The *PBL* driver hosts 7 WRF plus 2 new (CAM, UW) PBL schemes, all of which are integrated with the ORO

Correspondence to: Xin-Zhong Liang, Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD; E-mail: xliang@umd.edu.

module accounting for orographic turbulence stress and gravity-wave drag. The *cumulus* driver provides the hub for 7 WRF plus 6 new (GR, ZML, CSU, GFDL, MIT, ECP) deep cumulus schemes, all of which are conjunctive with a shallow convection scheme (UW). A consistent switch is added to control whether shallow convection is activated internally in 8 deep cumulus schemes or done externally by the UW scheme. The *microphysics* driver harnesses 11 microphysics schemes of the WRF.

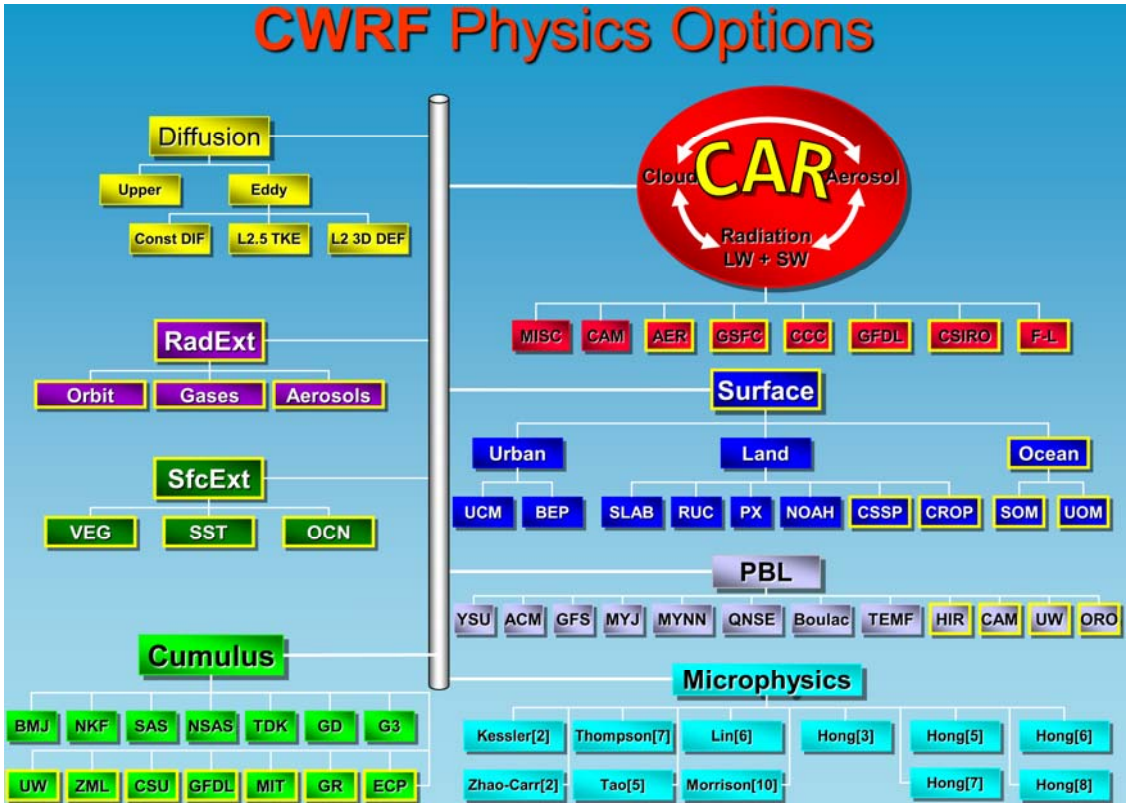


Fig. 1 The schematic of the current CWRF physics options and executing sequence from the top down. The CAR ensemble system and all modules or schemes outlined in yellow are additions specifically developed for the CWRF, while others are inherited from the WRF.

Importantly, we strove to make all alternative schemes in the CWRF fully coupled across all drivers with plug-and-play interfaces. Even without counting the grand CAR ensemble, the CWRF currently contains over 10^6 configurations modeling the surface, PBL, cumulus, and microphysics processes and their interactions. To achieve this, substantial efforts have been made to scrutinize all individual schemes for consistency and incorporate suitable algorithms for missing variables to enable the overall system coupling. Particular care has been taken to ensure continuous model integration that can be restarted at any interval while resulting in bit-by-bit agreement. This is not trivial, especially if time intervals differ among executing individual physics drivers. A seamless averaging procedure is implemented to replace cumulative variables, while pertaining to model prediction, by their averages between two consecutive steps of the driver at work. This is especially effective for precipitation fields (convective/resolved rainfall/snowfall) that are used for different purposes in the *cumulus*, *microphysics*, *surface*, *cloud*, and *aerosol* drivers. Other cumulative variables for diagnostic outputs, such as surface water and energy budget fields, can be set to zero at any restart check point to reduce truncation errors. As such, the CWRF can be run safely for a long-term climate simulation with frequent restarts as needed and with varying time steps for all 7 physics drivers. On the other hand, the WRF[1] with

[1] Note that the WRF can be configured to many versions using different combinations of physics schemes. The reported WRF configurations are limited. The statement was drawn upon our own experience with the WRF runs and through review of several journal manuscripts of others.

several tested configurations has been reported to result in numerical instability or serious drift that are prohibitive for continuous climate simulations.

CWRF advanced terrestrial hydrology prediction

The CWRF incorporates a Conjunctive Surface-Subsurface Process model (CSSP) to predict soil temperature/moisture distributions, terrestrial hydrology variations, and land-atmosphere flux exchanges. The CSSP is rooted in the Common Land Model (CoLM; Dai *et al.* 2003, 2004) with a few updates from the Community Land Model version 3.5 (CLM3.5; Oleson *et al.* 2008). It is built upon realistic distributions of surface (soil and vegetation) characteristics (Liang *et al.* 2005a,b), and with significant improvements in representing surface energy and hydrology processes. These include an improved dynamic-statistical parameterization of land surface albedo (Liang *et al.* 2005c); a 3-D subsurface hydrologic model with a scalable representation of subgrid topographic control on soil moisture (Choi *et al.* 2007); an explicit treatment of surface-subsurface flow interaction (Choi 2006; Choi and Liang 2010; Choi *et al.* 2011); an unconfined aquifer below the bedrock (Yuan and Liang 2011a). The CSSP integrates vertical water exchange (precipitation, evaporation, transpiration, infiltration) and horizontal water movement (across grids) to predict surface and subsurface runoff resulting from rainfall excess, saturation depletion and lateral flows due to resolved and subgrid topographic controls.

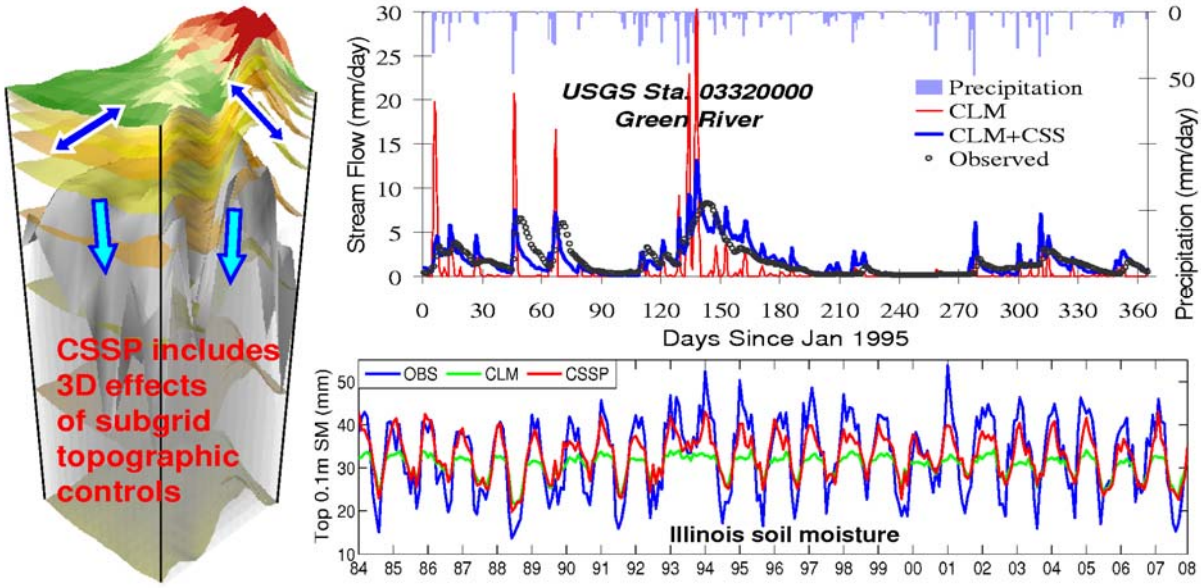


Fig. 2 The CSSP improves terrestrial hydrology prediction over the CLM. This includes incorporation of 3D effects of subgrid topographic controls (*left*), depiction of realistic streamflow variations over major watersheds (*right top*), and capture of seasonal-interannual variations of soil moisture observed in Illinois (*right bottom*).

A comprehensive evaluation against observations at regional-local scales over the contiguous U.S. has demonstrated that the CSSP overall performance is superior to both the CoLM and CLM3.5 (Yuan and Liang 2011a). A recent comparison of offline integrations driven by observational reanalysis data also revealed that the CSSP has clear advantages in modeling the U.S. terrestrial hydrology (soil moisture, runoff) over the NOAH used in the NCEP CFS. As a result, the CWRF using the CSSP generates not only more realistic phase (higher correlations) but also better amplitude (deviation ratios closer to 1) of the soil moisture seasonal-interannual variations throughout the root zone than the WRF using the NOAH (Liang *et al.* 2011). Figure 2 depicts an example CSSP improvement over its origin CLM, which produces streamflow pulse fluctuations as a result of quick response to rainfall events, causing no recession time, overall runoff underestimation, and weak seasonal-interannual soil moisture variability. This advance in representing the terrestrial hydrology by the CSSP over NOAH has other major climate benefits.

CWRF physics ensemble skill enhancement

The CWRF incorporates a massive suite of alternative numerical schemes for microphysics, convection, cloud, aerosol, radiation, surface, turbulence and transport processes, all of which are fully coupled with nonlinear interactions to ultimately determine its climate prediction. No single scheme can capture all aspects of the observed climate, but produce predictive skill dependent of climate regimes (Liang *et al.* 2004a,b; Mapes *et al.* 2004). Consensus prediction based on the ensemble of multiple models or multiple physics configurations of a model may offer significant skill enhancement (Krishnamurti *et al.* 1999; Peng *et al.* 2002; Palmer *et al.* 2004; Liang *et al.* 2007; Wang *et al.* 2009).

Here we tested an extremely limited subset of the CWRF full ensemble, focusing on the control configuration and all major alternative schemes across each physics driver, altered one at a time. In total, there are 26 CWRF simulations. Each simulation is driven by the NCEP Reanalysis-2 LBCs and integrated from 1 November 1992 to 31 December 1993. During the summer of that year, record flooding occurred in the Mississippi River basin. Figure 3 illustrates spatial frequency distributions of pointwise correlation coefficients and root-mean-square errors (*RMSE*) of daily mean rainfall variations between observations and all CWRF simulations. Shown are also the ensemble results as the averages of all runs with equal or optimal weights. The optimal weight results from local *RMSE* minimization, and the skill score depicts the upper limit of daily rainfall predictability that can be achieved from the best optimization of the ensemble. Clearly, the ensemble average of the alternative physics configurations using an equal weight substantially increases the predictive skill over all individuals, with more frequent occurrences of higher correlation coefficients and smaller *RMSE*. The improvement by the ensemble is realized because distinct regions are identified where each configuration complementarily captures certain but not all observed signals. The skill enhancement is most pronounced in summer, followed by autumn and spring, whereas rather weak in winter (Liang *et al.* 2011). Note that the ensemble average using the localized optimal weights has predictive skill significantly higher than that using the equal weight as well as the individuals throughout the entire year. Thus, there exists substantial room to further enhance that skill through intelligent optimization.

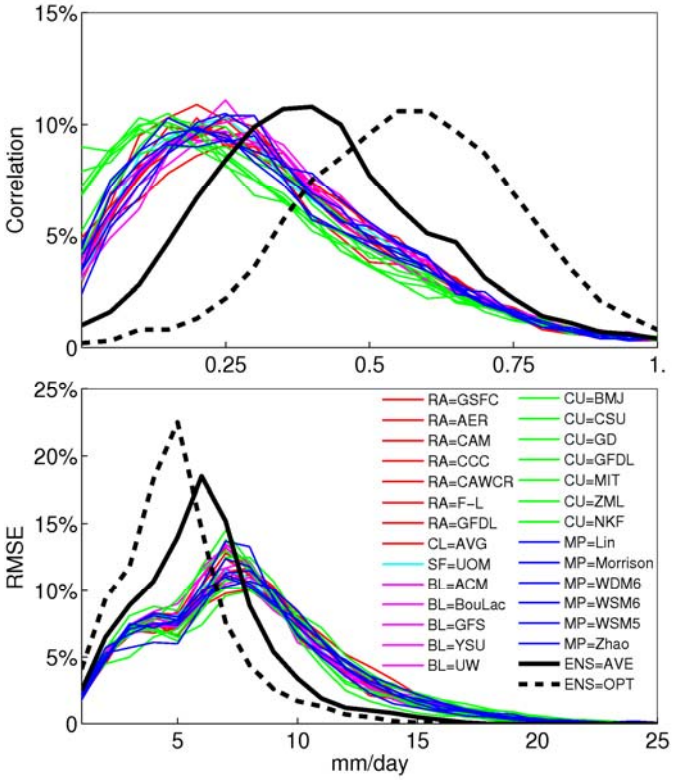


Fig. 3 Spatial frequency distributions of correlations (top) and rms errors (bottom) between CWRF and observed daily mean rainfall variations in summer 1993. Each line depicts a specific configuration in group of the five key physical processes (color). The ensemble result (ENS) is the average of all runs with equal (Ave) or optimal (OPT) weights, shown as black solid or dashed line.

CWRF downscaling improvement to CFS climate prediction

We have recently demonstrated that the mesoscale CWRF downscaling produces significant skill enhancement to the driving NCEP CFS seasonal forecast for winter precipitation during 1982-2008 (Yuan and Liang 2011b). Figure 4a compares spatial frequency distributions of *RMSE* of seasonal mean precipitation interannual variations predicted by the CFS and downscaled by the CWRF. The statistics are based on all land grids over the entire inner domain (U.S., southern Canada, and northern Mexico) from 5 realizations. All CWRF results consistently reduce CFS forecast errors. The reduction is obvious at all

forecast lead times, with the *RMSE* peaks decreased by about 0.5 mm/day. Figure 4b illustrates the CWRF minus CFS differences in equitable threat score (*ETS*) of seasonal mean precipitation forecasts. The CFS forecast skill decreases rapidly for heavy rainfall events, while the CWRF maintains a good level across the range. On average, the CWRF reduces CFS forecast *RMSE* by 22%, increases *ETS* by 0.08-0.15, and produces greater skill for heavy rainfall events.

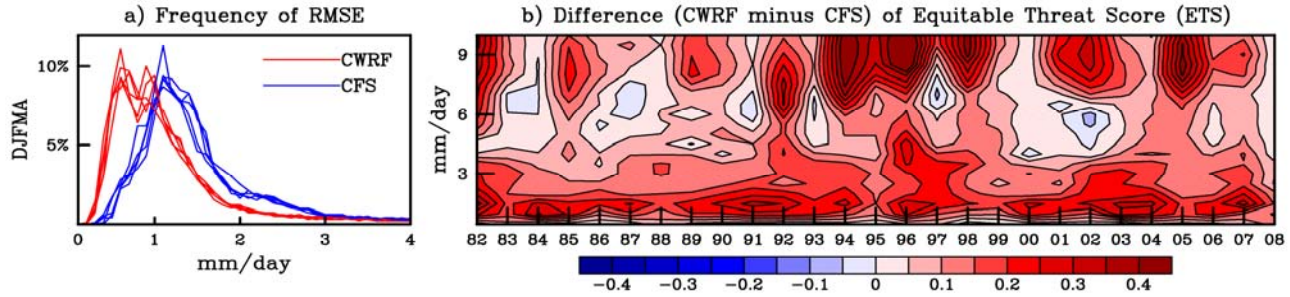


Fig. 4 a) Spatial frequency distributions of *RMSE* (mm/day) predicted by the CFS and downscaled by the CWRF and b) CWRF minus CFS differences in *ETS* for seasonal mean precipitation interannual variations. The statistics are based on all land grids over the entire inner domain for DJFMA from the 5 realizations during 1982-2008.

Note that the *ETS* differences are larger in ENSO-neutral years than in strong anomalous years. For instance, smaller enhancements are identified in years with La Niña (1984, 1988) and El Niño (1986, 1991, 2002). During these abnormal years, significant ENSO signals presented in the planetary circulation, and thereby the CFS has higher seasonal climate predictability, especially for wintertime when global anomalies are more intense. As a result, the advantage of the CWRF downscaling over the CFS forecast is relatively weaker than ENSO-neutral years. A further analysis (not shown) indicates that the CWRF simulates more accurate number of rainy days than the CFS over the northern and western U.S. due to the refined representation of orographic effect, shallow convection, and terrestrial hydrology, and also more realistically captures the broad region of extreme rainfall over the Gulf States and maximum dry spell length along the Great Plains, as well as their contrasts between El Niño and La Niña events. In conclusion, the CWRF downscaling exhibits significant advantages for regional precipitation prediction, especially during years with weak planetary anomalies.

CWRF application for climate service at regional-local scales

The CWRF has been coupled with detailed models of terrestrial hydrology, coastal ocean, crop growth, air quality, and recently expanding interactive water quality and ecosystem. As such, the CWRF has been designed for climate applications at regional-local scales, and can be used to translate GCM global climate simulations into regional-specific actionable information for local impacts. This can be done by nesting CWRF with selected (*e.g.*, NOAA, NASA) operational forecasts of seasonal-interannual climate anomalies, and CMIP (*e.g.*, NCAR, GFDL) projections of future decadal climate changes under a feasible range of emissions scenarios. The CWRF can be run at multiple nested grids finer than 30-km to resolve the synoptic, mesoscale and local processes that govern the climate and environmental anomalies and changes most relevant to end-users for decision making. In so doing, the CWRF will be able to integrate the global signals with regional characteristics into comprehensive information required for local impacts assessment. We can further constrain the CWRF by the advanced data assimilation to improve initialization and narrow uncertainty such that the final prediction will be the most reliable source for end-users. We anticipate that the CWRF ensemble of multiple alternative physics configurations, with optimal weights on individual members as constrained by their respective performance metrics against observations, will further increase the downscaling predictive skill over the driving GCM forecasts or projections with more reliable estimate of result uncertainty. The CWRF optimized physics ensemble downscaling approach will provide an unprecedented skill enhancement for predicting climate at regional-local scales.

References

Choi, H.I., 2006: 3-D volume averaged soil-moisture transport model: A scalable scheme for representing subgrid topographic control in land-atmosphere interactions. *Ph.D. Dissertation*, University of Illinois at Urbana-Champaign, 189 pp.

Choi, H.I., P. Kumar, and X.-Z. Liang, 2007: Three-dimensional volume-averaged soil moisture transport model with a scalable parameterization of subgrid topographic variability. *Water Resour. Res.*, **43**, W04414, doi:10.1029/2006WR005134, 15pp.

Choi, H.I., X.-Z. Liang, and P. Kumar, 2011: A conjunctive surface-subsurface flow representation for mesoscale land surface models. *Water Resour. Res*. (submitted).

Choi, H.I., and X.-Z. Liang, 2010: Improved terrestrial hydrologic representation in mesoscale land surface models. *J. Hydrometeorology*, **11**, 797–809.

Dai, Y., R.E. Dickinson, and Y.-P. Wang, 2004: A two-big-leaf model for canopy temperature, photosynthesis, and stomatal conductance. *J. Climate*, **17**, 2281-2299.

Dai, Y., X. Zeng, R.E. Dickinson, I. Baker, G.B. Bonan, M.G. Bosilovich, A.S. Denning, P.A. Dirmeyer, P.R. Houser, G. Niu, K.W. Oleson, C.A. Schlosser, and Z-L. Yang, 2003: The Common Land Model. *Bull. Amer. Meteor. Soc*., **84**, 1013-1023.

Krishnamurti, T. N., C. M. Kishtawal, T. E. LaRow, D. R. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil, and S. Surendran, 1999: Improved weather and seasonal climate forecasts from multimodel superensemble. *Science*, **285**, 1548-1550.

Liang, X.-Z., H. Choi, K.E. Kunkel, Y. Dai, E. Joseph, J.X.L. Wang, and P. Kumar, 2005a: Development of the regional climate-weather research and forecasting model (CWRF): Surface boundary conditions. *Illinois State Water Survey Scientific Research*, ISWS SR 2005-01, 32 pp. (http://www.sws.uiuc.edu/pubs/pubdetail.asp? CallNumber=ISWS+SR+2005%2D01).

Liang, X.-Z., H. Choi, K.E. Kunkel, Y. Dai, E. Joseph, J.X.L. Wang, and P. Kumar, 2005b: Surface boundary conditions for mesoscale regional climate models. *Earth Interactions*, **9**, 1–28.

Liang, X.-Z., L. Li, A. Dai, and K.E. Kunkel, 2004a: Regional climate model simulation of summer precipitation diurnal cycle over the United States. *Geophys. Res. Lett*., **31**, L24208, doi:10.1029/2004GL021054.

Liang, X.-Z., L. Li, K.E. Kunkel, M. Ting, and J.X.L. Wang, 2004b: Regional climate model simulation of U.S. precipitation during 1982-2002. Part 1: Annual cycle. *J. Climate*, **17**, 3510–3528.

Liang, X.-Z., M. Xu, W. Gao, K.E. Kunkel, J. Slusser, Y. Dai, Q. Min, P.R. Houser, M. Rodell, C.B. Schaaf, and F. Gao, 2005c: Development of land surface albedo parameterization bases on Moderate Resolution Imaging Spectroradiometer (MODIS) data. *J. Geophys. Res*., **110**, D11107, doi:10.1029/2004JD005579.

Liang, X.-Z., M. Xu, K.E. Kunkel, G.A. Grell, and J. Kain, 2007: Regional climate model simulation of U.S.-Mexico summer precipitation using the optimal ensemble of two cumulus parameterizations. *J. Climate*, **20**, 5201–5207.

Liang, X.-Z., M. Xu, X. Yuan, T. Ling, H.I. Choi, F. Zhang, L. Chen, S. Liu, S. Su, F. Qiao, J.X.L. Wang, K.E. Kunkel, W. Gao, E. Joseph, V. Morris, T.-W. Yu, J. Dudhia, and J. Michalakes, 2011: Regional Climate-Weather Research and Forecasting Model (CWRF). *Bull. Amer. Meteor. Soc.* (submitted).

Mapes, B.E., T.T. Warner, M. Xu, and D.J. Gochis, 2004: Comparison of cumulus parameterizations and entrainment using domain-mean wind divergence in a regional model. *J. Atmos. Sc.*, **61**, 1284-1295.

Oleson, K. W., G.-Y. Niu, Z.-L. Yang, D. M. Lawrence, P. E. Thornton, P. J. Lawrence, R. Stöckli, R. E. Dickinson, G. B. Bonan, S. Levis, A. Dai and T. Qian, 2008: Improvements to the Community Land Model and their impact on the hydrological cycle. *J. Geophys. Res.*, **113**, G01021, doi:10.1029/2007JG000563.

Palmer, T. N., A. Alessandri, U. Andersen, P. Cantelaube, M. Davey, P. Délécluse, M. Déqué, E. Díez, F. J. Doblas-Reyes, H. Feddersen, R. Graham, S. Gualdi, J.-F. Guérémy, R. Hagedorn, M. Hoshen, N. Keenlyside, M. Latif, A. Lazar, E. Maisonnave, V. Marletto, A. P. Morse, B. Orfila, P. Rogel, J.-M.

Terres, and M. C. Thomson, 2004: Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER)**.** *Bull. Amer. Meteorol. Soc*., **85**, 853–872.

Peng, P., A. Kumar, H. van den Dool, and A. G. Barnston, 2002: An analysis of multimodel ensemble predictions for seasonal climate anomalies, *J. Geophys. Res.*, **107**(D23), 4710, doi:10.1029/2002JD002712.

Skamarock, W.C., J.B. Klemp, J. Dudhia, D.O. Gill, D.M. Barker, M.G. Duda, X.-Y. Huang, W. Wang, and J.G. Powers, 2008: *A Description of the Advanced Research WRF Version 3*. NCAR Technical Note, NCAR/TN-475+STR, 113 pp.

Wang, B., J.-Y. Lee, I.-S. Kang, J. Shukla, C.-K. Park, A. Kumar, J. Schemm, S. Cocke, J.-S. Kug, J.-J. Luo, T. Zhou, B. Wang, X. Fu, W.-T. Yun, O. Alves, E.K. Jin, J. Kinter, B. Kirtman, T. Krishnamurti, N.C. Lau, W. Lau, P. Liu, P. Pegion, T. Rosati, S. Schubert, W. Stern, M. Suarez, and T. Yamagata, 2009: Advance and prospectus of seasonal prediction: assessment of the APCC/CliPAS 14-model ensemble retrospective seasonal prediction (1980-2004). *Clim. Dyn.*, **33**, doi: 10.1007/s00382-008-0460-0.

Yuan, X., and X.-Z. Liang, 2011a: Evaluation of a Conjunctive Surface-Subsurface Process model (CSSP) over the contiguous U.S. at regional-local scales. *J. Hydrometeorol*, **12**, 579-599, doi:10.1175/2010JHM1302.1.

Yuan, X., and X.-Z. Liang, 2011b: Improving cold season precipitation prediction by the nested CWRF-CFS system. *Geophys. Res. Lett.*, **38**, L02706, doi:10.1029/2010GL046104.

ABBREVIATIONS AND ACRONYMS

ACM	Asymmetric Convective Model
AER	Atmospheric and Environmental Research
ARL	NOAA Air Resource Laboratory
BEP	Building Environment Parameterization (multilevel urban model)
BMJ	Betts-Miller-Janjic cumulus parameterization
BouLac	Bougeault-Lacarrère PBL scheme
CAM	NCAR Community Atmosphere Model
CAR	CWRF Cloud-Aerosol-Radiation Ensemble Modeling System
CAWCR	Centre for Australia Weather and Climate Research
CCCMA	Canadian Centre for Climate Modeling and Analysis
CFS	NCEP Climate Forecast System
CLM3.5	Community Land Model version 3.5
CMM5	Climate Extension of the PSU/NCAR Mesoscale Model generation 5
CoLM	Common Land Model
CMIP	Coupled Model Intercomparison Project
CROP	Dynamic crop growth modeling system
CSSP	Conjunctive Surface-Subsurface Process model
CSU	Colorado State University
CWRF	Climate extension of the Weather Research and Forecasting model
ECP	Ensemble Cumulus Parameterization modified from G3
ENSO	El Niño-Southern Oscillation
ESSIC	Earth System Science Interdisciplinary Center, University of Maryland
ETS	Equitable Threat Score
FLG	Fu-Liou-Gu radiation transfer scheme
G3	Grell-3 ensemble cumulus parameterization
GCM	General Circulation Model
GD	Grell-Dvénéyi ensemble cumulus parameterization
GFDL	Geophysical Fluid Dynamics Laboratory
GR	Grell cumulus parameterization
GSFC	NASA Goddard Space Flight Center
HIR	High Resolution PBL scheme

LBCs	Lateral Boundary Conditions
Lin	Lin et al. microphysics scheme
MISC	Miscellaneous (obsolete) radiation schemes
MIT	Massachusetts Institute of Technology
Morrison	Morrison et al. two-moment microphysics scheme
MYJ	Mellor-Yamada-Janjic PBL scheme
MYNN	Mellor-Yamada PBL scheme modified by Nakanishi-Niino
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCAS	NOAA Center for Atmospheric Sciences
NCEP	National Centers for Environmental Prediction
NKF	New Kain-Fritsch cumulus parameterization
NOAA	National Oceanic and Atmospheric Administration
NOAH	NCAR-NCEP unified land surface model
ORO	Module for orographic turbulence stress and gravity-wave drag
PBL	Planetary Boundary Layer
PX	Pleim-Xiu land surface model
QNSE	Quasi-Normal Scale Elimination PBL scheme
RadExt	CWRF module for external radiative conditions (solar constant, atmospheric gas volume mixing ratios, aerosol distributions)
RCM	Regional Climate Model
RMSE	Root Mean Square Errors
SAS	Simplified Arakawa-Schubert cumulus parameterization
SBCs	Surface Boundary Conditions
SfcExt	CWRF module for external surface and subsurface conditions
SOM	Simple Ocean Model
SST	Sea Surface Temperature
Tao	Tao et al. microphysics scheme
TEMF	Total Energy–Mass Flux boundary layer scheme (Angevine et al. 2010)
Thompson	Thompson et al. microphysics scheme
UCM	Urban Canopy Model
UOM	Multilevel Upper Ocean Model
UW	University of Washington
WDM6	WRF Double-Moment 6-class microphysics scheme
WSM5	WRF Single-Moment 5-class microphysics scheme
WSM6	WRF Single-Moment 6-class microphysics scheme
WRF	Weather Research and Forecasting model
YSU	Yonsei University
ZML	Zhang-McFarlane-Liang cumulus parameterization

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

A Verification Framework for Interannual-to-Decadal Prediction Experiments

Lisa Goddard[+], Paula Gonzalez[+], Simon Mason[+], Arthur Greene[+],
and the US CLIVAR Working Group on Decadal Predictability

[+]*International Research Institute for Climate & Society (IRI), Columbia University*

1. Introduction: motivations and overview of the work

This work resulted from the efforts of the US CLIVAR Working Group on Decadal Prediction (DPWG) (http://www.usclivar.org/wgdp.php). The authors listed explicitly on this paper are scientists at IRI that participated in the working group. This limited lifetime working group addressed two objectives during our 2009-2011 tenure. The first focused on methodologies that attempt to separate natural from forced climate changes (Solomon *et al.* 2011). The second objective was aimed at the validation and verification of dynamical decadal hindcasts through a set of targeted metrics. The work presented here represents an illustration of the resulting hindcast verification work (see Goddard *et al.* 2012 for more complete details).

Verification of forecasts is needed not only to indicate their expected skill for those who may wish to apply them. Verification also allows improvements in prediction systems to be tracked over time, and allows for comparison across different prediction systems and/or different prediction approaches. The value of a verification framework is provision of guidelines for a common format of the results across hindcasts from different prediction centers, such as common observational data for verification, common period over which the hindcasts are verified and common period against which anomalies are calculated, common metrics, and even common graphical presentation. Through the efforts of our working group, these results are being collected on a central website: http://dpwg-clivar.iri.columbia.edu).

2. Data and methodology/experimental design

2a. Data

The dynamical decadal hindcasts presented here include: the perturbed physics experiments from Hadley Centre using an updated version of the DePreSys forecast system (Smith *et al.* 2010) and the hindcasts from the Canadian Climate Centre using CanCM4 (Arora *et al.* 2011, Merryfield *et al.* 2011). The verification dates are those dictated by the CMIP5 experimental design (Taylor *et al.* 2011), which contain 10 sets of hindcasts – initialized near the end of 1960, 1965, 1970, 1975, 1980, 1985, 1990, 1995, 2000, and 2005. The two hindcast datasets both include more years than this nominal set, and it should be noted that skill metrics generally indicate better performance when the more complete set of hindcasts is used. However, since the common denominator for the CMIP5 hindcasts are these start dates, this is what is shown. Visit the DPWG verification site for more complete information.

The verification is performed for air temperature and precipitation, with the following observational data sets: (1) Air temperature: Hadley Centre/Climate Research Unit Temperature version 3 variance‐adjusted (HadCRUT3v; available on a 5°x5° grid). Preference is given to the HadCRUT3v data because missing data is indicated as such. This can make verification over time more difficult, but it also provides a more realistic view of where forecasts can be verified with gridded data. (2) Precipitation: Global Precipitation Climatology Centre version 4 (GPCCv4). This dataset covers the period 1901-2007, at a resolution of 2.5°x2.5° grid, although the data is provided also at higher resolutions.

The model data are first interpolated to the resolution of the observations prior to the calculation of verification metrics. Thus, the "grid-scale" analysis shown in the results is done at a resolution of 5x5 degrees resolution for temperature and 2.5x2.5 degrees resolution for precipitation. We also do the verification analysis on a smoothed version of the model and observational data. A balance between skill improvement

Correspondence to: Lisa Goddard, International Research Institute for Climate & Society (IRI), Columbia University, Palisades, NY; E-mail: goddard@iri.columbia.edu.

signal-to-noise retention suggests that 5º latitude x 5º longitude represents a reasonable scale for smoothing precipitation, and 10º latitude x 10º longitude for temperature (Räisänen and Ylhäisi 2011). At these scales, grid-scale noise is reduced while retaining the strength of the climate signal and increasing the skill of the verification. Both grid-scale and spatially-smoothed verification contain useful information, and in combination provide guidance on the robustness of signals on different spatial scales.

2b. Methodology: Verification metrics

Verification metrics are chosen to answer specific questions regarding the quality of the forecast information. Our questions address the accuracy in the forecast information (Q1) and the representativeness of the forecast ensembles to indicate forecast uncertainty (Q2). Specifically, the questions are:

Q1: Do the initial conditions in the hindcasts lead to more accurate predictions of the climate? If so, on what time scales?

Q2: Is the model's ensemble spread an appropriate representation of forecast uncertainty on average?

In both cases skill scores are used, which allow for comparison against a chosen baseline. The first question (Q1) is assessed using the mean squared skill score (MSSS, Murphy 1988). When the baseline is the climatological average (common baseline in seasonal-to-interannual predictions), the MSSS can be decomposed into the square of the correlation and the square of the conditional bias. To answer Q1, the baseline skill is defined by the uninitialized climate change projections from the same model as the initialized hindcasts. The MSSS thus quantifies the improvement in the mean squared error (MSE) for the initialized versus uninitialized hindcasts.

To address Q2, we use a measure of probabilistic quality: the continuous ranked probability skill score (CRPSS). The CRPSS is based on the continuous ranked probability score, analogous to the relationship between MSSS and MSE. The CRPS is a measure of squared error in probability space. A continuous score is preferable to a categorical score in the context of a non-stationary climate, where trends may lead to a chronic forecast of, say above-normal temperatures, and offer little discrimination among forecasts, particularly the relative risk of attaining or exceeding some threshold. Following Q2, we assess whether a model's average ensemble spread is suitable for quantifying forecast uncertainty compared to the standard error of the mean forecast, once corrected for conditional bias.

2c. Methodology: Statistical significance

Statistical significance must be estimated, and is of particular importance when sampling uncertainty is likely to be large, such in this case with only 10 hindcasts over a 45-year period for phenomena with a timescale of 20+ years. There is no single way to assess significance. The DPWG verification framework uses a non-parametric bootstrap approach that takes serial autocorrelation into account. (see Goddard *et al.* 2012 for more details)

3. Results

Deterministic (Q1):

The MSSS comparing the initialized and uninitialized temperature hindcasts is shown at the top of Figure 1 based on the spatially smoothed data for the forecast target of the average over years 2-9, or equivalently a 1-year lead-time for a decadal-average prediction. Additionally, the MSSS for each of those hindcasts relative to a climatology baseline is shown in the lower panels. For both prediction systems, the MSSS for temperature from the initialized predictions and the uninitialized projections show positive values over much of the map, suggesting that the trend plays an important role in the MSSS when using a climatological reference forecast. Most of the places where the MSSS is worse (negative or blue areas in the figure) than the reference forecast of climatology (Figure 1, middle and bottom row) is where the temperature trend has been weak or negative. However, many of the regions of negative MSSS referenced against climatology is where the conditional bias (Figures 2&3, right), is large; these are areas where the strength of the model response is too large compared to the observations for a given correlation.

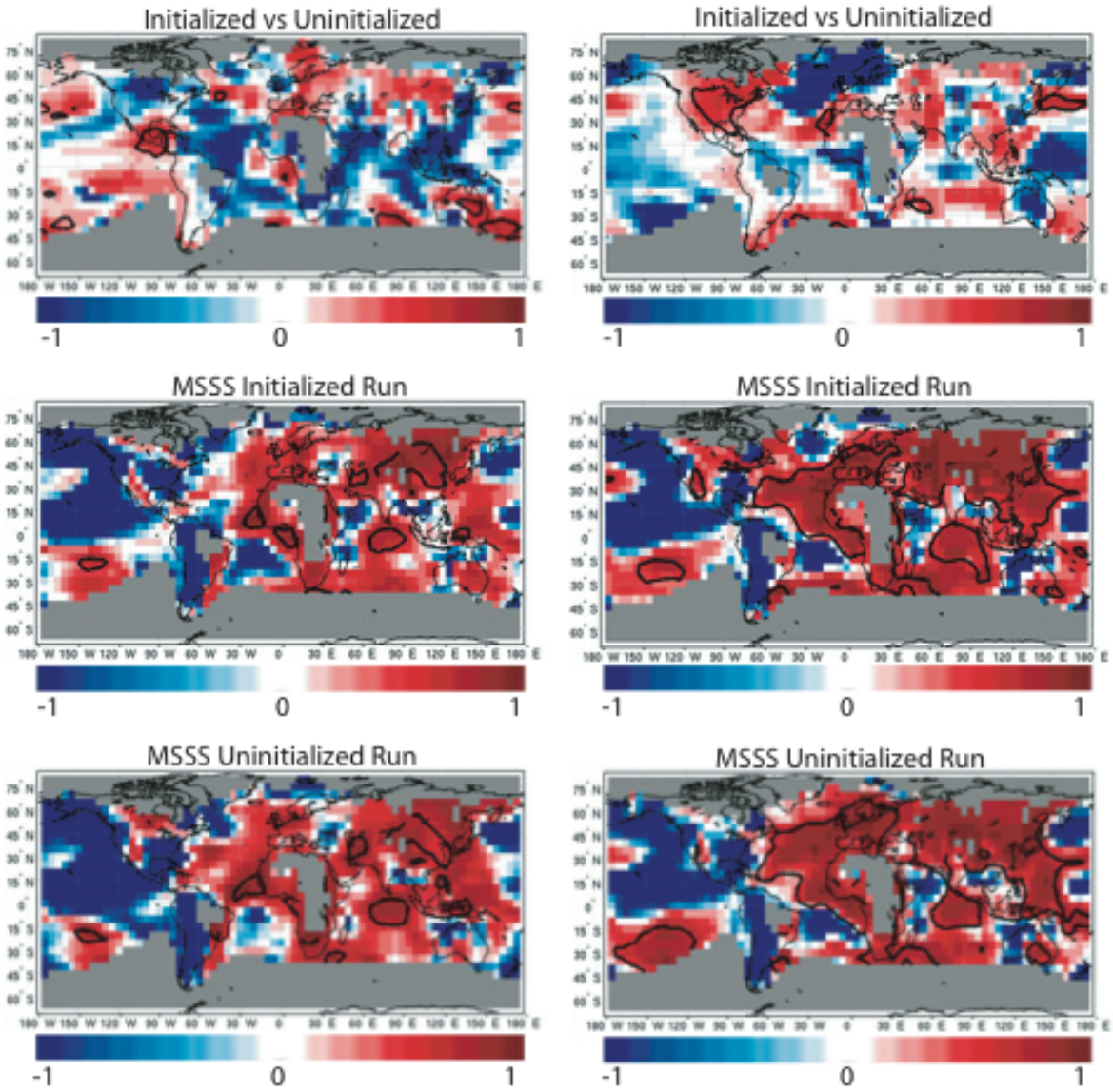


Fig. 1 Mean squared skill score (MSSS) for decadal temperature hindcasts from the DePreSys prediction system of the Hadley Centre (left) and the CanCM4 prediction system of the Canadian Climate Centre (right). Top row: MSSS comparing the initialized hindcasts ("forecasts") and the uninitialized hindcasts ("reference") as predictions of the observed climate; middle row: MSSS comparing the initialized hindcasts ("forecasts") and the climatological mean ("reference"); bottom: MSSS between the unitialized hindcasts ("forecasts") and the climatological mean ("reference"). Observed and model data has been smoothed as described in text. The forecast target is year 2-9 following the initialization every 5 years from 1961-2006 (*i.e.* 10 hindcasts). Contour line indicates statistical significance that the MSSS is positive at the 95% confidence level.

The MSSS of the initialized hindcasts referenced to the uninitialized ones shows that areas of improved skill due to initialization differ between the two models (Figures 1, top panels, the positive or red areas). For example, over the Atlantic the initialized DePreSys hindcasts for temperature improve over the uninitialized hindcasts in the North Atlantic, whereas in the CanCM4 temperature hindcasts the improvement is seen in the tropical Atlantic. That different prediction systems differ in where they are skillful is a common situation in seasonal-to-interannual prediction. It should also be noted that in the case of the Atlantic neither of these improvements are deemed statistically significant, which is shown by the heavy contour line enclosing the

positive skill areas. For the CMIP5 experimental design, very few cases are available and thus it cannot be ruled out that some differences of skill between the two systems result from sampling variability.

The MSSS for the precipitation hindcasts (not shown) are not significantly better than the reference forecast of climatology, anywhere. There are regions where the MSSS of the initialized hindcasts are significantly better than the uninitialized ones, but these areas are small, and though point-wise significant may still be related to the small sample size. Even in regions where improvement between the initialized and uninitialized hindcasts is seen, this improvement must be viewed together with the actual skill (*i.e.* relative to climatology) from the initialized hindcasts.

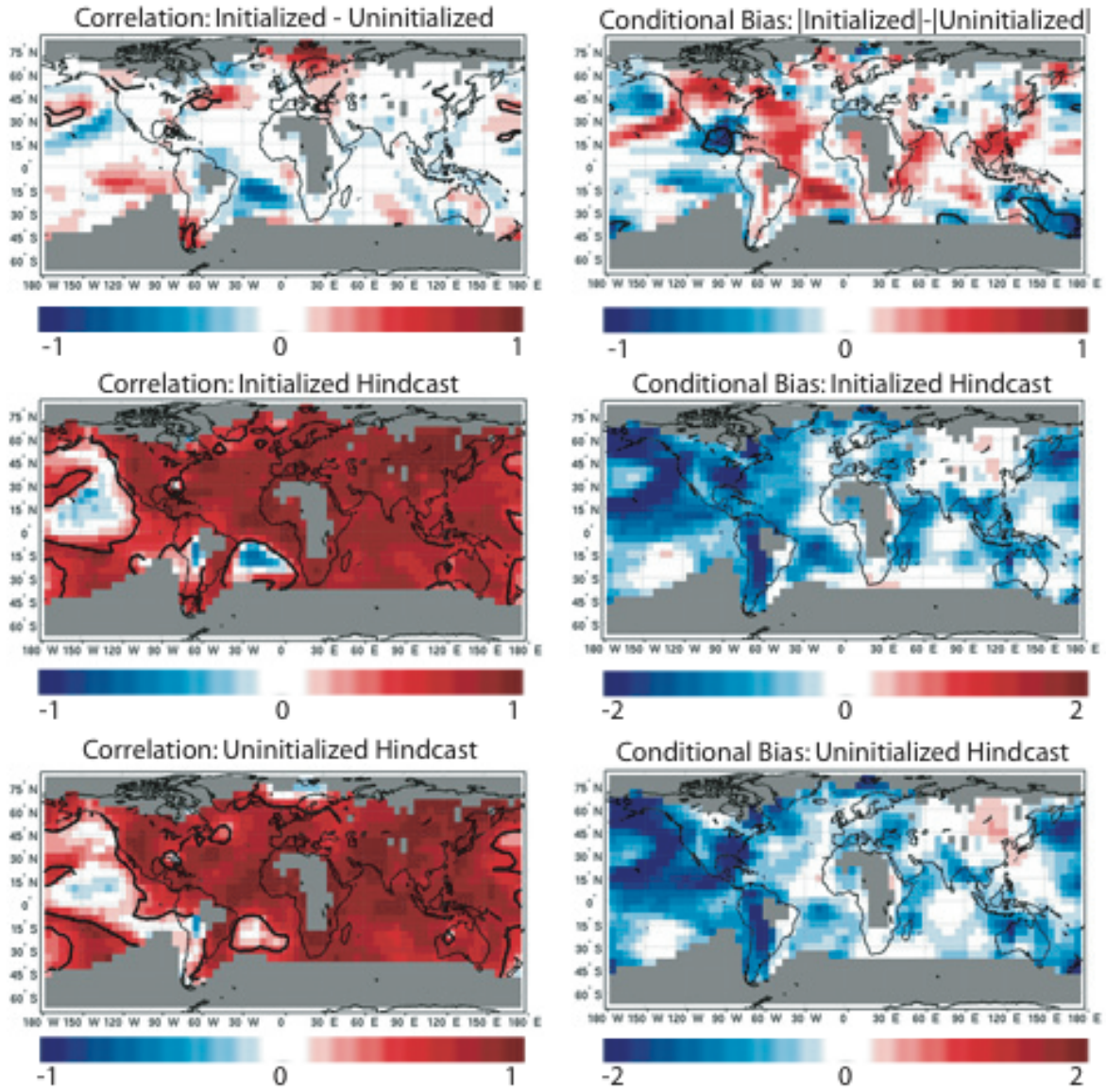


Fig. 2 Skill metrics related to MSSS decomposition for DePreSys temperature hindcasts. Left: Anomaly correlation coefficients with top row depicting the difference between the correlation of the initialized hindcasts (middle row) and that of the uninitialized hindcasts (bottom). Right: Conditional bias, with top row depicting the decrease in magnitude of conditional bias between the initialized hindcasts (middle) relative to that of the uninitialize hindcasts (bottom). Observed and model data has been smoothed as described in text. The forecast target is year 2-9 following the initialization every 5 years from 1961-2006 (*i.e.* 10 hindcasts). Contour line on the correlation maps indicates statistical significance that the value is positive at the 95% confidence level.

For temperature the areas where MSSS shows greater accuracy in the initialized hindcasts compared to the uninitialized (red areas, top of Fig. 1) comes from a reduction in conditional bias (blue areas, Figs. 2&3 right) rather than an increase in correlation (red areas, Figs. 2&3 left). This suggests that at least for this forecast target and in these prediction systems, increased accuracy is not due to the capture of signals in climate variability. For precipitation, improvements may come from both increased correlation and decreased bias, though again the improvements are not typically associated with forecasts that are skillful in their own right.

Probabilistic (Q2):

The CRPSS of the temperature hindcasts show very similar patterns (not shown) whether one estimates the uncertainty in a given forecast from the average ensemble spread or the standard error of the mean. Negative CRPSS values dominate the comparative metric that tests the uncertainty from the ensemble members against the uncertainty from the standard error. This indicates that the use of the ensemble spread leads to less reliable forecasts. The CRPSS for the precipitation forecasts yields very similar results.

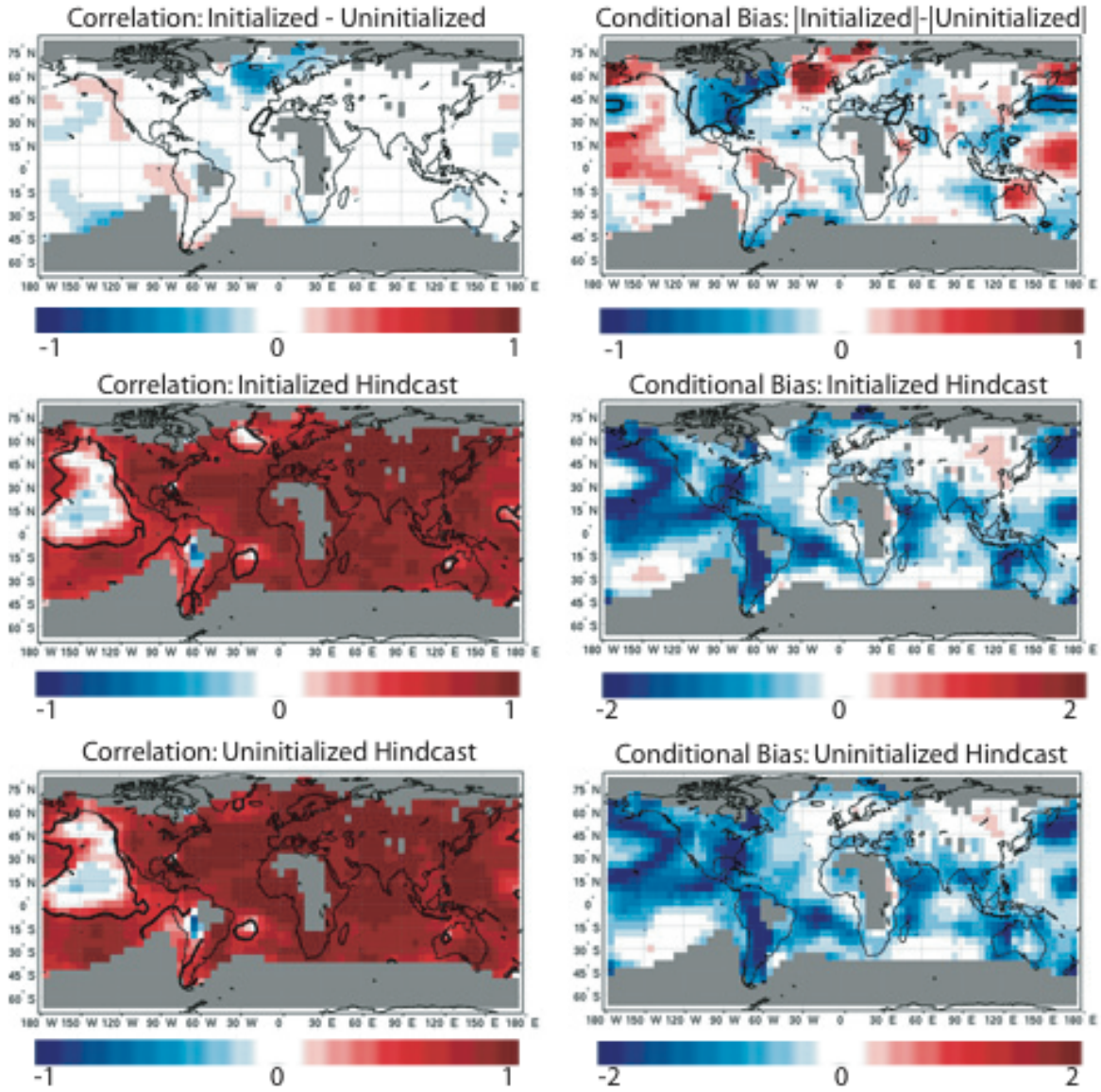


Figure 3. Same as Figure 4, but for CanCM4 hindcasts.

4. Concluding remarks /discussions

US CLIVAR Working Group on Decadal Predictability has developed a framework for verification of decadal hindcasts that allows for common observational data, metrics, temporal structure, spatial scale, and presentation. The framework addresses specific questions of the hindcast quality and offers suggestions for how they might be used. Considerable complementary research has aided this effort in areas of bias and forecast uncertainty, spatial scale of the information, and stationarity impacts on reference period.

The results from the hindcast verification performed on the two prediction systems yield some features that are also common to seasonal-to-interannual predictions. First, temperature is better predicted than precipitation. In this case the dominant signal is due to the upward trends, which are captured reasonably well by both systems over most of the world. In general, precipitation is a more localized variable in both space and time, and thus subject to larger noise-like variability that is not predictable. Second, forecasts from different prediction systems often differ in where they perform well. Some common areas of good and poor performance are seen in both prediction systems. However, many differences exist as well, especially for precipitation, and also for the impact of initialization.

Although these results may be sobering, they should not be viewed as a conclusion that there is no decadal predictability. Decadal prediction is very much an experimental activity, including how best to initialize the predictions. One positive result is the reduction in conditional bias that is seen for some areas in the initialized predictions, which is improved information about anthropogenic climate change. Those interested in these predictions should also visit the DPWG verification website to examine whether other time horizons might have more useable information. Even the more detailed paper that outlines the framework (Goddard *et al.* 2012) cannot show all the results, but there are instances of statistically significant skill obtained at the 1-year lead or 2-5 year period that do not appear in the decadal-scale results shown here. It is also possible that gains in prediction quality may be made by multi-model ensembling, as has been realized for seasonal prediction. Preliminary results based on just the two models used in this study show mixed results (not shown). Statistical post-processing, or calibration, of model predictions may also improve forecast quality. However, to do that robustly will require larger ensemble sizes and more forecast cases (*i.e.* more start dates) than was mandated for CMIP5. Finally development of improved models, and improved understanding of the processes that must be modeled well, is ongoing throughout the scientific community, and should be expected to improve the quality of decadal-scale climate information.

5. References

Arora, V., and Co-authors, 2011: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases. *Geophys. Res. Lett.*, **38**, L05805, doi:10.1029/2010GL046270.

Goddard, L., and Co-authors, 2012: A verification framework for interannual-to-decadal prediction experiments, *Clim. Dyn.*, submitted.

Merryfield, W. J., and coauthors, 2011: The Second Coupled Historical Forecasting Project (CHFP2): I. Models and Initialization. In preparation.

Murphy, A.H., 1988: Skill scores based on the mean squared error and their relationships to the correlation coefficient. *Mon. Wea. Rev.*, **116**, 2417-2424.

Räisänen, J., and J.S. Ylhäisi, 2011: How much should climate model output be smoothed in space? *J. Climate*, **24**, 867-880.

Smith, D. M., and Co-authors, 2010: Skilful multi-year predictions of Atlantic hurricane frequency. *Nature Geoscience*, doi: 10.1038/NGEO1004.

Solomon, A., and Co-authors, 2010: Distinguishing the roles of natural and anthropogenically forced decadal climate variability: Implications for prediction. *Bull. Amer. Meteor. Soc.*, doi: 10.1175/2010BAMS2962.1.

Taylor, K.E., R.J. Stouffer, and G.A. Meehl, 2011: An overview of CMIP5 and the experiment design, *Bull. Amer. Meteorol. Soc.*, doi: http://dx.doi.org/10.1175/BAMS-D-11-00094.1.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Development of a Seasonal Climate and Streamflow Forecasting Testbed for the Colorado River Basin

Andy Wood and Kevin Werner

NOAA/NWS Colorado Basin River Forecast Center, Salt Lake City, Utah

1. Motivation

Many groups have documented a consistent need for climate forecasts from one season to two years lead time to support a variety of applications, and particularly for streamflow forecasting for water, energy and agricultural management. Forecasts across these relatively long time scales are particularly valuable in the Colorado Basin (Figure 1) where the reservoir capacity is approximately four times the mean annual discharge of the river. Yet the Colorado River basin (CRB) presents a challenge due to the limited forecast skill that can be harnessed from traditional sources (*e.g.*, ENSO indices) even at shorter lead times for runoff-generating headwaters in the upper basin, the source of most of Colorado River flow. Nonetheless, management and planning objectives related to the larger reservoirs that the U.S. Bureau of Reclamation (hereafter Reclamation) manages apply streamflow projections for lead times up to two full years. Motivated to improve these forecasts, Reclamation has funded university research that has shown some promise for developing climate and streamflow predictions in the CRB based on a more expansive range of climate system indices and state variables. Other federal and state-funded research focused on the western US shares these goals and complements the Reclamation effort.

Fig. 1 Colorado River Basin

The NOAA/NWS Colorado Basin River Forecast Center (CBRFC) is the primary official provider of streamflow forecasting information products to Reclmation and other water management entities in the CRB, in some cases as a result of legal agreements between the seven western states that use Colorado River water. Reclamation and other agencies are unlikely to produce streamflow forecasts operationally, but rather depend on NOAA river forecast centers to provide forecasts that serve as input to their water operations and management activities. Consequently, the primary avenue for leveraging such research activities to support CRB water management is to evaluate the findings with respect to water management relevant watersheds, in comparison to current practice (the baseline), and where warranted, operationalize the new techniques into CBRFC streamflow forecasting efforts.

2. CBRFC streamflow forecasting

Operational forecasts that support reservoir management include one-week and peak flow predictions to manage high flows and recreation, but major water allocation decisions in the large storages of the CRB (Lakes Mead and Powell) depend primarily on "water supply forecasts", *i.e.*, predictions of runoff volumes into the major reservoirs of the upper CRB (listed in Table 1) at lead times of 1 to 24 months into the future. Water supply forecasts (WSF) currently rely on two primary techniques, statistical water supply (SWS)

Correspondence to: Andy Wood, NOAA/NWS Colorado Basin River Forecast Center, Salt Lake City, UT 84116; E-mail: andy.wood@noaa.gov.

prediction and Ensemble Streamflow Prediction (ESP). The former relates predictors such as observed snow water equivalent (SWE), water year precipitaton and runoff to date to future runoff, and is practiced in both the NOAA NWS RFCs and the NRCS National Water and Climate Center (NWCC). SWS methodology has remained essentially unchanged over the past half-century, despite two notable upgrades: (a) the deployment of the automated snow telemetry (SNOTEL) network for remote, real-time monitoring of SWE and precipitation, starting in the late 1970s (complementing manual snow course measurements); and (b) the upgrade of the statistical method from multiple linear regression to principal components regression (PCR, Garen 1992). ESP is an ensemble forecast approach that uses a continuous conceptual hydrologic model. ESP was introduced in the late 1970s (Twedt *et al.*, 1977), but was not widely used in NWS until the 1990s. Like SWS, ESP forecast skill derives primarily from the snowpack observed on the ground at the time the forecast is made. For the CRB, CBRFC subjectively merges its SWS and ESP WSFs into a preferred NWS forecast, and then subjectively coordinates with NWCC's SWS forecasts to arrive an an official WSF. During this process, forecasters from both agencies also discuss the current and forecast conditions and adjust the forecast in consideration of other factors deemed relevant. The resulting official forecast describes a probability distribution using the 90th, 50th, and 10th percentiles (non-exceedence probabilities).

Neither agency currently uses a climate forecast in the forecast process for the upper CRB, although both agencies are technically capable of doing so (*i.e.*, the software or approaches can admit a climate forecast as input). Short term weather forecasts (1-5 day QPF and 1-10 day temperature forecasts) are occasionally incorporated into the CBRFC ESP, at the subjective discretion of the forecaster, but not into the SWS from either agency. The primary reason that CBRFC and NWCC do not use climate forecasts (*e.g.*, CPC Official climate outlooks or ENSO indices) for prediction is that such forecasts have been found to have low to non-existent skill in the upper CRB. That the intermountain region of the western US exhibits low climate predictability has been well documented by agency investigation and academic research over the last decade.

3. CBRFC Climate and Streamflow Prediction Testbed

A workshop funded by the Colorado Water Conservation Board (CWCB) and co-sponsored by NIDIS was held at CBRFC in March, 2011 to discuss (1) research on seasonal to year 2 climate and flow predictability within the CRB and (2) formation of a testbed for comparing research results with the operational forecasts generated by CBRFC and used by Reclamation. The climate and reseach methods of interest from external researchers were primarily statistical, leveraging large-scale climate system information (*e.g.*, SST and geopotential height patterns), climate system indices, and other predictors to estimate future

Watershed Name	NWS ID	USGS ID	Drainage Area (km2)
Gunnison R abv Blue Mesa	BMDC2	09124800	9,092
San Juan River nr Navajo Res Archuleta	NVRN5	09355500	8,476
Green R at Flaming Gorge Res Flaming Gorge Dam	GRNU1	09234400	11,076 (50,310)
Gunnison R at Morrow Point Res	MPSC2	--	--
Taylor R at Taylor Park Res	TPIC2	09107000	332
Green R Nr Fontanelle Res Fontanelle	GBRW4	09211200	11,128
Gunnison R at Crystal Res	CLSC2	09127800	--
Los Pinos Nr Vallecito Res Bayfield	VCRC2	--	--
Major Upper Colorado River Basin Areas			
San Juan R nr Bluff	BFFU1	09379500	59,800
Green R nr Green R	GRVU1	09315000	116,610
Gunnison R nr Grand Junction, CO	GJNC2	9152500	19,958
Colorado R at Lake Powell Glen Cyn Dam	GLDA3	--	637,000

Table 1 The watershed data in the table above encompass the drainage areas of the 8 major river basins that directly support Reclamation probabilistic forecasting. The last four drainage areas are additional areas of interest: the outlets of three major tributaries above Lake Powell, and the entire Lake Powell drainage.

climate and streamflow in the Colorado River area. Typical methodologies this genre are described in Bracken *et al.* (2010), Grantz *et al.* (2005; 2007), Moradkhani and Meyer (2010), Najafi *et al.* (2011), Switanek *et al.* (2009) and Wang *et al.* (2009). ***The primary goal of the testbed is to focus collaborative efforts to improve prediction for the management of Colorado Basin water resources on the key climate and flow datasets involved in this enterprise.*** The central testbed website is located at http://www.cbrfc.noaa.gov/testbeds/si_y2/.

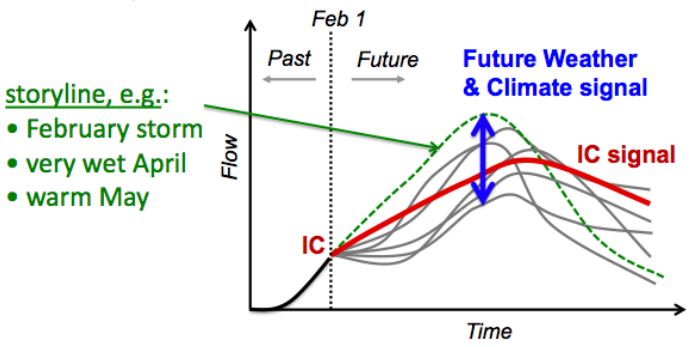


Fig. 2 Illustration of the signal components in a water supply streamflow forecast, which vary in importance throughout the year and may involve combinations of climate variables.

A basin-oriented focus

A guiding principle of the climate and flow forecasting testbed is that the evaluation of climate forecasts must be oriented toward river basins, and in particular, basins which are important to water management in the upper CRB. Climate forecasts are typically and sometimes exhaustively analyzed by forecast producers, but at space and time scales, and in a separate rather than covariational framework, that often does not offer clear connection to their potential for streamflow prediction. The complex connection of climate to hydrologic forecast outcomes is illustrated in Wood & Lettenmaier (2008), which shows the varying influences of initial hydrologic state and future climate on future streamflow. In a nutshell, ***every hydrologic anomaly has a story line***. As a water year progresses, for instance, past weather and climate are incorporated into the hydrologic initial conditions. In the CRB, for water supply forecasts, these initial conditions provide increasing signal to the runoff prediction, while the importance of future climate diminishes. Hydrologic extremes often involve pattern persistence (*e.g.*, a sequence of wet months), but can arise from more complicated, multi-variate climate phenomena, as illustrated in Figure 2.

Basin-oriented climate forecast evaluation means assessing climate variable predictions not only with respect to catchment areas of interest, but also with consideration of times of interest – *i.e.*, periods that are important for hydrologic response. For instance, predictions of precipitation anomalies during winter (*i.e.*, snow accumulation season) and of temperature anomalies during spring (*i.e.*, snow melt season) are more critical for water prediction purposes than at other times. Forecast accuracy can be more important if hydrologic conditions are anomalous (*i.e.*, during extreme years) for some users, than in normal years. The testbed will stress such considerations, in addition to evaluating the translation of climate prediction into streamflow predictions that improve over existing operational baselines. Although the temporal focus of the testbed evaluations must still be determined, the watersheds of interest are the drainages of the key upper CRB flow locations forming inputs for Reclamation reservoirs that influence water allocation decisions between the upper and lower CRB each year. These are listed in Table 1 and illustrated in Figure 3.

Testbed components and constraints

The testbed will contain the following elements, which are being populated as time permits by the lead author of this article.

- hindcast results sufficient to establish the current operational baseline for prediction skill
- timeseries that define the climatologies of the variables of interest – primarily precipitation, temperature and streamflow
- experimental forecast results in tabular and graphical form
- documentation of current forecast methods (*e.g.*, equations used for SWS)
- forecast evaluation metrics that are relevant to CBRFC operations (*e.g.*, the error metrics associated with April-July streamflow volume).

The current operational baseline for climate forecasting at CBRFC is essentially the use of historical climatologies of precipitation and temperature for the watershed areas for which CBRFC predicts flow. The

baseline for streamflow prediction at different lead times and for different predictands is the ESP approach, forced with the baseline (historical) precipitation and temperature inputs. For some predictands, the SWS water supply forecasts also form an operational baseline.

Two experimental climate and flow prediction approaches are now being developed CBRFC for seasonal/interannual lead times. One is the use of the CPC objective consolidation forecasts that are produced operationally in support of the official monthly CPC climate outlook. The other is the use of NCEP Climate Forecast System (CFS, soon to be CFSv2) precipitation and temperature, downscaled and calibrated to the watershed scale via statistical methods documented in Seo *et al.* (2006) and Wu *et al*. (2011).

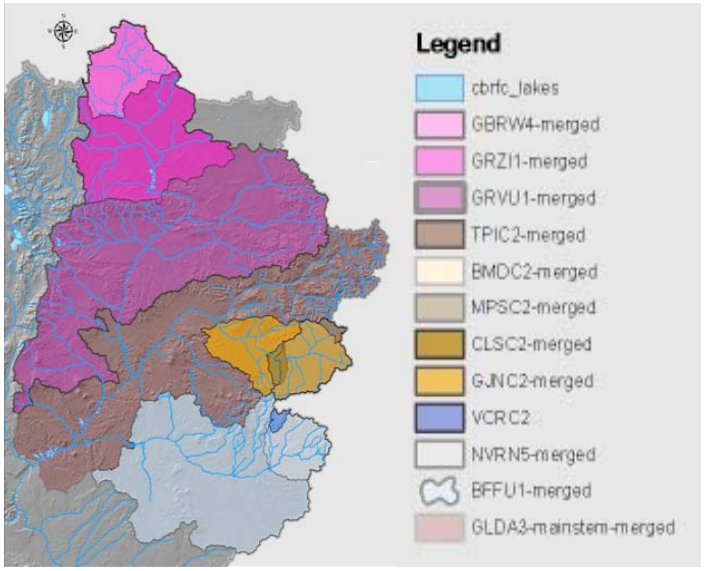


Fig. 3 Testbed watershed areas

A number of avenues for translating climate forecasts into streamflow forecasts exist. For instance, improved climate forecasts can be translated to streamflow forecast via a trace-weighting of ESP ensembles, in which the weightings for each trace (corresponding to a historical meteorological sequence) are derived from a climate forecast (*e.g.*, Werner *et al.*, 2004). Another approach might involve the modification of historical precipitation and temperature forecast time series to match climate forecast characteristics before input to ESP. More elaborate approaches involving, *e.g.*, synthetic weather generation and hydrological modeling, are under development by NOAA-funded collaborators (such as Dr. Balaji Rajagopalan at the University of Colorado). Climate predictors can also be related directly to predicted streamflow characteristics.

While CBRFC continues to populate the testbed from within, researchers are encouraged to train statistical climate forecast technique using historical data (observations or hindcasts) from one or more of the key watersheds provided, and for one or more forecast initialization dates (*e.g.*, October 1). The specific challenges raised in the testbed are to demonstrate that new climate prediction approaches:

- provide superior climate forecast skill (precipitation and/or temperature) relative to baseline forecast approaches available to CBRFC
- lead to improved long lead streamflow forecast skill, *i.e.*, when implemented via trace-weighting or alternative approach.

CBRFC will be focusing on applying the climate forecasts that warrant attention (via positive performance and abide by the constraints listed below) to streamflow predictions, though researchers are welcome to tackle that part of the challenge as well. Because an RFC is an operational center providing real-time climate and water information services, the testbed includes several constraints:

- The source datasets must be available with low enough latency to support a real-time prediction.
- The methods must be reasonably automatable, and configuration or setup steps must be "teachable" rather than arcane.
- The methods must ultimately (though not for prototyping) use non-proprietary software: *e.g.*, R is preferred to Matlab.

4. Discussion

The S/I to Year 2 Climate and Flow Forecasting Testbed is a CBRFC-led effort to engage and educate external research collaboration toward improving water management via advanced climate and streamflow forecasting in the western U.S. Although CBRFC is not a funding entity, this goal has been prominent in the

objectives of a number of federal and state grant opportunities, thus CBRFC seeks to attract collaborators to the testbed both by leveraging existing funded projects and by pursuing new ones through joint (co-investigative) proposals. While CBRFC has some capacity to engage in new research and development toward the objectives of the testbed, and can lean on some assistance from NWS and NOAA laboratories, integrated efforts that team CBRFC personnel with external groups is an ideal way to build capacity in CBRFC, educate external groups about operational forecasting, and ensure that applies research effort toward ostensible operational uses is properly focused and evaluated.

Although the major results of the SI/Y2 testbed are still evolving, the concept of a water-oriented climate and streamflow prediction testbed have broad advantages beyond the Colorado River basin. Water is a primary sector in which benefits from improved climate forecasting can be derived, and the evaluation of climate prediction through the prism of relevance to water prediction can add a valuable context for climate forecast producers, and a compelling demonstration for climate forecast users in the water sector. One can envision a nationwide testbed of this type as a framework to integrate researchers and forecasters from climate science to hydrology to water resources. The literature is full of research results that claim to have potential to advance water management, yet have never been implemented operationally – this gap may arise from a failure of integration that the testbed effort is designed to address.

Acknowledgements. The authors acknowledge funding support for the aforementioned workshop from the Colorado Water Conservation Board and the National Integrated Drought Information System (NIDIS).

References

Bracken, C., B. Rajagopalan, and J. Prairie, 2010: A multisite seasonal ensemble streamflow forecasting technique. *Water Resour. Res.*, **46**, Art. No. W03532, issn: 0043-1397, ids: 578UP, doi: 10.1029/2009WR007965.

Grantz, K., B. Rajagopalan, E. Zagona, and M. Clark, 2007: Water management applications of climate-based hydrologic forecasts: Case study of the Truckee-Carson River Basin. *J. Water Resour. Plan. Manage*.-ASCE, **133** (4), 339-350, issn: 0733-9496, ids: 180IZ, doi: 10.1061/(ASCE)0733-9496(2007)133:4(339)

Grantz, K., B. Rajagopalan, M. Clark, and E. Zagona, 2005: A technique for incorporating large-scale climate information in basin-scale ensemble streamflow forecasts. *Water Resour. Res.*, **41**, W10410, doi:10.1029/2004WR003467.

Moradkhani, H., and M. Meier, 2010: Long-Lead Water Supply Forecast using Large-scale Climate Predictors and Independent Component Analysis, *J. of Hydrologic Engineering*, **15**(10), doi: 10.1061/ASCE-HE.1943-5584.0000246.

Najafi, M., H. Moradkhani, and S. Wherry, Statistical Downscaling of Precipitation using Machine Learning with Optimal Predictor Selection, *J. of Hydrologic Engineering*, in press.

Seo D.-J., H.D. Herr, and J.C. Schaake, 2006: A statistical post-processor for accounting of hydrologic uncertainty in short-range ensemble streamflow prediction. *Hydrology and Earth System Sciences Discussions*, **3**, 1987-2035.

Switanek, Matthew B., Peter A. Troch, and Christopher L. Castro, 2009: Improving Seasonal Predictions of Climate Variability and Water Availability at the Catchment Scale. *J. Hydrometeor*, **10**, 1521-1533.

Wang, S.-Y., R. R. Gillies, J. Jin, and L. E. Hipps, 2009: Recent rainfall cycle in the Intermountain Region as a quadrature amplitude modulation from the Pacific decadal oscillation. *Geophys. Res. Lett.*, **36**, L02705, doi:10.1029/2008GL036329

Werner, K., D. Brandon, M. Clark, and S. Gangopadhyay, 2004: Climate Index Weighting Schemes for NWS ESP-Based Seasonal Volume Forecasts. *J. Hydrometeor*, **5**, 1076-1090.

Wood, A. W., and D. P. Lettenmaier, 2008: An ensemble approach for attribution of hydrologic prediction uncertainty. *Geophys. Res. Lett.*, **35**, L14401, doi:10.1029/2008GL034648.

Wu, Limin, D.-J. Seo, J. Demargne, J.D. Brown, S. Cong, and J. Schaake, 2011: Generation of ensemble precipitation forecast from single-valued quantitative precipitation forecast for hydrologic ensemble prediction. *J. Hydrology* (in press).

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Performance Characteristics of Forecasts Based on a Lagged Ensemble

David A. Unger

Climate Prediction Center, NCEP/NWS/NOAA, Camp Springs, Maryland

1. Introduction

The Climate Forecast System Version 2 (CFSv2) is a coupled atmospheric-oceanic global climate model run by the National Centers for Environmental Prediction (NCEP). The CFSv2 is run operationally each 6 hours and makes forecasts out to around 6 months lead time. Uncertainty is estimated by a forecast ensemble composed of recently available initial times and aligned according to the target period in question. Thus, an ensemble forecast set is formed from a series of runs lagged backwards in time from the run based on the most recent data, and is hence referred to as a lagged ensemble (Hoffman and Kalnay 1983). This contrasts with the approach commonly used for shorter range forecasting of producing an ensemble forecast from multiple runs initialized from perturbed initial states based on the most recent data.

For lead times typical of climate forecasts, usually measured in months, the forecast benefits of a larger ensemble more than compensates for losses due to the inclusion of negligibly less skillful members from runs a few hours or days old. However, there is eventually a point where the trade-off between ensemble size and loss of skill due to the inclusion of older members becomes important. At some point the inclusion of less accurate ensemble members from older runs will begin to detract from the information available from members based on more recent data. An objective procedure to build a lagged ensemble set is proposed in this paper. This method is based on the skill characteristics of the candidate ensemble members. The result is an objective weighting method for ensemble members from variously lagged initial times.

2. Data

Sea Surface Temperature (SST) forecasts for the Nino 3.4 region from the CFSv2 will be used to test the method. Hindcast data initialized every 5th day from 1982 – 2010 is available from the CFSv2. The hindcast data was adjusted for a discontinuity in forecast performance beginning in late 1998. A simple bias correction, stratified by initial time and lead, was applied to the earlier forecasts (1982-1998) to make their mean bias similar to the recent forecasts (1999-2010). Forecasts are expressed as anomalies relative to the 1981-2010 Nino 3.4 climatology.

The 4 ensemble members obtained from initial times 6 hours apart on any given day were assumed to be equal in skill and always were grouped together, treating them as if they were perturbations from a daily initial state. A lagged ensemble set was built from the series of 4-member daily runs available from CFSv2 hindcast data going backward in time from the run closest to the end of each calendar month, defined as lag 0. Members from 5-days prior to lag 0 were labeled as lag 5, 10 days prior as lag 10 and so on. The forecasts were for the monthly mean Nino 3.4 SST for the month following the lag 0 forecast (defined as lead 1), the subsequent month (lead 2), and successive months out to lead 6. Up to 20 4-member ensemble sets were offered for potential inclusion in a lagged ensemble (out to lag 95), adjusting the lead time of earlier runs as necessary to align the forecast relative to lag 0.

3. Methods

a. Ensemble Regression

A regression procedure specifically designed for ensemble forecasting was used to process the CFSv2 forecasts. Ensemble regression (Unger, *et al.* 2009) estimates the expected values of the coefficients of a regression equation relating the closest member of a set of ensemble solutions to the observation, together with its expected error. In addition to the usual requirements for the appropriate application of linear

Correspondence to: David A. Unger, Climate Prediction Center, NCEP/NWS/NOAA, Camp Springs, MD 20746; E-mail: David.Unger@noaa.gov.

regression and its error estimate (a reasonably linear fit, Gaussian distributed residual errors, case to case independence of errors, etc.), ensemble regression requires a reasonable a priori estimate of the probability that each ensemble member will be closest to the observation, usually taken to be equal for all members. Once this estimate is supplied, a regression equation can be derived that gives theoretically expected values of the coefficients, together with an expected value of the standard deviation of the errors, assumed to be Gaussian, about the regression estimate. This equation, together with the error estimate, is applied to each ensemble member.

Ensemble regression was used to calibrate the forecasts for an ensemble consisting of the 4 runs available for a calendar day. Each of these ensemble members were assumed equally likely to be closest to the observation (best member). In its application, ensemble regression produces an estimate of the forecast probability density function (PDF) of the combined ensemble. It is similar to a PDF obtained from a Gaussian kernel density estimate (Roulston and Smith 2003), except that the kernels are centered around the regression estimates of each ensemble member, rather than their original values. The standard deviation of the kernel distributions are specified by the regression procedure.

Regression equations were derived for the predicted monthly mean Nino 3.4 SST from each of 73 initial times (every 5 days referred to here as pentads) throughout the year. The equations for each pentad were based on the 29 cases for that initial time in the 1982-2010 period. A separate equation was derived for each monthly lead time. Results reported here are based on dependent data (no cross-validation).

Figure 1 illustrates the forecast PDF from two separate groups of ensemble forecasts. This is a forecast for the monthly mean Nino 3.4 SST for July, 2010. One forecast (indicated by blue) was initialized on June 30, 2010, only 1 day prior to the start of the valid period. The solid line is the combined PDF from the 4 kernel distributions representing individual members, shown by dashed lines. The forecasts at this lead were highly skillful, indicated by closely grouped, narrow kernels. A forecast made 3-months earlier (April 1, lag 90) for the same target period is shown in red. Note that the expected errors around the best member (illustrated by kernels) are only slightly less skillful than the shorter lead time as evident from the kernel width. This indicates that most of the difference in forecast uncertainty is accounted for by the increased amount of spread compared to the earlier run. The green vertical line shows the observed mean SST for July 2010.

b. A procedure for weighting a lagged ensemble

Figure 1 also illustrates a principle for optimizing a lagged ensemble. The PDF from the ensemble regression in the vicinity of the observation provides a measure of how closely the forecast and observation match. The PDF from each candidate ensemble forecasts (one for each lag) is used to provide a measure of which was "best" in a given case. The ensemble set with the highest PDF value at the observation is, by inference, the one that is most likely to be correct. The regression calibrates

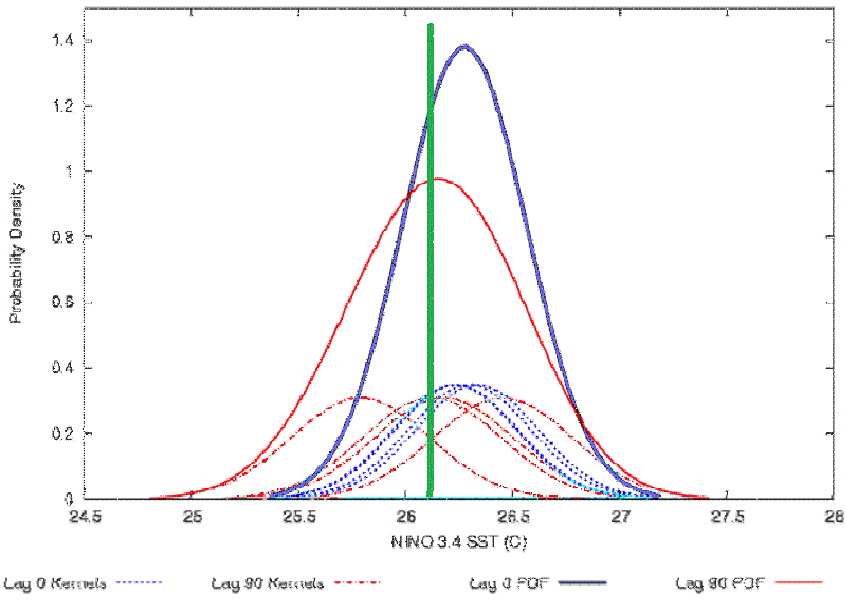


Fig. 1 Forecast PDFs based on ensemble regression from CFSv2 forecasts of Nino 3.4 SST from two initial times. Blue lines represent forecasts of monthly mean Nino 3.4 SST initialized on June 30, 2010 valid for July, 2010. Red lines are for the Lag 90 forecast (initialized April 1). Kernels represent individual members and are shown by dashed lines, with the combined PDF in solid. The green vertical line represents the observed SST.

the entire ensemble in relation to its performance on hindcast data. So, in this case, the single best member of the 8-member ensemble formed from the lag 0 and lag 90 runs happened to be from the lag 90 set. However, the regression processing accounts for both the spread and the skill to better match the forecasts and observations. The regression calibrated PDF indicated that, considering its spread, and the skill of CFSv2 forecasts from similar leads on the hindcast data, the more recent run, in fact, was the better match to the observation.

The potential ensemble system consisting of the 20 most recently available 4-member lagged ensemble forecast sets were evaluated on the dependent (hindcast) data by passing through the data twice, once to derive the ensemble regression equations, and once to evaluate forecasts and produce a separate PDF for each lag. The ensemble set that produced the "best member" was assumed to be the one with the highest PDF in each case. The probability that each of the 20 candidate ensemble sets produced the best member (P(best)) was estimated from the fraction of times out of the 29 cases (one per year) that a given lag was best. The results were smoothed under the assumption that P(best) decreases with increasing lag time. If this was not so, the older members were assumed equally likely to produce a best member than any more recent ensemble set with a lower P(best). The P(best) for each lag was then re-evaluated with the combined ensemble, effectively averaging the P(best) values over any inconsistent lead times. This way P(best) of an older run was assured to be equal to or lower than a more recent run.

4. Results

The top panel of Fig. 2 shows P(best) for lagged ensemble forecasts for SSTs initialized near the end of each of 12 calendar months, for a lead-3 forecast. The vertically stratified bins (columns) correspond to the initial month, and the boxes along the vertical (rows) correspond to the lag relative to the most recent run. The bottom row represents the four member ensemble closest to the end of the month in the corresponding column. Each higher row represents an ensemble set initialized 5 days earlier than the row below. The inset numbers show P(best) for that ensemble set, objectively determined from the hindcast data as described above. P(best) is an appropriate weighting for that ensemble set in ensemble regression theory. The final forecast of a weighted ensemble would be formed from a kernel density approach similar to that shown in Fig. 1 except with the area of each of the 4 kernel proportional to ¼ of the total weight assigned to that lag.

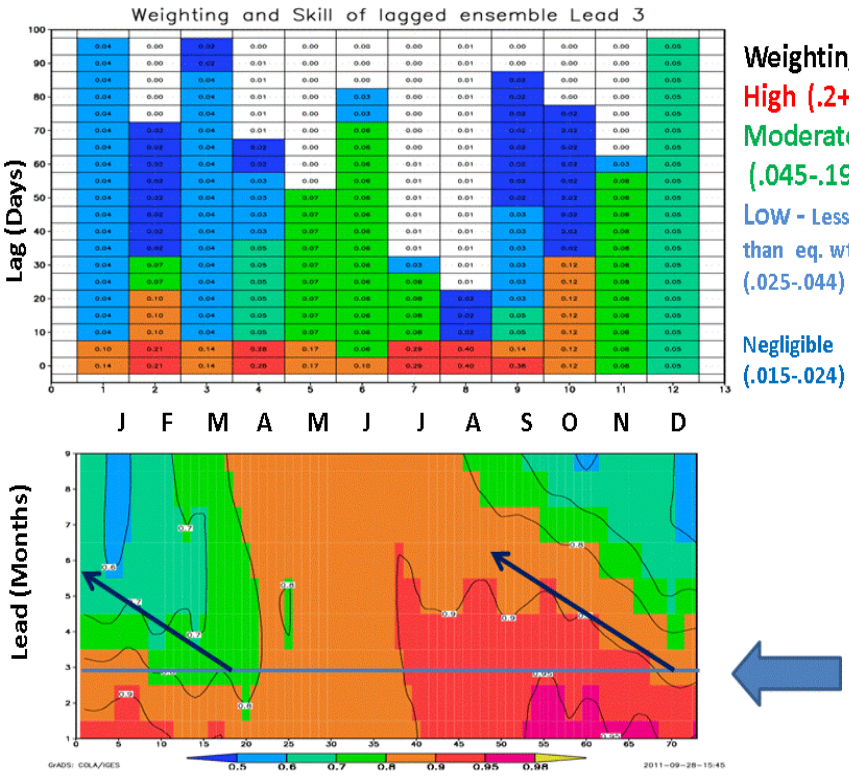


Fig. 2 Weighting for each lag (P(best)) stratified by initial month and lead time (top panel), with Hovmoller plot of regression skill by lead and initial time in pentads (bottom panel). Inset values in the bottom panel represent the correlation between the ensemble mean and the observation for a forecast for monthly mean Nino 3.4 SST from the CFSv2. See text for further explanation.

For example, a forecast for July, initialized at the end of June (column 7, bottom row), was the "best"of any of the 20 lagged ensemble sets offered 29 percent of the time. This run was initialized on June 30 of each

year of hindcast data. The run initialized 10 days before that, on June 20, was best 8 percent of the time. The oldest run to contribute significantly to lagged ensemble was from lag 30 (initialized on May 31) and accounted for only 3% of the best members (was best once out of the 29 year sample). Members initialized at lags between 35 and 55 days only sporadically contributed to the combined ensemble, accounting for trace weighting (1%), and no run 60-days or older contributed to the ensemble at all. Colors indicate the relative contribution, with red signifying a weighting significantly higher than an equally weighted 20-set ensemble. Green indicates moderate weighting, and blue indicating significantly less weight than would be expected for equal weighs among candidate ensemble sets.

The bottom panel of Fig. 2 reveals the possible explanation for the ensemble weighting. It shows the Hovmoller plot of the skill of a regression equation based on the ensemble mean (an important component of the ensemble regression). The 73 initial pentads are along the horizontal axis, with integer lead time in months relative to the calendar month of initialization shown on the vertical. The correlation coefficient of the regression relationship is contoured within the diagram. Even with smoothing, some evidence of a saw-tooth pattern is visible in the plot, indicating that the skill is higher for runs initialized late in the month than those initialized earlier on. The horizontal line illustrates the lead time (3-month) used for the ensemble weighting diagram on the top panel. The arrows indicate the direction in which the lagged ensemble is built for runs initialized in March and December. The skill of most recent run is shown near the tail of the arrow, with older runs represented toward the arrow's head.

When the skill of the ensemble set falls off rapidly with increasing lead, as indicated by skill decreasing along the arrow towards its head, the objective ensemble weighting procedure heavily front loads the lagged ensemble set. The weight drops rapidly and becomes trivial even after a few lags. This is common around the time of the spring predictability barrier indicated by areas of strong gradients on the diagram. On the other hand, skill in predicting Nino 3.4 SST does not depend much on lead time for runs initialized late in the year. The skill at the head of the arrow (oldest runs) is not much different from the newest runs. The lagged ensemble at this time of year is fairly evenly weighted, in spite of lags of 30 days or more. Thus, late in the year a larger ensemble group is adding information to the system, compared to a smaller ensemble consisting only of more recent runs. This is not the case in the boreal spring, where the bulk of the weight is on the most recent members, and older members are trivially weighted at best. At this time of year, the newer runs are more skillful than older ones, and the ensemble members from older runs provide little if any information not represented by the more recent runs.

5. Conclusion

The procedure described in this paper provides an objective way to analyze the information in a time-lagged ensemble. It can be used to help determine when ensemble members from older runs can meaningfully contribute to an ensemble consisting of members from more recent runs. The method provides an objective weighting procedure for the lagged ensemble forecasts, and provides results consistent with expectations based on forecast skill. This method is not restricted to lagged ensembles, since it can be used as a basis for combining multi-model ensembles as well. Future work will focus on evaluating the skill of this procedure in a fully cross-validated framework.

References

Hoffman, R., N., and E. Kalnay, 1983: Lagged average forecasting, an alternative to Monte Carlo forecasting. *Tellus*, **35A**, 100-118.

Roulston, M. S., and L. A. Smith, 2003: Combining dynamic and statistical ensembles. *Tellus*, **55A,** 16-30.

Unger, D. A., H van den Dool, E. O'lenic, and D. Collins, 2009: Ensemble Regression. *Mon Wea. Rev*. **137,** 2365 – 2379.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Enhancing Hydrological Seasonal Forecast by Downscaling CFSv2

Eric F. Wood, Xing Yuan, Joshua K. Roundy

Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ

1. Introduction

Due to the improvements in data assimilation techniques, computing resources, and numerical models such as representing large-scale climate teleconnections, forecasting seasonal climate by using coupled atmosphere-ocean-land general circulation models (CGCMs) has been an emerging area since mid-1990s. Now the major operational weather and climate forecast centers around the world are producing real-time seasonal climate predictions with CGCMs up to nine months. The progress in dynamical seasonal forecasts provides potential opportunity to predict hydrologic variables (*e.g.*, streamflow, soil moisture) at long lead times, which is important for agriculture and water resources management, drought and flood detection and mitigation. In this context, seasonal hydrologic forecast plays an important role in transitioning the scientific advances from the climate research community to the end users of society (Yuan *et al.* 2011b).

Recently, NCEP has updated its operational seasonal forecast system with a new CGCM, the second version of CFS (CFSv2), where a number of new physical packages for cloud-aerosol-radiation, land surface, ocean and sea ice processes, as well as a new atmosphere-ocean-land data assimilation system have been incorporated (Saha *et al.* 2010, 2011). Therefore, it is necessary to conduct comprehensive seasonal hydrologic reforecasts, and to investigate whether or how much of the improvement from climate forecast models can propagate into the land surface hydrologic forecast. Here, we summarize a deterministic assessment and probabilistic evaluation of CFSv2 by comparing with CFSv1 and European seasonal forecast models, and the CFSv2-based seasonal hydrologic forecast results. Full documentations are available in Yuan et al. (2011a, 2011b).

2. Data and method

To have a first look at CFSv2 in a hydrologic forecast perspective, we used the data as follows: 1) the 28-year (1982-2009) ensemble retrospective forecast data set from CFSv2 with 24 members at T126 resolution (Saha *et al.* 2011), and the nine-month reforecast for CFSv1 with 15 members at T62 resolution covering the same period 1982-2009 (Saha *et al.* 2006); 2) the EUROSIP (European Operational Seasonal to Interannual Prediction) model (ECMWF, Météo France (MF), and UK Met Office (UKMO)) reforecasts at 2.5°×2.5° resolution from 1960-2005 (Weisheimer *et al.* 2009); and 3) surface air temperature at 0.5° from the Climate Research Unit (CRU) TS3.1 dataset for the period 1901-2009 (Mitchell and Jones 2005), and precipitation at 0.5° from the Climate Prediction Center (CPC) Unified Gauge-Based Analysis for 1979-2009 (Chen *et al.* 2008).

We downscaled the precipitation and temperature hindcasts from CFSv1 and CFSv2 to 1/8 degree over CONUS by using the Bayesian method described in Luo *et al.* (2007) and Luo and Wood (2008). The downscaled precipitation and temperature fields are used as inputs to the VIC model (in water balance mode) to provide 6-month, 20-member ensemble hydrologic reforecasts starting on the 1st of February, May, August and November of each year during 1982-2008, with initial conditions from a 62-year (1949-2010) offline simulation driven by merged data from 31-year (1949-1979) University of Washington dataset (Maurer *et al.* 2002) and 31-year (1980-2010) dataset from the North-American Land Data Assimilation System Project Phase 2 (NLDAS-2; Xia *et al.* 2011).

3. Results

Figure 1 illustrates the geographic distributions of the month-1 predictive skill in terms of correlation for surface air temperature forecasts over land grids for CFSv1, CFSv2, ECMWF, MF, UKMO and a multi-

Correspondence to: Eric F. Wood, Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ; E-mail: efwood@princeton.edu.

model forecast (Yuan *et al.* 2011a). As compared with CFSv1, CFSv2 shows overall improvement, especially in the cold season. Wang *et al.* (2010) identified the cold bias of CFSv1 over northern hemisphere mid-high latitudes in the real-time seasonal forecast during the warm season, while the CFSv2 reduced the bias for the August forecast by 53%, averaged over the globe, and the reduction was more pronounced over the high latitudes in Eurasia (not shown). As compared with the ENSEMBLES EUROSIP models, CFSv2 has similar performance to ECMWF, where the former has higher skill over North America in November, while the latter has higher skill over northern Africa in February. The equally-weighted multi-model combines the advantages of individual models, and presents generally improved predictability. However, neither the multi-model nor the individual models in this study have significant predictability over western Russia in May and August, which is an issue that needs further investigation. On average, the global mean (excluding Antarctica) correlations of surface air temperature for the four months are as follows: CFSv1, 0.38; CFSv2, 0.52; ECMWF, 0.51; MF, 0.39; UKMO, 0.44; and the multi-model, 0.54. Note that the global mean correlation for multi-model in November is slightly smaller than the CFSv2, which may result partly from the low skill of MF and UKMO (Figure 1).

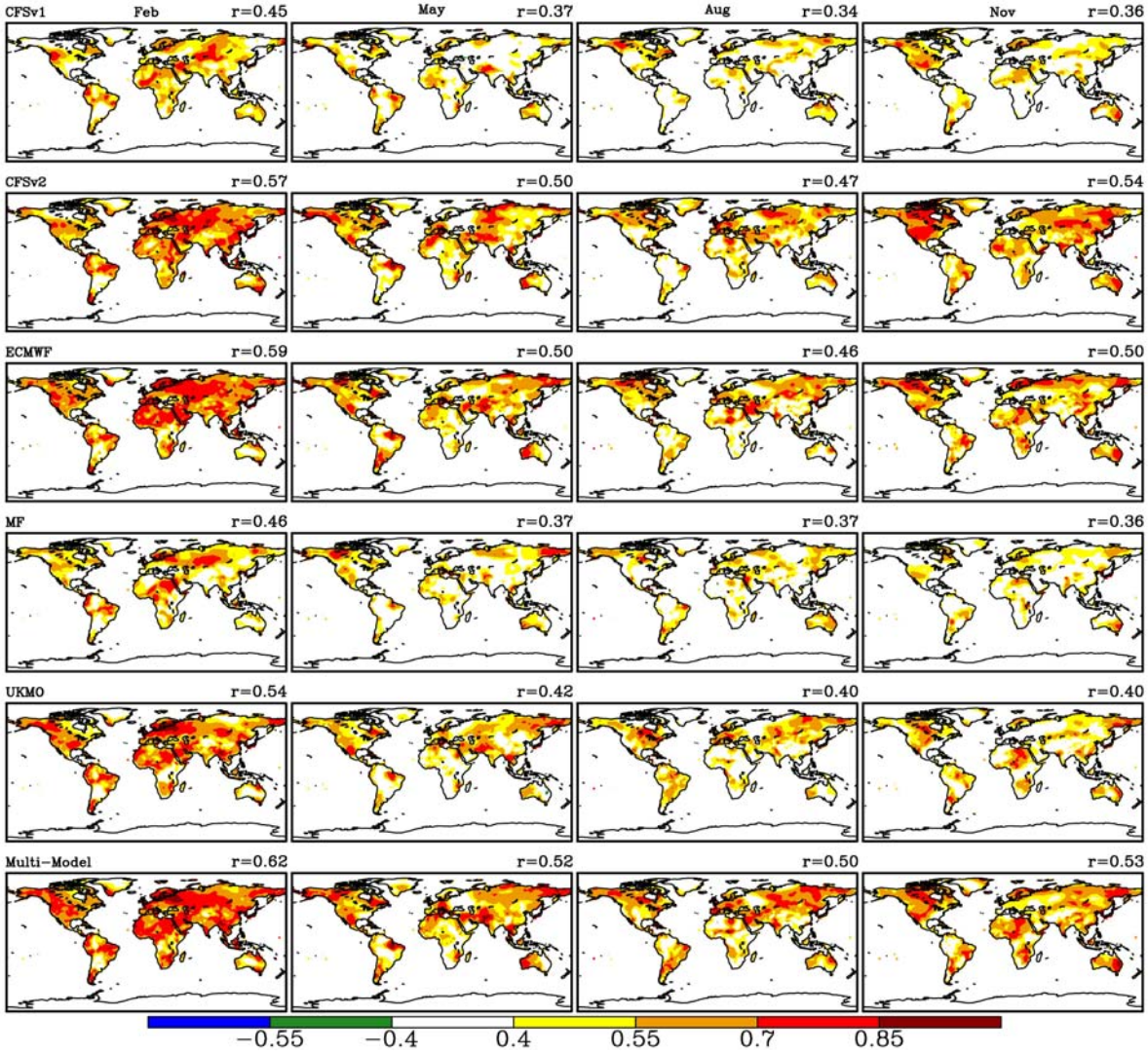


Fig. 1 Month-1 predictive skill of monthly surface air temperature forecasts over land grids in February, May, August and November during 1982-2005. The multi-model is the average among CFSv2, ECMWF, MF and UKMO. Non-white colors represent significant correlation at 0.05 levels. The numbers are the global mean correlations (Yuan *et al.* 2011a).

Figure 2 presents the percentage of the positive RPSS for monthly surface air temperature and precipitation anomaly over the global land area (excluding Antarctic). For the surface air temperature anomaly, the CFSv2 has a higher percentage of forecasts with skill beyond climatology than the other models. Even out to two months, more than 54% of the forecasts from the CFSv2 produce useful predictions. For the precipitation anomaly, the performance of the CFSv2 is comparable to the ECMWF, and both are much better than the other models. Similarly to the deterministic evaluation, the skill of probabilistic forecasts for precipitation for each individual model and the multi-model drops greatly beyond one month.

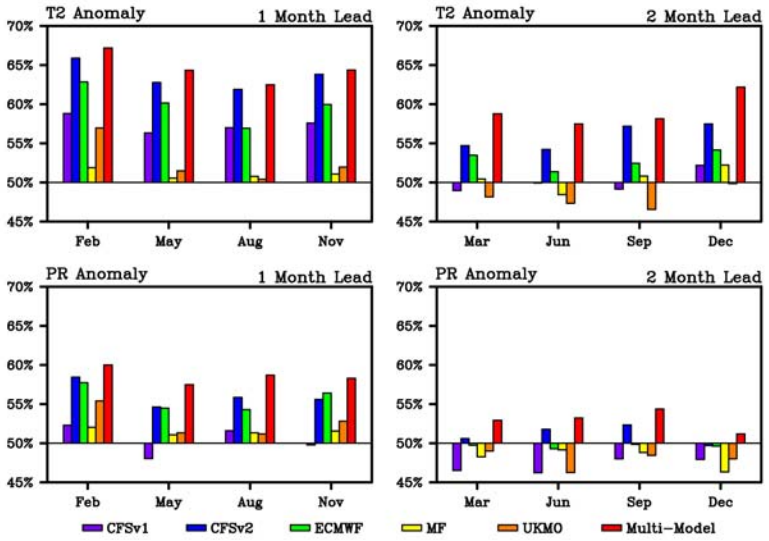


Fig. 2 Percentage of positive Ranked Probability Skill Score (RPSS) for monthly surface air temperature and precipitation anomaly over the global land area during 1982-2005 (Yuan *et al.* 2011a).

Figure 3 shows the correlation between NLDAS-2 observation and downscaled month-1 precipitation forecasts at 1/8 degree over conterminous U.S. in February, May, August and November during 1982-2008 (Yuan *et al.* 2011b). The patterns are similar to the original CFSv1 and CFSv2 forecast results at 2.5 degree (Yuan *et al.* 2011a), where CFSv2 is generally better than CFSv1 especially in cold seasons. For instance, there are obvious improvements over Pacific Northwest, New Mexico-Oklahoma-Texas region and eastern coast for the forecasts in February and November (Figure 3a,d). In May, CFSv2 has higher skill than CFSv1 over southwestern U.S., Ohio basin and Montana, but has lower skill over the southeast (Figure 3b). In the summer time, both CFSv1 and CFSv2 have limited skill, so the CFSv2 has negligible improvement (Figure 3c). On average, CFSv2 increases percentage of grid cells with significant predictive skill (>0.38) from CFSv1 by 61%, 45%, 14% and 60% for the four months, respectively.

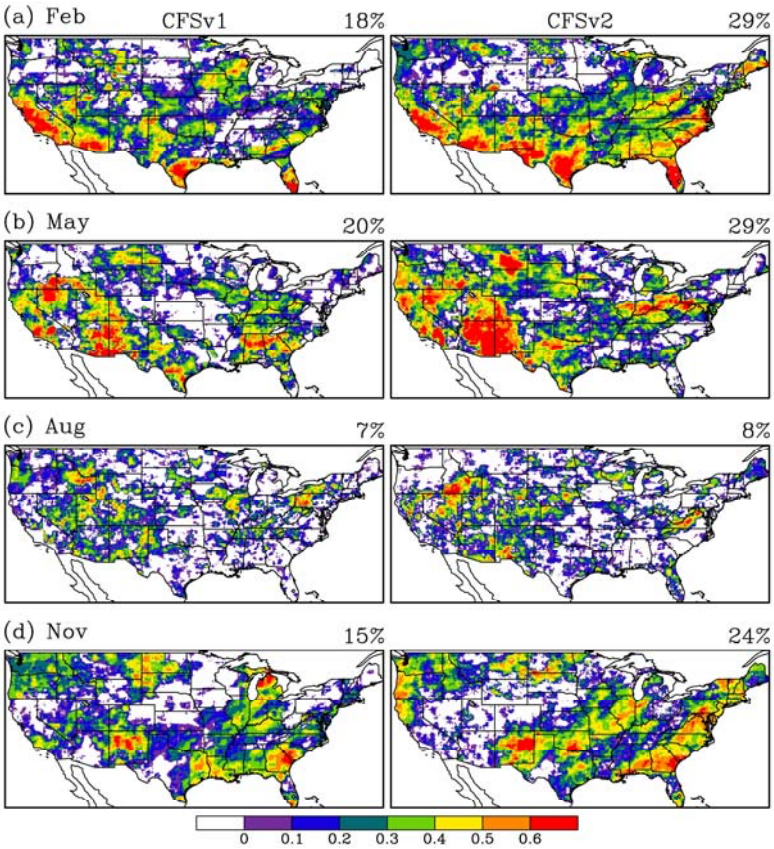


Fig. 3 Correlation between NLDAS-2 observation and Bayesian downscaled CFSv1 and CFSv2 month-1 precipitation forecasts at 1/8 degree over conterminous U.S. in February, May, August and November during 1982-2008. The numbers are percentages of grid cells with significant predictive skill (Yuan *et al.* 2011b).

Before using the downscaled forecast forcing to provide seasonal hydrologic forecast, a 62-year (1949-2010) offline simulation was performed to generate initial condition and reference data for

validation (*e.g.*, soil moisture, runoff), and test the capability of VIC model in capturing the streamflow interannual variations given observed forcing. To match the reforecast period, we calculated the Nash-Sutcliffe efficiency coefficients for monthly streamflow of offline VIC simulations at 416 U.S. Geological Survey (USGS) gauges during 1982-2008. Figure 4 presents locations of the gauges and their corresponding efficiency coefficients (Yuan *et al.* 2011b). Most of the gauges used in this study are in eastern U.S., and the highest coefficients are mainly over Ohio basin, Northeast and lower Mississippi. Among the 416 gauges, there are 315 (75%) and 130 (31%) gauges with coefficients larger than 0.3 and 0.7, respectively. Figure 5 shows interannual variations of streamflow from simulation and observation at four selected gauges with drainage areas ranging from 1027 to 20300 square miles (Yuan *et al.* 2011b). Consistent with high Nash-Sutcliffe coefficients, the VIC model captured seasonal fluctuation of streamflow quite well at the four gauges during 1982-2008, even though we started the model from 1949. Given the limited long-term in situ soil moisture observations over CONUS and the hypothesis that well-calibrated hydrologic model could provide reasonable soil moisture, we used the offline simulated soil moisture as a reference data to validate the hydrologic forecast.

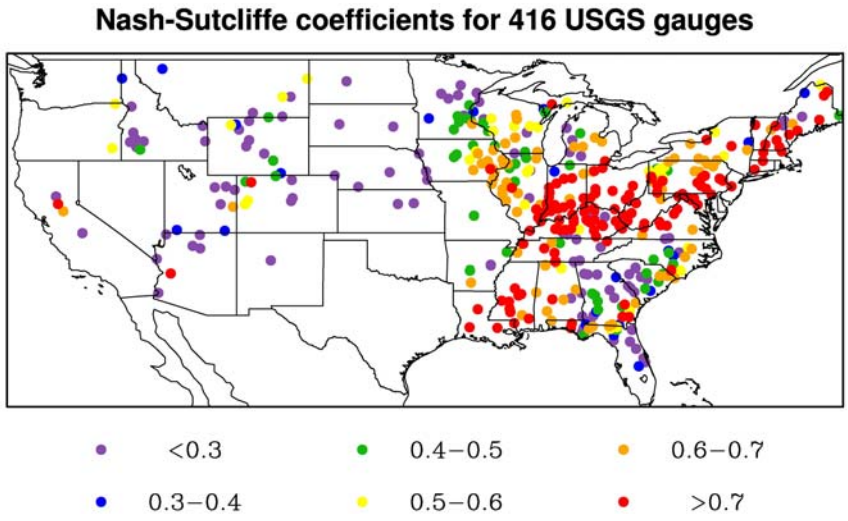


Fig. 4 Efficiency coefficients of monthly streamflow from offline VIC simulations at 416 USGS gauges during 1982-2008 (Yuan *et al.* 2011b).

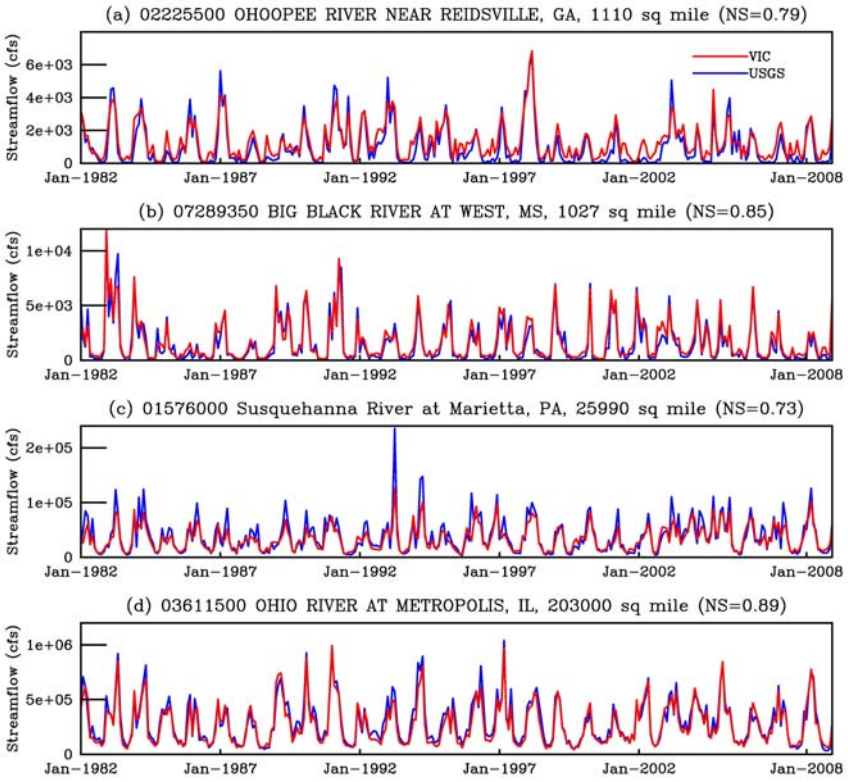


Fig. 5 Interannual variations of VIC offline simulated monthly streamflow compared with USGS observation at four selected gauges (Yuan *et al.* 2011b).

Figure 6 shows the Relative Operating Characteristic (ROC) diagram for low and high flow forecasts (Yuan *et al.* 2011b). The results were calculated for 130 gauges with NS coefficients large than 0.7 (Figure 4). All forecasts were more skillful than random forecast in the first three months. Low flow forecasts were a little better than high flow forecasts in the first month, regardless if the method is ESP or the dynamical climate model; however, their differences diminished beyond month-1. CFSv1 had no advantage in distinguishing low flows over ESP in the first month due to the strong impact on the forecast from initial condition. As the effect of initial conditions decreased and climate model still maintaining some skill, CFSv1 outperformed ESP in the second month;

while their performance became similar again because of the skill decrease in the climate model (Figure 6). With improved skill, CFSv2 was consistently better than ESP in the first three months, and increased the area under ROC curve by 4-7%. For high flows, the climate forecast model-based approach only limited skill beyond ESP in the first month (Figure 6). The moderate advantage manifested the importance of initial hydrologic conditions in identifying low and high flows, and the need of improvement for climate forecasting beyond month-1.

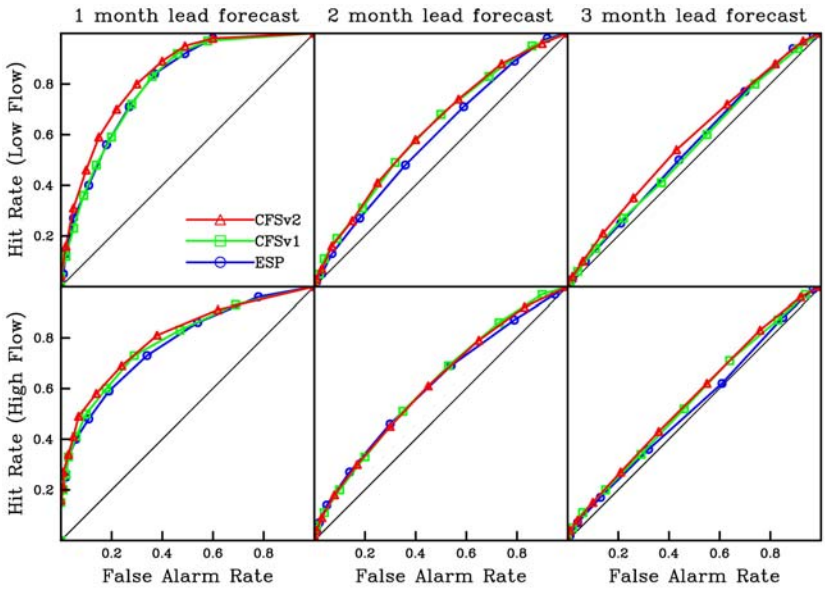


Fig. 6 Relative Operating Characteristic (ROC) diagram for low and high flow forecasts averaged at 130 gauges in the first three months (Yuan *et al.* 2011b).

To investigate the performance for drought severity and duration forecasts, Severity-Area-Duration (SAD) plots were shown in Figure 7 for 3-month duration (Yuan *et al.* 2011b). Severity (S) is defined as $S = (1-\Sigma P/t)*100\%$, where ΣP is the summary of monthly percentile of soil moisture over t months. Given that we were validating drought forecast, the SAD plot was a little different from its traditional way (Andreadis *et al.* 2005). We defined the drought grid cells based on the offline simulation results, no matter whether they were under drought or not in the forecasts. Therefore, the severity was used to quantify the difference between the models' forecasted droughts and the offline simulation conditioning on drought area. The offline simulated severity values were around 0.9 (Figure 7), which was similar to previous studies for 3-month drought duration (Andreadis *et al.*, 2005). Except for 1988, central U.S. drought forecasts had very low skill (Figure 7b) due to the underestimated drought area. CFSv1 and CFSv2 provided more accurate severity values than ESP (Figure 7a,c-d), indicating the added values from climate forecast besides initial conditions. Unlike monthly drought area analysis where the forecast could produce larger drought areas than the offline simulation, the SAD plot demonstrated that the forecasted severity values were generally lower than offline simulation due to the under-prediction of drought areas and/or intensities. For the 1988 drought, all three forecast approaches under-predicted severity compared to the offline simulation by about 40%. Averaged over the other three droughts, ESP underestimated the severity by 31%, CFSv1 by 24%, and CFSv2 by 15%.

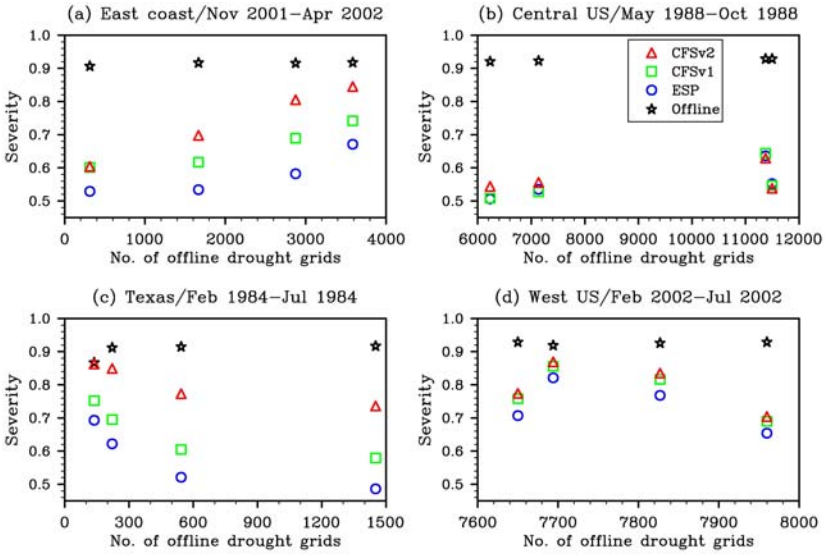


Fig. 7 Severity-Area-Duration (SAD) plots for 3-month duration. The 3-month drought grid cells are identified from offline simulation (Yuan *et al.* 2011b).

4. Conclusion

We provided a first look at the capability of the NCEP's latest operational seasonal forecast model CFSv2 by comparing its hindcast forecast skill with the CFSv1, ENSEMBLES EUROSIP models. CFSv2 shows significant skill enhancement for land surface air temperature and precipitation from CFSv1 for month-1 forecasts, and has comparable result to ECMWF, where the former and the latter have slightly higher skill in temperature and precipitation respectively (Yuan *et al.* 2011a).

CFSv1 and CFSv2-based seasonal hydrologic forecasts were generally more skillful than ESP in the first three months, and CFSv2 increased the skillful forecast percentage from CFSv1 by 10% on average. The climate model-based approach could outperform ESP out to three months in identifying low flow, but not for high flow beyond one month. CFSv2 had better discrimination than CFSv1 for the month-1 streamflow forecast. The offline simulated soil moisture was used to validate short-term drought forecasts. The SAD plots for 3-month duration illustrated the underestimation of drought severity, but the CFSv2 had the least error (Yuan *et al.* 2011b).

Acknowledgements. The research was supported by the NOAA Climate Program Office through Grants NA17RJ2612 and NA10OAR4310246.

References

Andreadis, K. M., E. A. Clark, A. W. Wood, A. F. Hamlet, and D. P. Lettenmaier, 2005: 20th Century drought in the conterminous United States, *J. Hydrometeorol.*, **6**, 985– 1001.

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.

Luo, L., E. F. Wood, and M. Pan, 2007: Bayesian merging of multiple climate model forecasts for seasonal hydrological predictions. *J. Geophys. Res.*, **112**, D10102, doi:10.1029/2006JD007655.

Luo, L., and E. F. Wood, 2008: Use of Bayesian merging techniques in a multimodel seasonal hydrologic ensemble prediction system for the eastern United States. *J. Hydrometeorol.*, **9**, 866-884.

Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen, 2002: A long-term hydrologically-based data set of land surface fluxes and states for the conterminous United States, *J. Climate*, **15**, 3237– 3251.

Mitchell, T. D., and P. D. Jones, 2005: An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int. J. Climatol.*, **25**, 693-712.

Saha, S., and Coauthors, 2006: The NCEP climate forecast system. *J. Climate*, **19**, 3483–3517, doi:10.1175/JCLI3812.1.

Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Am. Meteorol. Soc.*, **91**, 1015-1057.

Saha, S. and Coauthors, 2011: The NCEP Climate Forecast System Version 2, to be submitted to *J. Climate* (http://cfs.ncep.noaa.gov/)

Wang, W., M. Chen, and A. Kumar, 2010: An assessment of the CFS real-time seasonal forecasts. *Weather and Forecasting*, **25**(3), 950-969.

Weisheimer, A., and Coauthors, 2009: ENSEMBLES: A new multi-model ensemble for seasonal-to-annual predictions-Skill and progress beyond DEMETER in forecasting tropical Pacific SSTs. *Geophys. Res. Lett.*, **36**, L21711, doi:10.1029/2009GL040896.

Xia, Y., K. Mitchell, M. Ek, J. Sheffield, B. Cosgrove, E. F. Wood, L. Luo, C. Alonge, H. Wei, J. Meng, B. Livneh, D. Lettenmaier, V. Koren, Q. Duan, K. Mo, Y. Fan, and D. Mocko, 2011: Continental-scale water and energy flux analysis and validation for the North-American Land Data Assimilation System Project Phase 2 (NLDAS-2), Part I: Intercomparison and application of model products. *J. Geophys. Res.* (submitted).

Yuan, X., E. F. Wood, L. Luo, and M. Pan, 2011a: A first look at Climate Forecast System version 2 (CFSv2) for hydrological seasonal prediction. *Geophys. Res. Lett.*, **38**, L13402, doi:10.1029/2011GL047792.

Yuan, X., E. F. Wood, J. K. Roundy, M. Pan, and L. Luo, 2011b: CFSv2-based seasonal hydrologic forecasts over conterminous United States. *J. Geophys. Res.* (to be submitted)

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Low-Frequency SST variability in CMIP5 Historical Integrations

Lydia Stefanova and Timothy LaRow

Center for Ocean-Atmospheric Prediction Studies, The Florida State University

1.Introduction

Anthropogenic-driven change in hurricane frequency is obscured, among other things, by low-frequency natural sea surface temperature (SST) variability. Our goal is to set up a modeling system to estimate changes in North Atlantic hurricane frequency in response to changing climate, separating the anthropogenic from the natural-variability components of the projected SST warming. Prerequisites for this undertaking are: an atmospheric general circulation model with demonstrated skill in reproducing the observed North Atlantic hurricane frequency and trend, and SST projections from a global model with a degree of realism in its representation of the physical mechanisms for low-frequency variability.

The FSU/COAPS AGCM satisfies the first of these prerequisites. LaRow *et al.* (2008, 2010) have documented the skill of the FSU/COAPS AGCM in simulating the observed historical hurricane frequency when driven with either observed or model-produced (namely, CFS) SSTs. The FSU model accurately simulates the interannual variability in response to time varying SSTs, and is also able to reproduce the trend in Atlantic hurricane counts since 1982.

One theory suggests that the North Atlantic hurricane activity is strongly modulated by the phase of the Atlantic Multidecadal Oscillation (AMO) (Goldenberg *et al.* 2001). However, climate model estimates of future hurricane activity have large uncertainties; one source for the uncertainty is changes in low frequency (multidecadal) SST variability. In addition, the low frequency SST variability is difficult for many climate models to correctly simulate.

In order to address the low-frequency modulation of hurricane variability in a future climate, we need SST projections from the "best" IPCC AR5 CMIP5 model simulations to use as boundary forcing in the FSU/COAPS atmospheric model. To identify the "best" models**,** we evaluate CMIP5 historical baseline (1850-2005) simulations. Our assessment focuses on the models' 20th century SST trend and ENSO-, and AMO-related variability.

2. Data and methodology

Table 1 shows the models that are currently available (as of October 2011) from the IPCC CMIP5 historical experiment (obtained from http://pcmdi3.llnl.gov/esgcet/home.htm). Also shown in the table are the models' native horizontal resolutions and the number of ensemble members. Before conducting the analysis on the models' sea surface temperatures, the horizontal resolution of each model was re-mapped to a 360x180 uniform resolution. We analyze the low-frequency component of

Center	Model	Resolution	Ensemble size
Met Office Hadley Centre	HadGEM2	360x216	1
Canadian Centre for Climate Modeling and Analysis	CanESM2	256x192	5
Norwegian Climate Centre	NorESM1-M	320x384	2
NASA Goddard Institute for Space Studies	GISS-E2-H	144x90	5
NASA Goddard Institute for Space Studies	GISS-E2-R	144x90	1
Institute for Numerical Mathematics (Russia)	INMCM4	360x340	1
Institut Pierre-Simon Laplace (France)	IPSL-CM5A-LR	182x149	2
Australian Commonwealth Scientific and Industrial Research Organization	CSIRO-Mk-3-6.0	189x192	10

Table 1 List of the CMIP5 models used in this assessment and their native resolution and ensemble size.

Correspondence to: Lydia Stefanova, Center for Ocean-Atmospheric Prediction Studies, The Florida State University, Tallahassee, FL; E-mail: lstefanova@fsu.edu.

SSTs in the CMIP5 historical simulations by comparing the 20th Century trend, the Atlantic Multidecadal Oscillation (AMO) and the ENSO characteristics (spatial distribution, frequency and magnitude) of the member models' SSTs to the corresponding observed characteristics during the period.

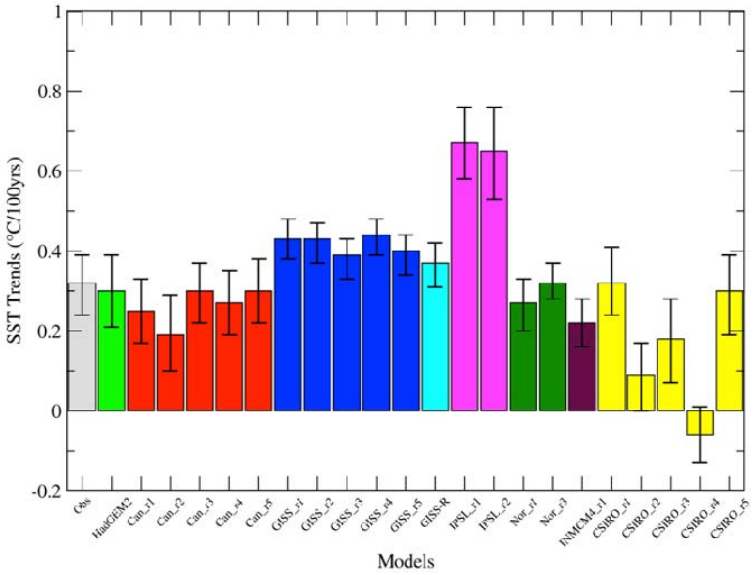


Fig. 1 North Atlantic SST trends in observations (grey bar) and CMIP5 models (colored bars; color coding as in Table 1

3. Results

3.1 Trend

All available models reproduce the observed increase in SSTs in both the North Atlantic (Fig. 1) and tropical Pacific (not shown). Models show marked increase in warming (~4x) in the simulated (RCP4.5) 21st Century compared to the 20th Century Historical simulations in the North Atlantic and tropical Pacific Oceans (not shown). The largest difference between the models' trends during the 21st Century is in the tropical Pacific.

3.2 ENSO

All models have a strong tropical Pacific signal in the 1.5-8 year band (Fig. 2). In the observations, the first complex EOF (CEOF1) accounts for 54.5% of the total variance for this frequency band. Amongst the CMIP5 model runs, this percentage ranges from 30.7% (INMCM4-r1) to 62.9 (NorESM1-M-r1). The variance explained by CEOF1 differs among the individual realizations of a given model by up to 5.6 percentage points. The mean observational ENSO period is 3.7 years. The AR5 models' ENSO period is generally underestimated, varying from a rapid return period of 2.7 years (CanESM2-r3) to 3.9 years

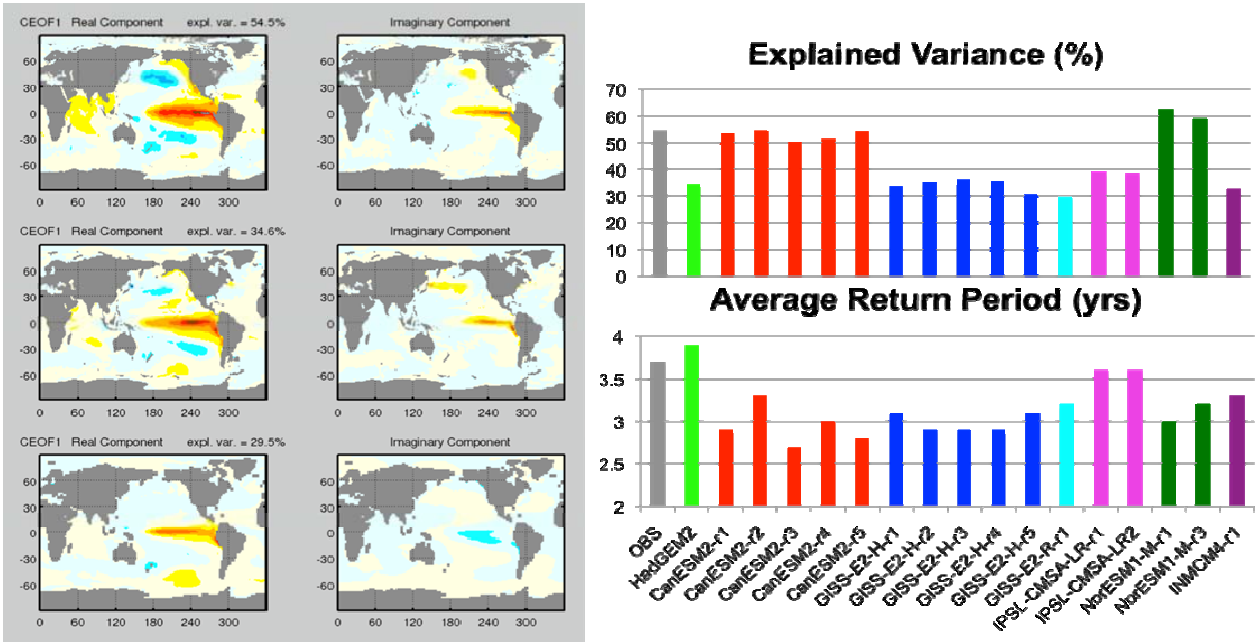


Fig. 2 Left: Real and imaginary components of the leading Complex EOF for the observed (top), best model (center) and worst model (bottom) historical period SSTs. Right: Explained variance and average return period of ENSO.

(HadGEM2). Individual realizations differ in their estimate of the average return frequency of ENSO by up to 0.5 years.

3.3 AMO

Most (but not all) of the analyzed models have a pronounced AMO signal in two dominant frequency bands ~70 years and ~25 years. In the observed data (HadISSTv1.1, Rayner *et al.* 2003), both frequencies are found in PC1 of the de-trended North Atlantic SSTs, while in the model data they tend to be distributed between PC1-3 (Fig. 3).

Some models (*e.g.*, GISS (R and H) models and NorESM1) have a very weak multi-decadal signal in the North Atlantic. Others (e.g. IPSL and HadGEM2) overestimate the power spectrum. The rest of the models are more or less comparable to observations.

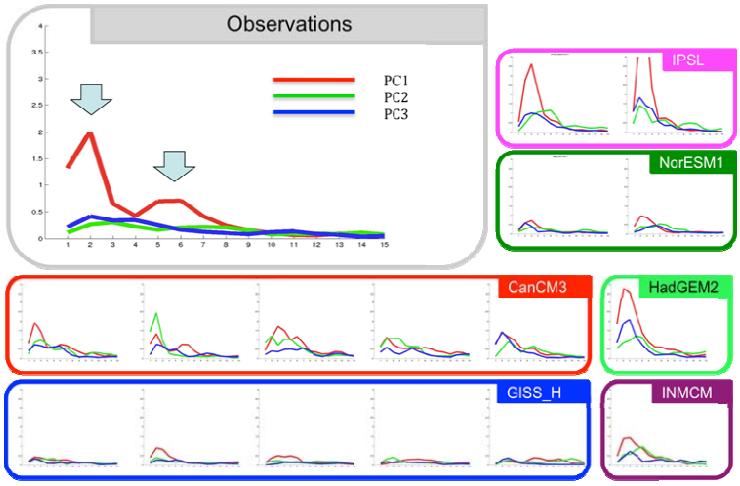


Fig. 3 Power spectra of the first three principal components of North Atlantic SSTs for observations and historical simulations. The abscissa denotes number of cycles per 140 years. Blue arrows indicate the observed spectral peaks at ~70 and ~25 years.

4. Summary and conclusions

All models have strong tropical Pacific SST variability in the 1.5-8 year range. The average model ENSO return period ranges from 2.7 to 3.9 years, compared to the observed 3.7 years. The models with most realistic ENSO representation are the CanESM2 and HadGEM2.

Most models have pronounced multi-decadal Atlantic variability with dominant modes ~70 and ~25 years. In the observations, both modes are contained in the Atlantic EOF/PC1; in the models, they are spread between EOF/PC1, 2 and 3. The best AMO representation amongst the models analyzed thus far is that of the CanESM2.

In addition, some models have spurious signal in the high latitudes of the southern hemisphere. Most models have pronounced multidecadal Pacific variability that we have not yet analyzed in detail.

As historical simulations from additional CMIP5 models become available, we will update the above evaluation. Our preliminary conclusion is that of the models available thus far, CanESM2 has the winning combination of trend, ENSO and AMO variability and its SSTs would be the most suitable for addressing Atlantic basin hurricane projections for the mid- to late-21st century.

Acknowledgements. This research is supported by a grant from DOE-CESD.

References

Goldenberg, S. B., C. W. Landsea, A. M. Mestas-Nuñez, and W. M. Gray, 2001: The recent increase in Atlantic hurricane activity: Causes and implications, *Science*, **293**, 474–479.

LaRow, T. E., Y. -K. Lim, D. W. Shin, E. Chassignet and S. Cocke, 2008: Atlantic Basin Seasonal Hurricane Simulations. *J. Climate*, **21**, 3191–3206.

LaRow, T. E., L. Stefanova, D. W. Shin and S. Cocke, 2010: Seasonal Atlantic Tropical Cyclone Hindcasting/Forecasting using Two Sea Surface Temperature Datasets. *Geophys. Res. Lett.*, **37**, L02804, doi:10.1029/2009GL041459.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L.V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 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 10.1029/2002JD002670

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Implementation of Land Information System in the NCEP Operational Climate Forecast System CFSv2

Jesse Meng[1], Rongqian Yang[1], Helin Wei[1], Michael Ek[1], and Pingping Xie[2]

[1]*Environmental Modeling Center, NCEP/NWS/NOAA, Camp Springs, MD*

[2]*Climate Prediction Center, NCEP/NWS/NOAA, Camp Springs, MD*

To support NOAA's mission in improving its climate prediction on the seasonal scale, NCEP has launched a new version of the NCEP Climate Forecast System (CFSv2) in 2011. The NASA Land Information System (LIS, Peters-Lidard *et al.*, 2007) is implemented within CFSv2 to execute the Global Land Data Assimilation System (GLDAS) to provide land surface analysis.

Numerical weather and climate prediction and assimilation systems generally use prescribed or simulated soil moisture values to formulate surface layer parameterizations within a land surface model (LSM) to calculate evapotranspiration and latent and sensible heat fluxes, and to solve for the land surface energy and water balance equations. In the first version of the NCEP CFS (CFSv1), the Oregon State University (OSU) LSM (Pan and Mahrt, 1987) was used to calculate the land surface states and fluxes. To outcome the deficiencies in the precipitation forcing from the atmospheric model and to prevent the calculated soil moisture drifting too far from a preferred climatology, an adjustment on the simulated soil moisture was required to compensate the differences between the simulated and observed precipitation. The Noah LSM (Ek *et al.*, 2003) was developed with substantial improvements from the OSU LSM as a collaborative effort between NCEP and numerous partners from the land hydrology community over the last decade. Noah was implemented in the NCEP North American Mesoscale Model (NAM) in 2003 and the NCEP Global Forecast System (GFS) in 2005 for operational regional and global medium-range weather forecasts. In CFSv2, Noah is implemented in both the fully-coupled land-atmosphere-ocean prediction model to make seasonal climate forecast and the semi-coupled GLDAS to perform the land surface analysis.

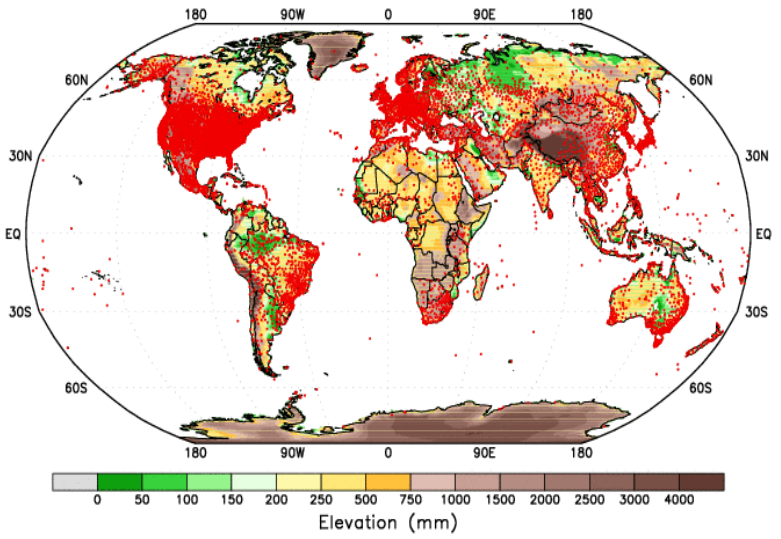


Fig. 1 Example of global gauge distribution used in the CPC unified global daily gauge analysis.

The NASA LIS is employed to execute the CFSv2 land surface analysis. Compared to CFSv1, this CFSv2 LIS uses observed global precipitation analyses as direct forcing to drive the Noah LSM to execute the land surface analysis. The adjustment on soil moisture is based on land surface physics responding to the adjustment on precipitation forcing, rather than the artificial adjustments previously applied in CFSv1. Two sets of observation-based global precipitation analyses are utilized as alternative forcing to drive the CFSv2 LIS. One is the pentad data of CMAP (Xie and Arkin, 1997) that also used in CFSv1; the other is the CPC unified global daily gauge analysis (Xie *et al.*, 2010). After reviewing the global gauges distribution (Figure 1), a merging approach with latitude-dependent weighting masks is designed that favors the gauge-only

Correspondence to: Jesse Meng, Environmental Modeling Center, NCEP/NWS/NOAA, Camp Springs, MD 20746; E-mail: Jesse.Meng@noaa.gov.

analysis in mid-latitudes, and the satellite-dominated CMAP analysis in tropical latitudes where gauge count is low. In high-latitudes where gauge count is also low and satellite estimate lacks of accuracy, the CFSv2 model predicted precipitation is given the most weights. In addition, the gauge analysis is given the most weight in the regions of high gauge density (*e.g.*, contiguous United States, Europe, and Australia).

In summary, NCEP CFSv2 is for the first time an LDAS strategy is used in a coupled land-atmosphere-ocean global climate prediction system to perform the land surface analysis. Observed precipitation fields are used to constrain the land surface simulation of soil moisture, soil temperature, and snowpack. This product can improve the current understanding of the global land surface energy and water cycles and their roles in the Earth system to further assist the operation and research in hydrology, weather, and climate.

References

Ek., M., K. E., Mitchell, Y Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land-surface model advances in the NCEP operational mesoscale Eta model. *J. Geophys. Res.*, **108**(D22), 8851, doi:10.1029/2002JD003296.

Pan, H.-L., and L. Mahrt, 1987: Interaction between soil hydrology and boundary-layer development. *Bound.-Layer Meteor.*, **38**, 185-220.

Peters-Lidard , C.D., P.R. Houser, Y. Tian, S.V. Kumar, J. Geiger, S. Olden, L. Lighty, B. Doty, P. Dirmeyer, J. Adams, K. Mitchell, E.F. Wood, and J. Sheffield, 2007: High-performance Earth system modeling with NASA/GSFC's Land Information System. *Innovations in Systems and Software Engineering*, **3**(3), 157-165. doi:10.1007/s11334-007-0028-x.

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.

Xie, P., M. Chen, A. Yatagai, T. Hayasaka, Y. Fukushima, and S. Yang, 2007: A gauge-based analysis of daily precipitation over East Asia. *J. Hydrometeor.*, **8**, 607 – 626.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Problem of Cloud Overlap in Radiation Process in JMA Global NWP Model

Ryoji Nagasawa

Climate Prediction Division, Japan Meteorological Agency, Tokyo, Japan

1. Introduction

The JMA global NWP model (JMA-GSM) tends to be optically thicker (thinner) in the tropics (extratropics) for the shortwave radiation compared with observations (Fig. 1). One of the causes is an insufficient treatment of a cloud overlap in the shortwave radiation calculation. To improve the treatment, the problem and solution is studied.

2. Problem of current cloud overlap in shortwave radiation calculation

In a shortwave radiation scheme of JMA-GSM, total cloud fraction is calculated by assuming a maximum-random overlap treating a column area as clear sky and cloudy areas separately. In the cloudy area of the column, a random overlap is always adopted to treat cloud multiple scattering effects. In the longwave radiation calculation, maximum-random overlap is adopted (NPD/JMA 2007).

In a case that a fraction of optically thin high level clouds (anvil) is large and a fraction of tower-shaped cumulus is small, which is often observed in the tropics, cloud optical thickness is overestimated in the shortwave radiation calculation (Fig. 2).

3. The solution

An Independent Column Approximation (ICA, *e.g.*, Cahalan *et al.* 1994) can take the cloud multiple scattering effects into account in shortwave radiation calculation (Fig. 3). This makes it possible to improve the aforementioned problem and to adopt maximum-random overlap both in the shortwave and longwave radiation

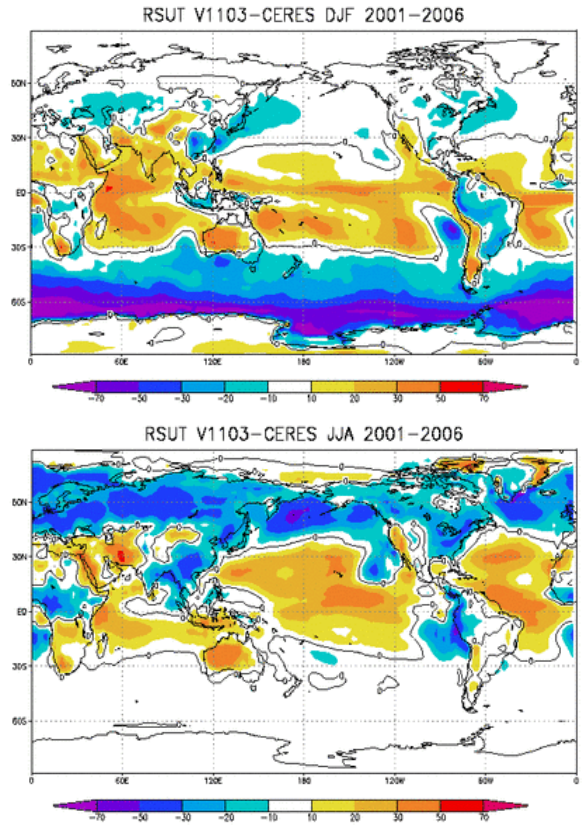


Fig. 1 Upward shortwave radiation flux at TOA (JMA-GSM - CERES) (Wm^{-2}). Left: JJA, right: DJF. 2001-2006 year average is shown.

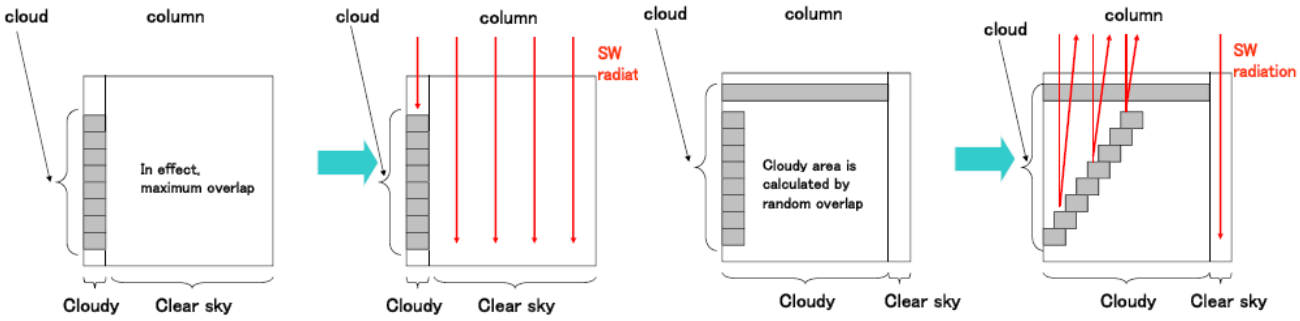


Fig. 2 Pattern diagrams of the case that cloud optical thickness is overestimated by current cloud overlap in shortwave radiation scheme. Left: Without anvil, Right: With anvil.

Correspondence to: Ryoji Nagasawa, Climate Prediction Division, Japan Meteorological Agency, Tokyo, Japan; E-mail: r-nagasawa@met.kishou.go.jp.

calculation. Although full ICA requires more computational cost than the current scheme, an efficient method was proposed by Collins (2001) without degrading accuracy of the radiation computation (Practical ICA: PICA). Essence of PICA is to ignore radiation calculation in sub-column whose contribution (width) is small (narrow).

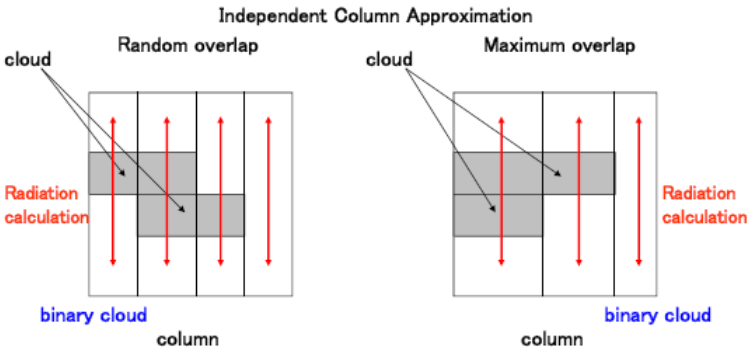


Fig. 3 Pattern diagrams of ICA (Independent Column Approximation) that only two clouds overlaps vertically. Left: in case of random overlap, Right: in case of maximum overlap.

4. Results

Fig. 4 and Fig. 5 show impacts of PICA with maximum-random overlap on JMA-GSM and simulated cloud distribution. Initial time of experiment is 12UTC on 10 Aug 2009 and 24 hour forecast was executed. Horizontal gird spacing and number of vertical layer of JMA-GSM are about 200 km and 60 respectively.

PICA with maximum-random overlap decreases cloud optical thickness around tropics and mid-latitude (Fig. 4). In addition PICA decreases shortwave heating in middle troposphere and increases that in lower

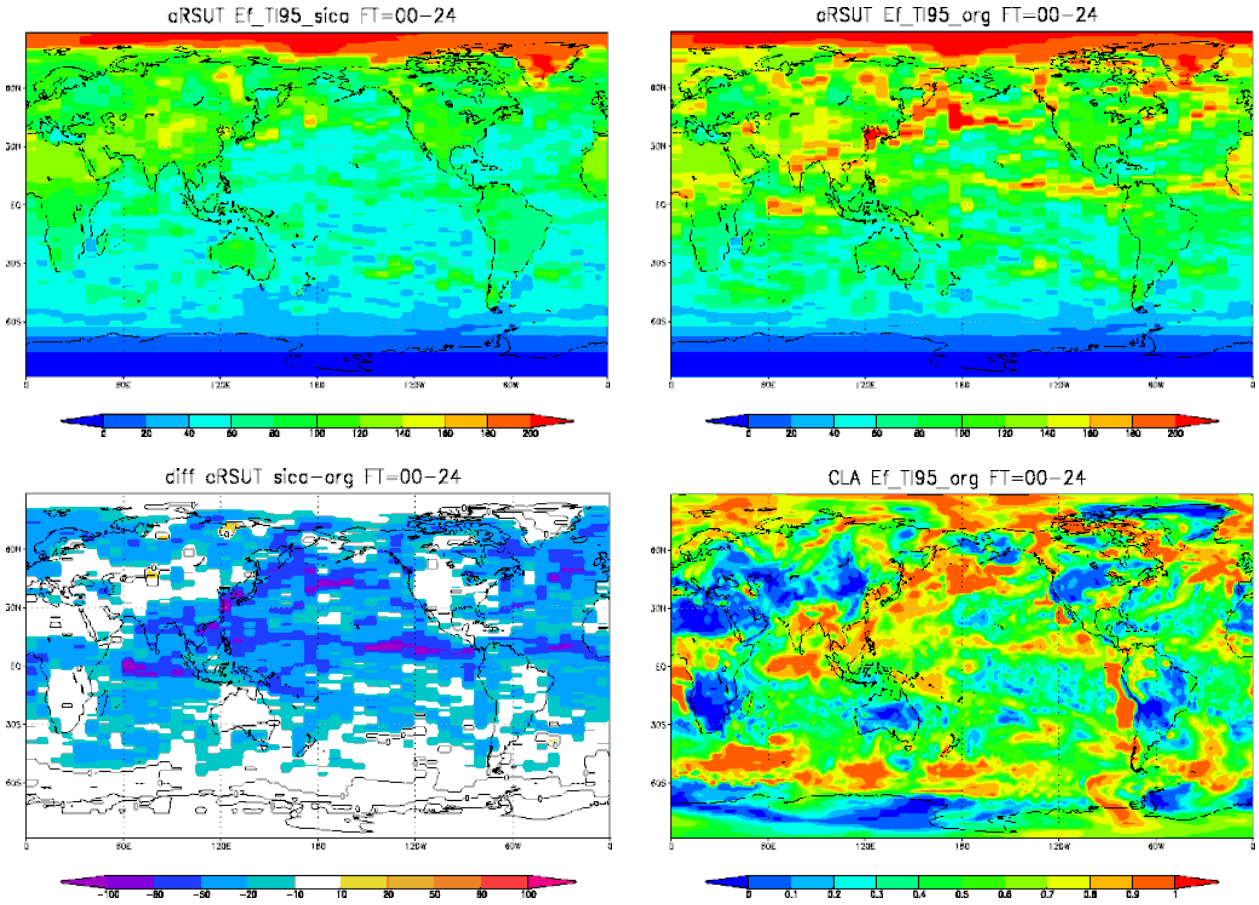


Fig.4 Impact of difference of cloud overlap assumption to upward shortwave radiation flux at TOA (Wm^{-2}). Upper left: TEST, upper right: CNTL, lower left: TEST – CNTL, lower right: total cloud fraction. Initial time is 12UTC on 10 Aug 2009. 24 hour forecast average is shown.

troposphere (Fig. 5). It is thought that decrease of excess shortwave radiation flux reflection in middle troposphere induces increase of downward shortwave radiation flux to lower troposphere and shortwave radiation absorption by cloud and water vapor below 900hPa.

5. Summary

JMA-GSM tends to be optically thicker (thinner) in the tropics (extratropics) for shortwave radiation flux compared with observation. One of the causes is an insufficient treatment of cloud overlap in the shortwave radiation calculation. Practical ICA, improved method to treat better cloud overlap in the shortwave radiation calculation with small computational cost is tested. The method with maximum-random overlap decreases cloud optical thickness extensively and increases (decrease) shortwave heating in middle (lower) troposphere. Practical ICA has to be tested in many cases and appropriate parameters have to be fixed considering computational cost and accuracy. And then description of cloud simulated by JMA GSM has to be improved.

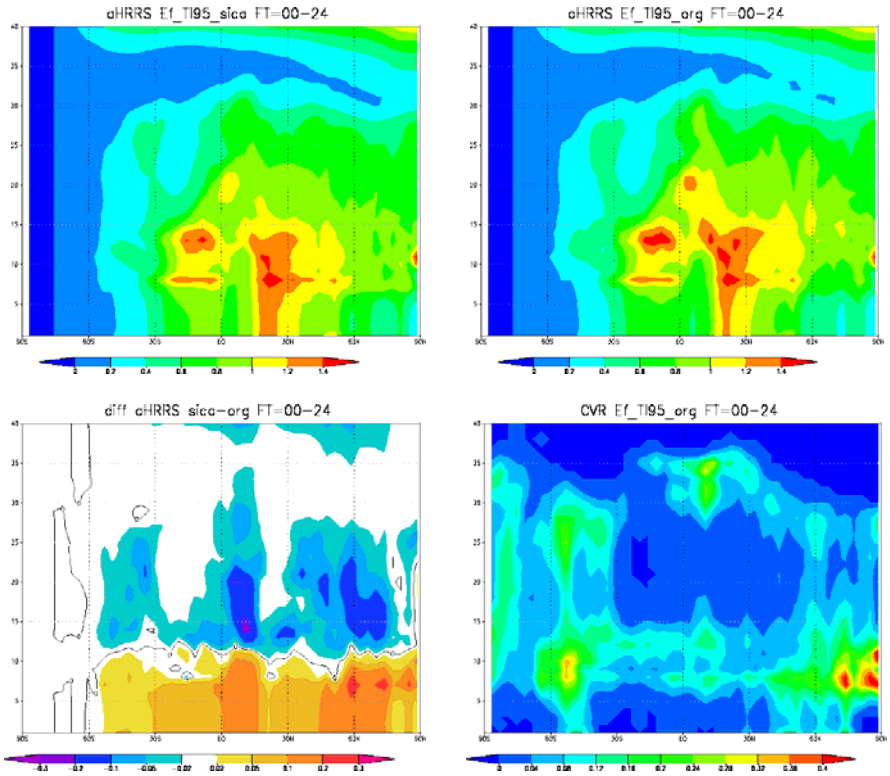


Fig. 5 Impact of difference of cloud overlap assumption to shortwave radiation heating rate (K/day). Upper left: TEST, upper right: CNTL, lower left: TEST – CNTL, lower right: cloud fraction. Initial time is 12UTC on 10 Aug 2009. Zonal mean and 24 hour forecast average are shown.

References

Cahalan, R.F., W. Ridgway, W.J. Wiscombe, S. Gollmer, and Harshvardhan, 1994: Independent pixel and Monte Carlo estimates of stratocumulus albedo. *J. Atmos. Sci.*, **51**, 3776-3790.

Collins, W.D., 2001: Parameterization of Generalized Cloud Overlap for Radiative Calculation in General Circulation Models. *J. Atmos. Sci.*, **58**, 3224-3242.

Numerical Prediction Division / Japan Meteorological Agency, 2007: Outline of the operational numerical weather prediction at the Japan Meteorological Agency. 194pp.
Available online at http://www.jma.go.jp/jma/jma-eng/jma-center/nwp/outline-nwp/index.htm

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Preliminary Evaluation of Multi-Model Ensemble System for Monthly and Seasonal Prediction

Qin Zhang[1], Huug van den Dool[2], Suru Saha[3], Malaquias Pena Mendez[3], Emily Becker[1], Peitao Peng[2] and Jin Huang[2]

[1]*Wyle Information System, Climate Prediction Center, NCEP/NWS/NOAA*
[2]*Climate Prediction Center, NCEP/NWS/NOAA*
[3]*Environment Modeling Center, NCEP/NWS/NOAA*

ABSTRACT

A collaborative prediction system, the National Muli-Model Ensemble (NMME), is under development through the NOAA Climate Test Bed (CTB) for experimental monthly and seasonal prediction at Climate Prediction Center (CPC). The CTB NMME project is funded by Climate Program Office MAPP Program (Modeling, Analysis, Prediction and Projection). In the current phase, seven models from different US institutes (NCEP-CFSv1, NCEP-CFSv2, GFDL-CM2.2, NCAR/U.Miami/COLA-CCSM3, NASA-GEOS5, IRI (ECHAM-a and ECHAM-f)) are participating. Three variables (monthly mean precipitation, sea surface temperature, and air temperature at 2 meters on a 1x1 degree grid), all with at least 29 years of hindcasts (1982-2010), are evaluated after removing their systematic errors, and then verified against the observations. Realtime experimental forecasts of the multi-model ensemble were first conducted in August 2011. The bias corrected multi-model ensemble prediction system is designed to contribute to the ongoing monthly and seasonal prediction in CPC.

1. Introduction

Monthly-to-seasonal time scale climate predictions are made at NCEP/CPC routinely by a number of tools, both statistical, *e.g.* canonical correlation analysis (CCA) and optimal climate normal (OCN) (O'Lenic *et al.* 2008), and dynamical models (Climate Forecast System version 1 &2). Although the ratio of signal to noise is low in the dynamical model due to long time integration and growth of the systematic error (Straus and Shukla 2002), the dynamical models have comparable forecast scores to the statistical models (DeWitt 2005; Saha *et al.* 2006). Successful monthly-to-seasonal prediction mostly depends on a revolution in our understanding of the coupled ocean-atmosphere system after the dramatic strong ENSO events in 1982/83 and 1997/98 (Barnston *et al.* 1999; Landsea and Knaff 2000; Shukla *et al.* 2009). This means monthly-to-seasonal predictability relies on the slowly evolving components of the climate system, like the ocean or land surface, that act as boundary conditions for the atmosphere with its shorter intrinsic time scales (Shukla *et al.* 2009; Goddard et al. 2001; Paolino *at al.* 2011). Two types of uncertainties are involved in the monthly-to-seasonal predictability: one is related to the uncertainty of the initial conditions (Keenlyside *et al.* 2005; Luo and Wood 2006), and the other is accounted for model errors in the physics processes related to the sub-grid parameterization (Palmer *at al.* 2004, Kirtman and Min 2009, De Witt 2005).

In recent years, the multimodel ensemble forecast has become a powerful tool for the monthly-to-seasonal time scale prediction to deal with both uncertainties (Krishnamurti *et al.* 2000, Kirtman *et al.* 2003; Peng *et al.* 2002; Hagedorn *et al.* 2005; Doblas-Reyes *et al.* 2005; Palmer *et al.* 2004; Lavers *et al.* 2009). For the monthly-to-seasonal forecast, multimodel prediction has successfully increased the spread by reducing the overconfident forecast of the individual model (Palmer *et al.* 2004; Weisheimer *et al.* 2009). Furthermore, recently research has shown that a multimodel ensemble, even in a simple equal weight combination, has

Correspondence to: Qin Zhang, Wyle Information System, Climate Prediction Center, NCEP/NWS/NOAA, Camp Springs, MD 20746; E-mail: Qin.Zhang@noaa.gov.

higher prediction skill scores than that of any individual model in the prediction of tropical SST anomaly (Kirtman and Min 2009). Several projects, like DEMETER and EUROSIP, have demonstrated the improvement of multimodel seasonal forecast reliability (Hagedorn *et al.* 2005, Mitchell *et al.* 2004, Palmer *et al.* 2004).

This work of reports the latest results from the experimental National Multi-Model Ensemble (NMME) prediction system (Phase1). The models and data involved in Phase 1are briefly described in section 2, and followed by a preliminary evaluation of the multimodel ensemble system for monthly and seasonal prediction. In the last section, we present a summary and discussion.

Model	Period	Members	Leads	Arrangement of Members
CFSv1	1981-2009	15	0-8 months	1st 0Z +/-2days, 21st0Z+/-2d, 11th0Z+/-2d
CFSv2	1982-2009	24(28)	0-9	4 members (0,6,12,18Z) every 5th day
GFDL-CM2.2	1982-2010	10	0-11	All 1st of the month 0Z
IRI-Echam4-f	1982-2010	12	0-7	All 1st of the month
IRI-Echam4-a	1982-2010	12	0-7	All 1st of the month
CCSM3.0	1982-2010	6	0-11	All 1st of the month
NASA-GEOS5	1982-2010	6 (8)	0-8	1 member every 5th day Additional 2 members on the beginning of month

Table 1 NMME models information

2. Models and data

Based on two Climate Test Bed (CTB) workshops (February 18 and April 8, 2011), a collaborative and coordinated implementation has been established under the frame work of the CTB project, called National Multi-Model Ensemble (NMME). The experimental realtime experimental forecast system made a first seasonal and monthly multimodel forecast in August 2011 in Climate Prediction Center as Phase 1 of the NMME project. Seven models, from NCEP (CFSv1&2), GFDL (CM2.2), IRI (ECHAM-a and ECHAM-f), NCAR/U.Miami/COLA-CCSM3 (Collins *et al.* 2006), and NASA-GEOS5 are participating. Three variables (monthly mean precipitation, sea surface temperature, and air temperature at 2 meters on a 1X1 degree grid) with 29 years of hindcasts (1982-2010), have been evaluated. The model climatology and prediction skill mask have been calculated after the systemic errors are corrected for each model in every leading month forecasts. More details of the NMME models are given in Table1.

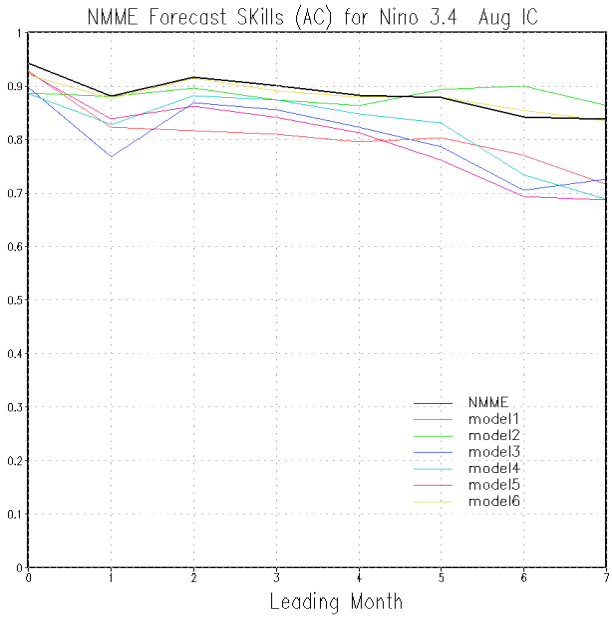


Fig. 1 Multimodel ensemble forecast skills of Nino3.4 (black line) and individual modes (color lines) for 7 lead months with August initial conditions.

The observation data used for verification are the NOAA Optimum Interpolation (OI) Sea Surface Temperature (SST) V2 (SST OISST-QD) (1982-2010) (Reynolds *et al.* 2002) for verification of model SST, CMAP (1982-2010) (Xie and Arkin 1997) for precipitation, and GHCN_CAMS (1982-2010) for temperature at 2m (Fan and van den Dool 2006).

3. Results

The first experimental realtime monthly-to-seasonal forecast of NMME-PHASE1 was made in August 2011. In the month before, hindcasts from 1982-2010 with August ICs for all models, except NASA's, were collected by FTP for all 3 variables: SST, precipitation and temperature at 2m. The evaluation of prediction skills for each model and equal-weight average of multimodel ensemble mean were done after calculating the model climatology for the systematic bias correction. Here, we are reporting some results related to the NMME prediction skill assessment for the subsequent realtime forecast in August.

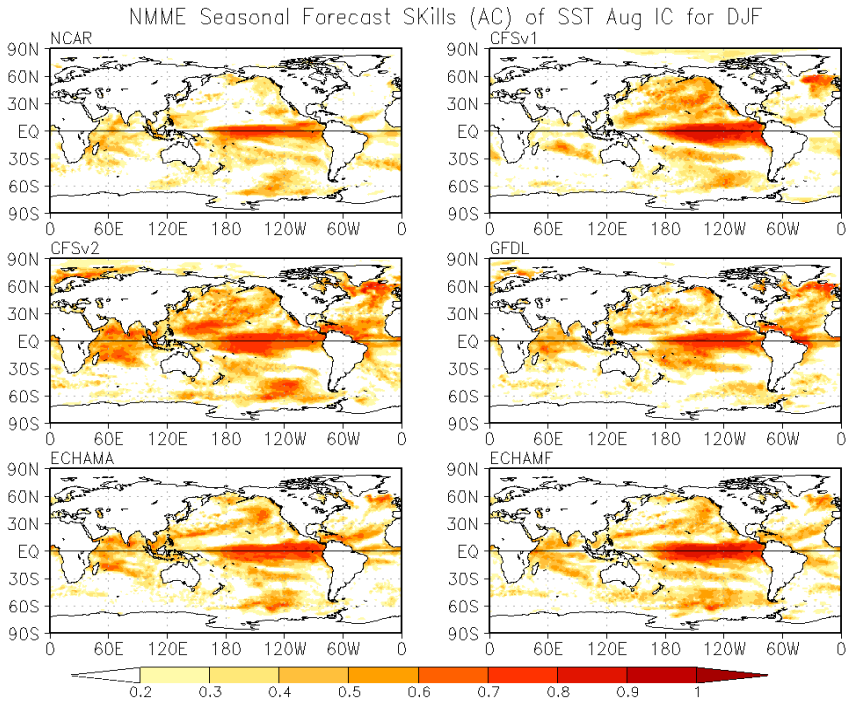


Fig. 2a Maps of anomaly correlation coefficient (ACC) of SST for individual models for prediction DJF with August ICs.

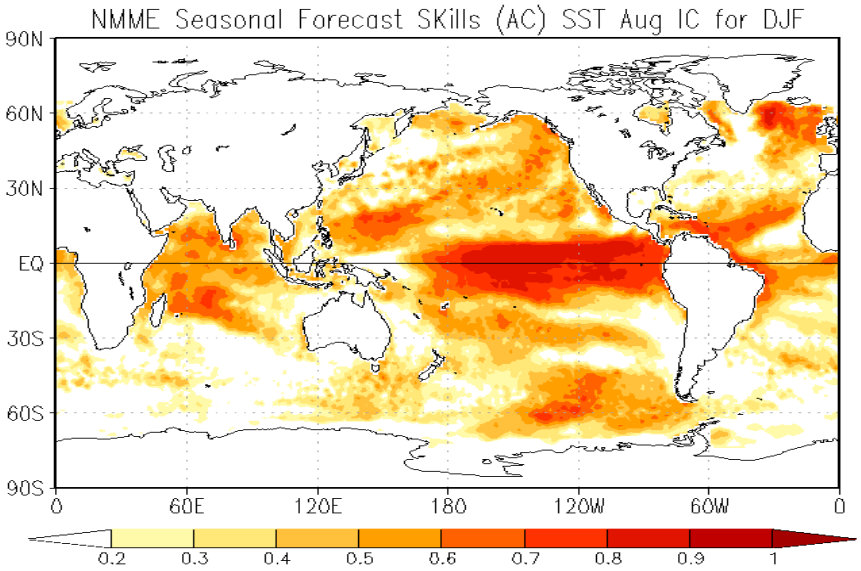


Fig. 2b SST anomaly correlation coefficient (ACC) of multimodel ensemble with observation (1982-2010) for prediction DJF with August ICs.

a. Nino3.4 and SST

Since SSTs in the tropical Pacific are a major source of climate predictability on monthly-to-seasonal time scales, model performance in the tropical Pacific is of particular interest. To demonstrate the typical level of skill in this area, Fig. 1 shows the anomaly correlation coefficient (ACC) of the ensemble mean for the single model ensemble (colored lines) and multimodel ensemble (black line) for the SST anomaly averaged over the Nino3.4 area for each lead in months. The results suggest that all single-model ensembles generally perform as well as El Nino-Southern Oscillation (ENSO) prediction systems. All single models have achieved above 0.7 for 7 months forecast, but the NMME multimodel ensemble system has an ACC above 0.83 for all the 7 months. In addition, note the higher correlation of the multimodel ensemble compared to all single models, except for the last two leads. This indicates the multimodel ensemble indeed has more skill than single model, as pointed out

by the previous studies (Palmer *et al.* 2004, Kertman and Min 2009). This improvement of the Nino3.4 prediction skill is greatly encouraging and has shown the value of the NMME efforts.

To further investigate the multimodel SST hindcast skill, maps of grid point ACC for the target season DJF with August initial conditions are shown for the six individual models (Fig. 2a) and the multimodel ensemble (Fig. 2b). Unsurprisingly, the ACCs of the SST anomaly for both the single models and the multimodel ensemble mean have higher scores over the central eastern Pacific, exceeding 0.8, and then gradually decrease along the equator westward and off the equator in the western tropical Pacific. It is interesting to see that while some models have low skills over the northern west Pacific, the NMME ensemble has skills comparable to the best performing model over the regions. Many areas like this can be found over the tropical Indian Ocean and northern Atlantic. In the extra-tropics, the correlations are generally low, but there are same notable high correlations (*e.g.* greater than 0.6) in the north Atlantic and South Pacific.

Carefully comparing each map in Fig. 2a and Fig. 2b, one may find that the ACC of the multimodel ensemble seems to take over skill scores wherever they are better among the single models. This hypothesis is confirmed by averaging the prediction skill over the global tropical band (from 30ºS-30ºN) shown in Fig. 3. The ACC of the multimodel is the second highest for the first 3 leads and the best one after that for longer lead month forecast. The root square mean error (RSME) is almost the lowest for all the leading months except the first month among all models. Consistent with DEMETER (Hagedorn *et al.* 2005), the NMME multimodel ensemble improvement in SST prediction achieved by the error compensation each other in individual models (Kirtman and Min 2009).

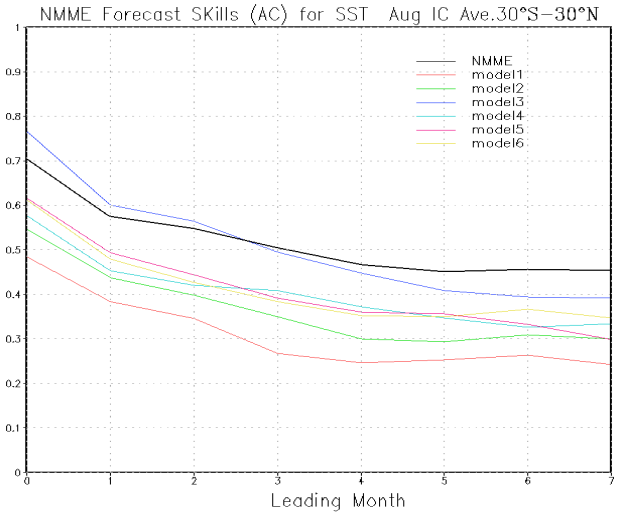


Fig. 3a NMME forecast skills (black line) and individual models (color lines) for SST averaged 30S-30N with August ICs.

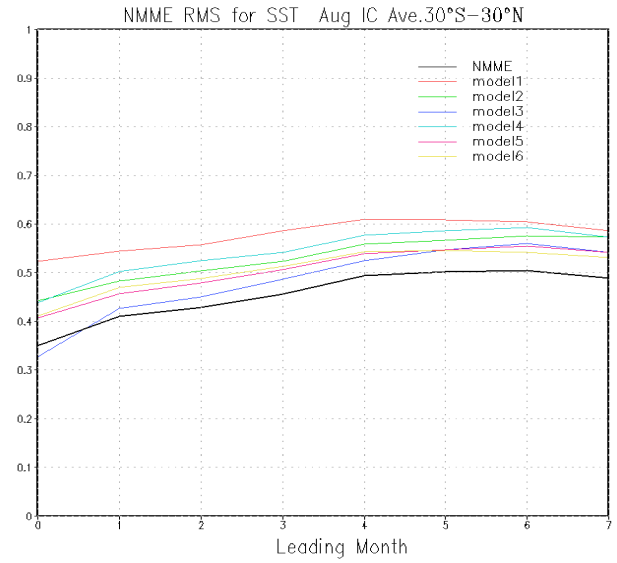


Fig. 3b NMME RMS error (black line) and individual models (color lines) for SST averaged 30S-30N with August ICs.

b. Precipitation

Corresponding to the high scores over the central eastern Pacific in SST anomaly prediction, the ACC of precipitation for each model and multimodel ensemble have the same narrow band of high forecast skills (Fig. 4) for DJF with August initial conditions. Scores quickly decrease westward and off the equator for the individual models and the multimodel ensemble. Only isolated scattered high score areas can be found in the extra-tropics. The ACC over the land is also low for the both individual models and multimodel ensemble. ACC scores averaged over the global tropical band are lower than 0.3 for all the leading months (Fig. 5a). However, the multimodel ensemble has the highest scores, and RSME is much smaller than for any of the single models (Fig. 5b). The superior performance of the multimodel ensemble for the prediction of precipitation indicates the benefit of the NMME approach even with a simple equal-weighting ensemble. This

achievement especially helps to improve the CPC's realtime monthly-to-seasonal forecast, since currently precipitation prediction scores are low on land and have large uncertainty.

c. Temperature at 2m

Now we explore the forecast skills of the atmospheric temperature at 2m for the northern hemisphere winter (DJF) in 4 months leading prediction. The strongest signals of the ACC are mainly over the tropical regions related to ENSO impacts on the monthly-to-seasonal time scales. Good scores can be found over South America and Africa in each individual model, indicating that state-of-the-art coupled models have caught the SST forcing and the right response to the forcing of the boundary conditions of ENSO (Fig. 6a). The multimodel ensemble has superior scores for these near equatorial regions, such as the east and west coasts of Australia and Sumatra in the western Indian Ocean and western equatorial Pacific (Fig. 6b). The prediction scores for temperature at 2m averaged over the tropical band (30ºS-30ºN, land only) for each lead show the multimodel ensemble is the second best compared to all models (Fig. 7a), and RMSE is almost as good as the best individual model for the all lead (Fig. 7b). We are very encouraged by these preliminary results for the multimodel ensemble forecast assessment. The prediction scores suggest the NMME will help improve the CPC realtime monthly-to-seasonal prediction, and will provide improved climate forecast to decision makers and downscaling and other user communities.

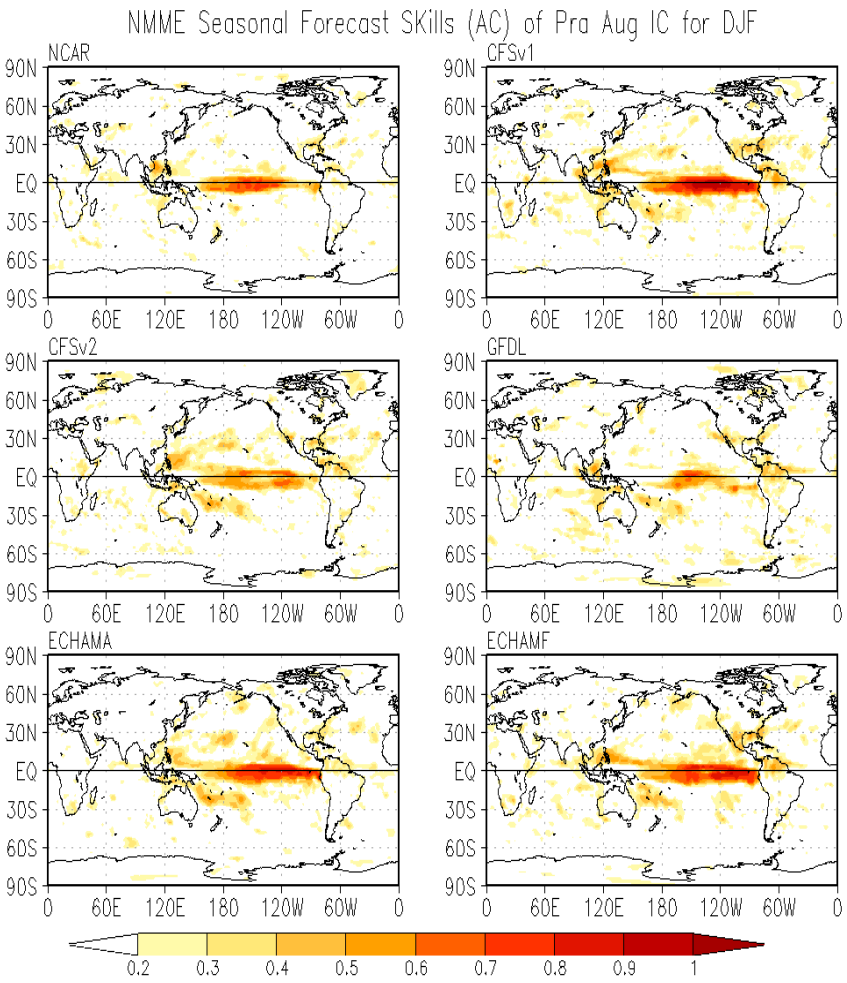


Fig. 4a Maps of anomaly correlation coefficient (ACC) of precipitation for individual models for prediction DJF with August ICs.

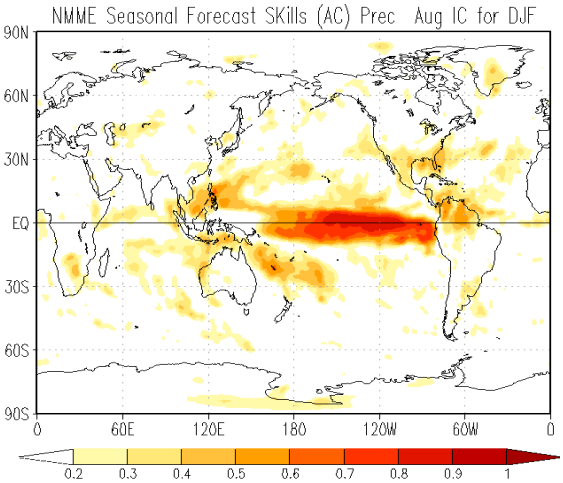


Fig. 4b Maps of anomaly correlation coefficient (ACC) of precipitation for multimodel ensemble for prediction DJF with August ICs.

4. Summary and discussion

Based on existing state-of-the-art US climate prediction models from these institutes (NCEP-CFSv1, NCEP-CFSv2, GFDL-CM2.2, NCAR/U.Miami/COLA-CCSM3, NASA-GEOS5 and IRI-ECHAM4), the Climate Test Bed has launched a phase-1 of National Multi-Model Ensemble project in February, 2011. The monthly-to-seasonal multimodel prediction system is under development and

experimental realtime forecasts have been made on the 8th of each month routinely since August, 2011 to adhere to the schedule of the operational monthly and seasonal forecast in NCEP Climate Prediction Center. Maps of the NMME experimental prediction can be found on the web site: http://www.cpc.ncep.noaa.gov/products/people/wd51yf/NMME. The graphical forecast guidance includes North America and the global domain of precipitation and temperature at 2m anomalies and SST anomaly. The plots are monthly and seasonal means, with or without skill mask, applied for 7 lead months or 5 lead seasons (3-month averages).

All NMME forecast are bias corrected by using the 29 years of hindcast data for each participating model. The model climatology and skill mask are calculated so as to apply to the realtime forecast in each month. The assessment of prediction scores of the three fields for both individual model and NMME ensemble are also given on the web for information about the confidence and reliability of the prediction. This report only discusses the preliminary evaluation of the NMME prediction system for forecast precipitation, temperature at 2m and SST for 2011/12 winter (DJF) with August initial conditions. More detailed information on forecasts and verifications on other months can be found at the web site mentioned above.

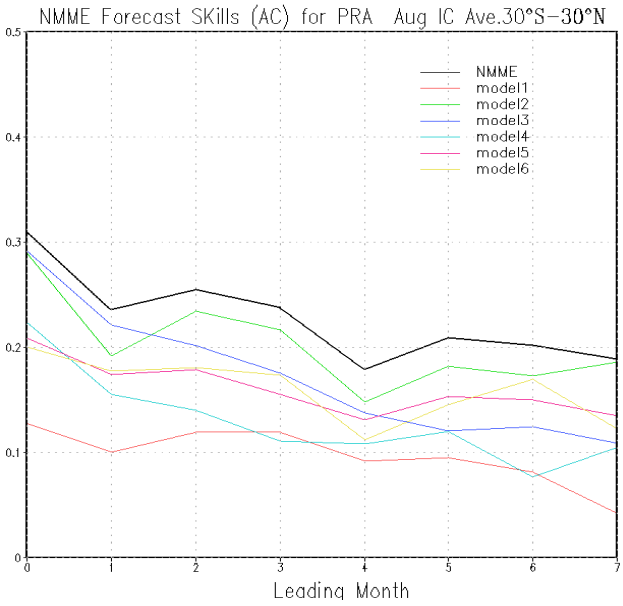


Fig. 5a NMME forecast skills (black line) and individual models (color lines) for precipitation averaged 30S-30N with August ICs.

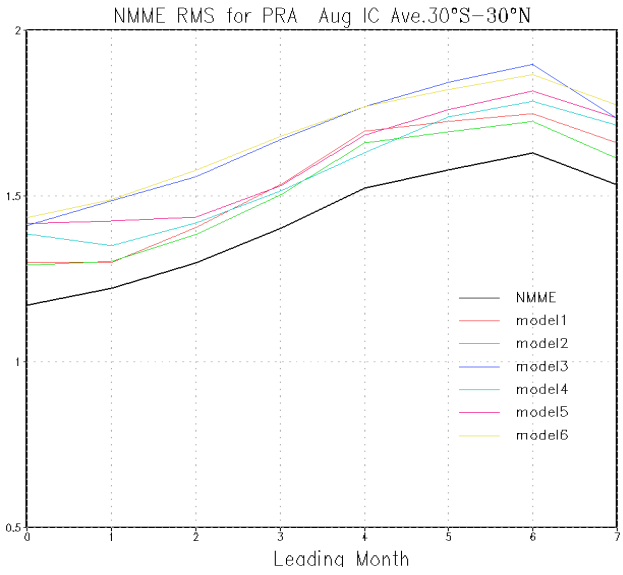


Fig. 5b NMME RMS error (black line) and individual models (color lines) for precipitation averaged 30S-30N with August ICs.

References

Barnston, A. G., M. Glantz, and Y. He, 1999: Predictive skill of statistical and dynamical climate models in SST forecasts during the 1997–98 El Nino and the 1998 La Nina onset. *Bull. Amer. Meteor. Soc.*, **80,** 217–243.

Collins, W. D., and Coauthors, 2006: The Community Climate System Model version 3 (CCSM3). *J. Climate*, **19,** 2122-2143.

DeWitt, D. G., 2005: Retrospective forecasts of interannual sea surface temperature anomalies from 1982 to present using a directly coupled atmosphere-ocean general circulation model. *Mon. Wea. Rev.*, **133**, 2972-2995.

Doblas-Reyes, F. J., R. Hagedorn, T. N. Palmer, 2005: The rationale behind the success of multimodel ensembles in seasonal forecasting - II. Calibration and combination *Tellus A* **57**, 234–252. doi:10.1111/j.1600-0870.2005.00104.x

Hagedorn, R. F. J. Doblas-Reyes, T. N. Palmer, 2005: The rationale behind the success of multimodel ensembles in seasonal forecasting - I. Basic concept *Tellus A*, **57**, 219–233. doi:10.1111/j.1600-0870.2005.00103.x

Goddard, L., and Coauthors, 2001: Current approaches to seasonal-to-interannual climate predictions. *Int. J. Climatol.,* **21,** 1111–1152.

Keenlyside, N.S., M. Latif, M. Botzet, J. Jungclaus and U. Schulzweida, 2005: A coupled method for initializing El Niño Southern Oscillation forecasts using sea surface temperature. Tellus, 57A, 340-356.

Kirtman, B. P., 2003: The COLA anomaly coupled model: Ensemble ENSO prediction. Mon.Wea. Rev., 131, 2324-2341.

Kirtman, B. P., and D. Min, 2009: Multi-model ensemble ENSO prediction with CCSM and CFS. *Mon. Wea. Rev*., DOI: 10.1175/2009MWR2672.1.

Krishnamurti, T.N., and Coauthors, 2000: Multi-Model ensemble forecasts for weather and seasonal climate. *J. Climate*, **13,** 4196-4216.

Landsea, C. W., and J. A. Knaff, 2000: How much skill was there in forecasting the very strong 1997–98 El Nino? *Bull. Amer. Meteor. Soc.*, **81,** 2107–2120.

Lavers, D., L. Luo, and E. F. Wood, 2009: A multiple model assessment of seasonal climate forecast skill for applications. *Geophys. Res. Lett.*, **36**, L23711, doi:10.1029/2009GL041365.

Luo, L., and E. F. Wood, 2006: Assessing the idealized predictability of precipitation and temperature in the NCEP Climate Forecast System. *Geophys. Res. Lett.*, **33**, L04708, doi:10.1029/2005GL025292.

Mitchell, K.E., and Coauthors, 2004: The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res.* **109**(D7). doi:10.1029/2003JD003823.

O'Lenic, Edward A., David A. Unger, Michael S. Halpert, Kenneth S. Pelman, 2008: Developments in Operational

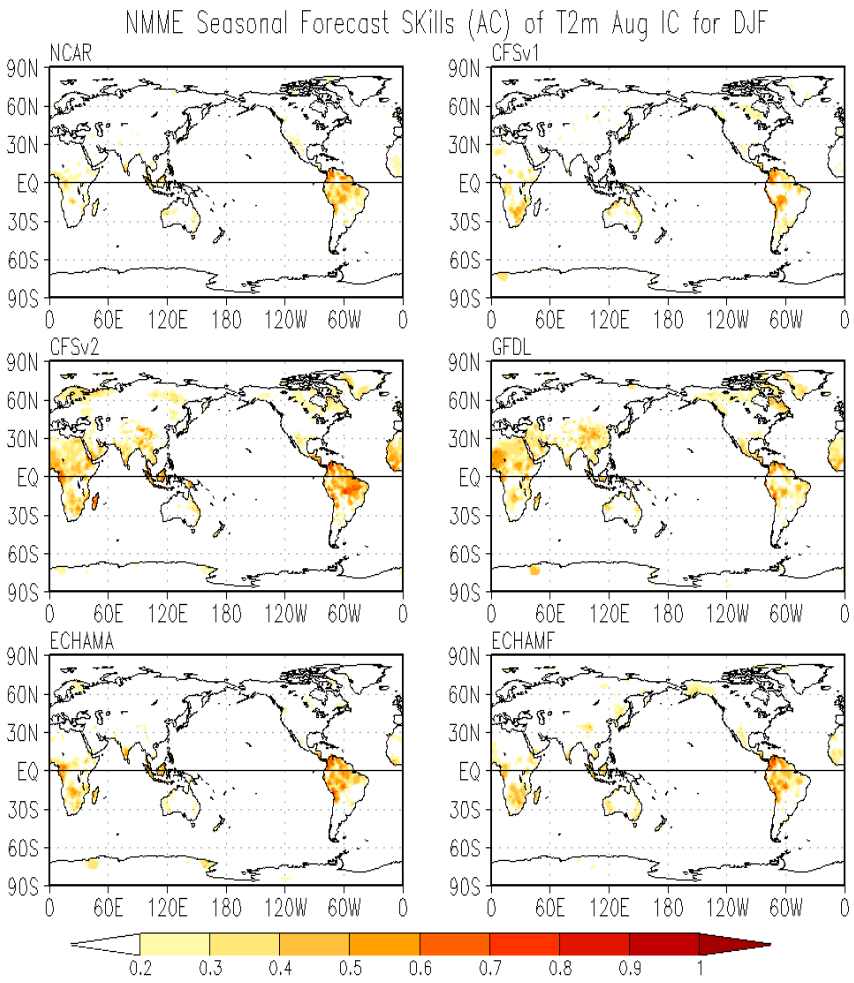


Fig. 6a Maps of anomaly correlation coefficient (ACC) of temperature at 2m for individual models for prediction DJF with August ICs.

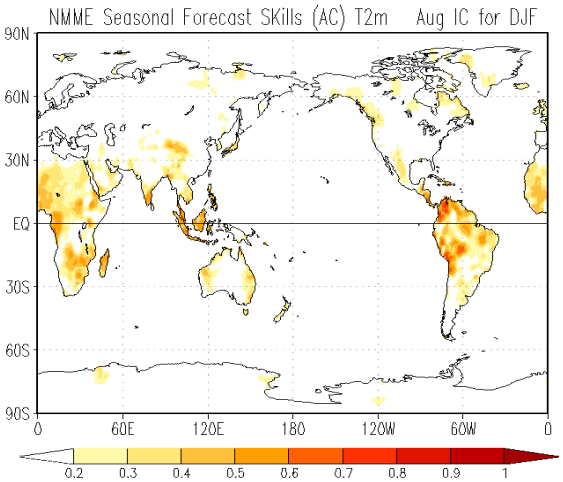


Fig. 6b Maps of anomaly correlation coefficient (ACC) of temperature at 2m for multimodel ensemble for prediction DJF with August ICs.

Long-Range Climate Prediction at CPC. *Wea. Forecasting*, **23**, 496–515. doi: 10.1175/2007WAF2007042.1

Palmer, T.N., and Coauthors, 2004: Development of a European multi-model ensemble system for seasonal-to-interannual prediction (DEMETER), *Bull. Amer. Meteor. Soc*., **85,** 853-872.

Paolino, D.A., and Coauthors, 2011: The Impact of Land Surface and Atmospheric Initialization on Seasonal Forecasts with CCSM. Journal of Climate, in press. doi:10.1175/2011JCLI3934.1

Peng, P., A. Kumar, and H. van den Dool, 2002: Ananalysis of multimodel ensemble prediction for seasonal climate anomalies. J. Geophys. Res., 107, 4710, doi:10.10129/2002JD002712.

Reynolds, R.W., N.A. Rayner, T.M. Smith, D.C. Stokes, and W. Wang, 2002: An improved in situ and satellite SST analysis for climate. J. Climate, 15, 1609-1625.

Saha, S., and Coauthors, 2006: The NCEP Climate Forecast System. *J. Climate*, **19**, 3483–3517. doi: 10.1175/JCLI3812.1

Shukla, J., R. Hagedorn, M. Miller, T. N. Palmer, B. Hoskins, J. Kinter, J. Marotzke, J. Slingo, 2009: Strategies: Revolution in Climate Prediction is Both Necessary and Possible: A Declaration at the World Modelling Summit for Climate Prediction. *Bull. Amer. Meteor. Soc.*, **90**, 175–178. doi: 10.1175/2008BAMS2759.1

Straus, D. M. and J. Shukla, 2002: Does ENSO Force the PNA? *J. Climate*, **15**, 2340–2358.

Weisheimer, A., F. Doblas-Reyes, T. Palmer, A. Alessandri, A. Arribas, M. D´equ´e, N. Keenlyside, M. MacVean, A. Navarra, and P. Rogel, 2009: ENSEMBLES: A new multi-model ensemble for seasonal-to-annual predictions—skill and progress beyond DEMETER in forecasting tropical Pacific SSTs. *Geophys Res Lett*. **36**, L21, 711, DOI 10.1029/2009GL040896.

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.

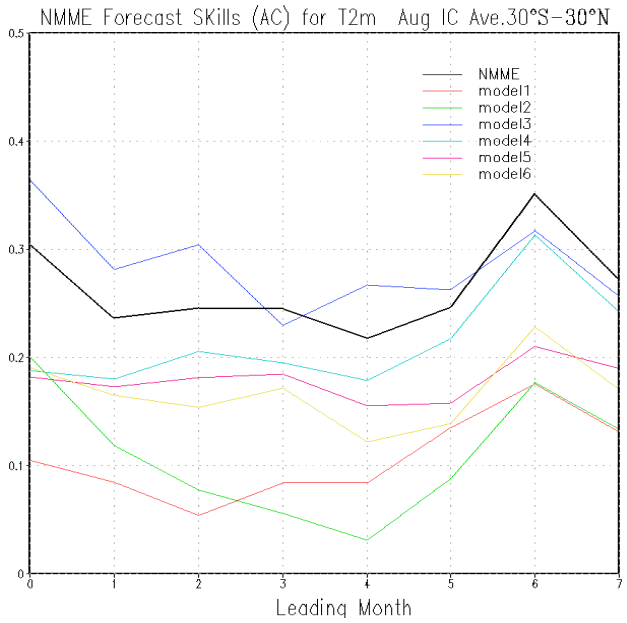


Fig. 7a NMME forecast skills (black line) and individual models (color lines) for temperature at 2m averaged 30S-30N with August ICs.

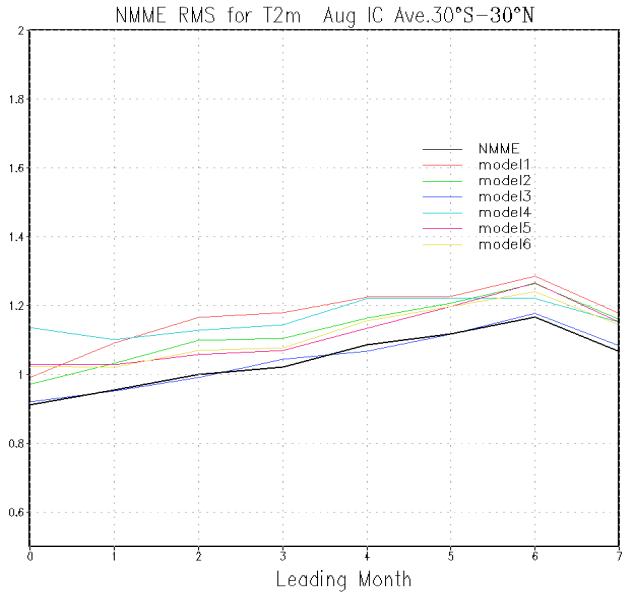


Fig. 7b NMME RMS error (black line) and individual models (color lines) for temperature at 2m averaged 30S-30N with August ICs.

5. CLIMATE VARIABILITY AND ATTRIBUTIONS

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Decadal variations in Eurasian snow: Relation with Circulation and Possible Influence on Spring Rainfall over China

Zhiyan Zuo[1], Renhe Zhang[1], Song Yang[2], Wanqiu Wang[2], Arun Kumar[2] and Yan Xue[2]

[1]*Chinese Academy of Meteorological Sciences, Beijing, China*

[2]*Climate Prediction Center, NOAA/NWS/NCEP, Camp Springs, MD, USA*

1. Introduction

There has been increasing evidence that the atmosphere influences land surface processes, which in turn affect the atmosphere via feedback mechanisms, particularly at decadal time scales. More than a century ago, Blandford (1884) documented an inverse relationship between winter snow over the Himalayas and subsequent all-India monsoon rainfall. There has been increasing evidence that snow may generate anomalous atmospheric forcing via changing the process of energy and water transfer between land surface and the atmosphere. The existence of snow-monsoon relationship was supported by subsequent studies with updated snow data and numerical models. For example, Wu and Kirtman (2007) showed that the enhanced spring snow cover in western Siberia corresponded to above-normal spring rainfall in southern China. Some studies emphasized the complex nature of the relationship between Eurasian snow and broader-scale Asian monsoon which is also strongly influenced by El Niño-Southern Oscillation (ENSO) and the Arctic Oscillation (AO). However, it has been found that the connection between Asian climate and ENSO is weakening in the recent decades. Several studies reported that Eurasian snow is playing a more important role in Asian climate variations than before.

Here we summarize two recent papers by Zuo *et al.* (2011) that examined the decadal variations in Eurasian snow, focusing on the relation with circulation and the influence on spring rainfall over China.

2. Data and Methods

The present analysis is based on monthly observed rainfall data recorded at 595 stations in China over the period 1958–2004 provided by the National Meteorological Information Centre of China. The monthly snow water equivalent (SWE) dataset (1979–2004) was provided by the National Snow and Ice Center (Armstrong and Brodzik 2005). The monthly mean winds and geopotential heights were obtained from the National Center of Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR) reanalysis version 1 outputs for the period 1948–2006. The 0-month lead (LM0) NCEP Climate Forecast System version 2 (CFSv2) hindcasts, the ensemble of 16 members, were used to investigate the predictability of the relationship between Eurasian SWE and rainfall in China (Saha *et al*. 2010).

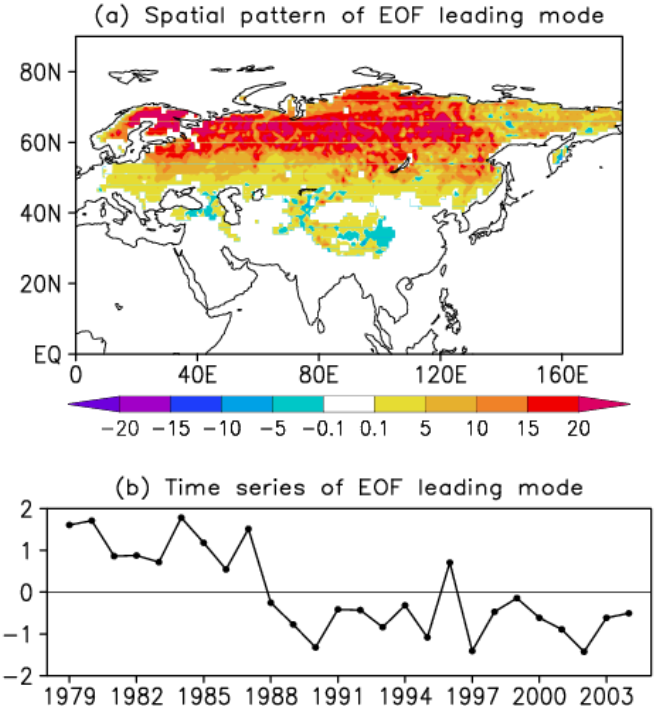


Fig. 1 (a) Spatial pattern of springtime SWE anomalies of the leading mode for the period 1979–2004, and (b) the corresponding time series.

Correspondence to: Zhiyan Zuo, Chinese Academy of Meteorological Sciences, Beijing 100081, China
E-mail: zyzuo@cams.cma.gov.cn.

Empirical orthogonal function (EOF) analysis was used to examine the spatial and temporal characteristics of variability in Eurasian SWE ('Eurasia' defined as the domain from 0°E to 180°E, north of 0°N in the present study). In addition, correlation and composite analyses were performed to investigate the relationship between SWE and spring rainfall in China, assessed using Student's *t*-test.

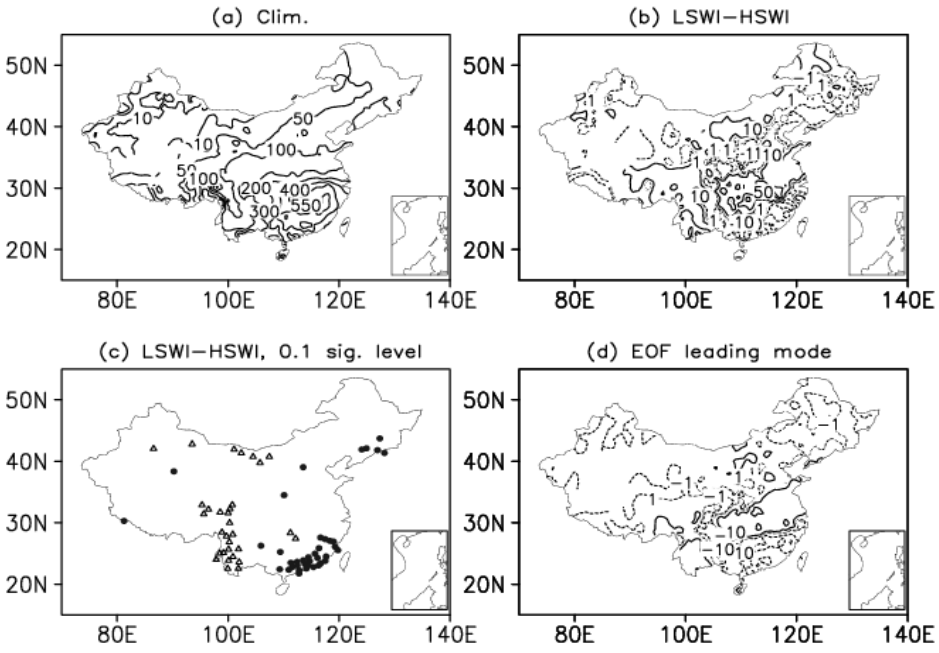


Fig. 2 (a) Climatology of the springtime accumulated precipitation (mm) during 1979–2004, (b) composite difference between the LSWI and HSWI cases, (c) distribution of stations for which the composite difference exceeds the 0.1 level of significance [triangles (dots) indicate significant positive (negative) difference], and (d) the spatial pattern of springtime rainfall anomalies of the leading EOF mode for the period 1979–2004.

3. Results

Figure 1 shows the first EOF mode of Eurasian spring (March-April-May) SWE and the corresponding time series. The variance percentage explained by the mode is 41.7%. Variation in spring snow in the leading EOF mode is homogeneous over large parts of Eurasia, with a maximum value of 20 mm in western and central Siberia, whereas the opposite variation is found over small areas such as the region south of Lake Balkhash and the eastern Tibetan Plateau (Fig. 1(a)). The leading mode of the SWE variability shows an apparent decadal shift in the late 1980s, with negative phases during 1988–2004 (hereafter LSWI) and a mainly positive phase during 1979–1987 (hereafter HSWI), which indicates a decreasing decadal trend in spring Eurasian snow from 1979 to 2004 (Fig. 1(b)).

Figure 2(a) shows the climatology of spring rainfall over China for the period 1979–2004. The rainfall shows a gradual increase from northwestern to southeastern China. Precipitation exceeds 300 mm in southeastern China, with a maximum value of 700 mm. To identify the relationship between rainfall in China and spring snow in Eurasian, the composite difference in spring precipitation was calculated between the LSWI and HSWI cases (Fig. 2(b)). Negative differences are seen across most of northwestern China and eastern China, except the Yangtze River valley and parts of Inner Mongolia. The maximum difference, greater than –100 mm, is seen in southeastern China. Figure 2(d) shows the first EOF mode of the spring rainfall anomaly. The first mode is similar to the composite difference between LSWI and HSWI. Note that the variance percentage explained by the mode is 23.5%, demonstrating that the bulk of rainfall variability is explained by decadal variations in Eurasian spring SWE. Figure 2(c) shows the distribution of stations for which the composite differences exceed the 0.1 level of significance. As expected, the composite differences of spring rainfall are significantly negative in southeastern China and positive in southwestern China. Although the difference is small in northern China, spring precipitation shows a significant decrease in southeastern area in Northeast China and a significant increase in northern Inner Mongolia.

Figure 3(a) shows the composite difference of wave activity flux between LSWI and HSWI. Downward and equatorward wave activity flux anomalies appear at mid- and high-latitudes below 200 hPa. The wave activity flux is a three-dimensional extension of the Eliassen–Palm flux (Edmon *et al.* 1980), which reveals the propagation of wave activity and the effect of waves on the flow. Thus, the reduced Eurasian snow acted

to weaken the upward and poleward wave activity flux in the troposphere over Siberia. The center of upward and poleward zonal-mean wave activity flux was also strongly weakened due to reduced snow in Eurasia.

Figure 3(c) shows the composite differences in SLP between LSWI and HSWI. A positive AO phase was associated with decreasing Eurasian snow. A large area of negative anomalies (maximum values exceeding –3 hPa) emerged over the polar cap, extending from Siberia to northern Canada and covering almost the entire polar cap at the land surface. Positive centers appeared in mid-latitude regions, with maximum values exceeding 2 hPa over the North Pacific. The negative SLP anomalies over the Arctic and positive anomalies over mid-latitude areas provided quantitative evidence that deficient snow over the Eurasian continent corresponded to a positive AO phase. The 500-hPa geopotential height field followed a similar pattern to that of SLP.

With a persistent decrease in Eurasian snow, the positive AO phase is associated with anomalous anti-cyclonic circulation over Siberia, resulting in turn in robust northerly anomalies over all of eastern China, with maxima in Northeast and southeastern China (Fig. 3(b)). Consequently, the warm and moist southerly is weakened, resulting in divergence anomalies over Northeast and southeastern China and reduced regional rainfall. In contrast, the northerly anomaly corresponds to greater amounts of water vapor being blocked in southwestern China, resulting in enhanced rainfall. Moreover, the westerly anomaly due to the anomalous anti-cyclone over Siberian is associated with enhanced water vapor convergence over northwestern China, resulting in a positive precipitation anomaly.

The CFSv2 captures the decreasing trend of spring SWE over Eurasia reasonably well (Figure 4(a)), with maximum negative anomalies of -30kg/m^2 in eastern Europe and Siberia. The reduced precipitation over southern China and surrounding oceans is also captured by the CFSv2 successfully, with maximum anomalies -1.2 mm/day over southeastern China (Figure 4(b)). The characteristics aforementioned indicate that the CFSv2 is skillful for predicting the relationships of the Eurasian SWE with the rainfall in southeastern China.

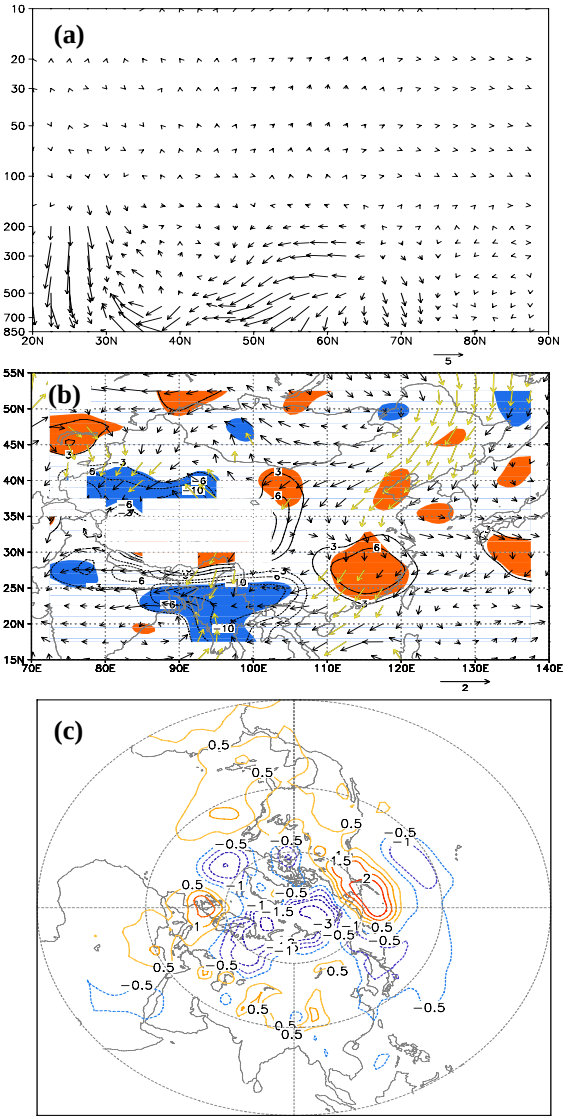


Fig. 3 Composite difference between the LSWI and HSWI cases in terms of meridional and vertical (×10) wave activity flux (m^2 s^{-2}) along 90°E (a), the horizontal wind field (vectors, m s^{-1}) and horizontal moisture flux divergence (contours and shading, g × (s × cm × mb)$^{-1}$ × 10^{-7}) at 850 hPa (b) and SLP (c).

4. Summary and Conclusions

The relationship between decadal variability in spring snow water equivalent (SWE) over Eurasia and spring rainfall over China is investigated using satellite-observed SWE, rainfall observations from 595 stations, and NCEP/NCAR reanalysis data. Decreasing spring SWE in Eurasia corresponded to reduced spring rainfall over southeastern and Northeast China, and more rainfall over southwestern and northwestern China. This relationship was supported by the feedback effect of snow in high-latitude areas to changes in background atmospheric circulation.

A decadal shift in spring Eurasian SWE occurred in the late 1980s, marked by a change from persistent positive phases in 1979–1987 to frequent negative phases. The reduction in Eurasian SWE resulted in reduced upward and poleward wave flux activity, which corresponded to anomalous negative heights/pressures in the Arctic and anomalous positive heights/pressures in mid-latitude regions from the upper-level troposphere to the surface. There was an anomalous anti-cyclonic circulation over Siberia and the western Pacific subtropical high was weakened, accompanied by an anomalous northerly in eastern China and westerly in northwestern China. The anomalous northerly resulted in reduced water vapor convergence in southeastern and Northeast China. Thus, negative rainfall anomalies developed over southeastern and Northeast China, and positive rainfall anomalies appeared over southwestern and northwestern China. Restropective forecasts from CFSv2 successfully simulated the relationship between Eurasian SWE and southeastern China rainfall.

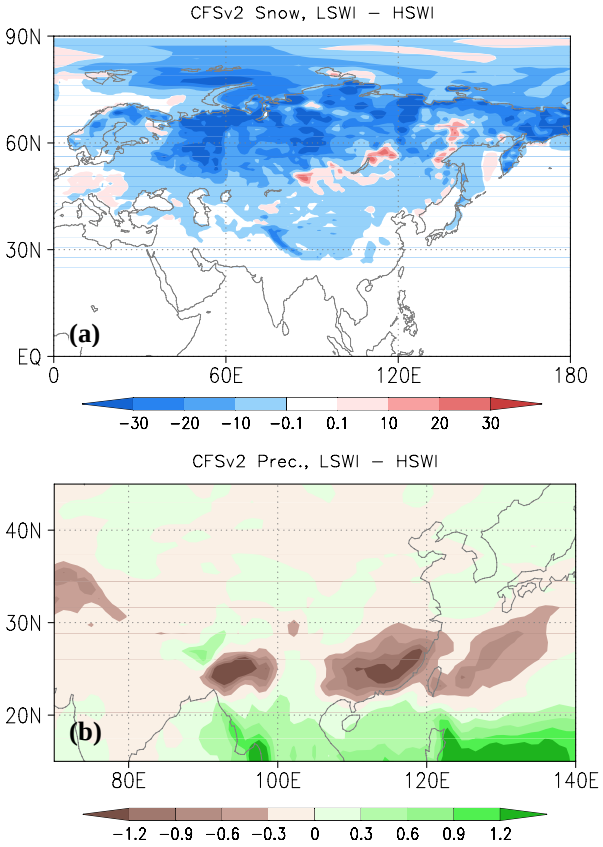


Figure 4 (a) Composite difference between LSWI and HSWI in terms of CFSv2 LM0 spring SWE (a) and precipitation (b).

References

Armstrong, R.L., and M.J. Brodzik, 2005: *Northern Hemisphere EASE-grid weekly snow cover and sea ice extent version 3. National Snow and Ice Data Center* (Boulder, CO, digital media).

Blanford, H.F., 1884: On the connexion of the Himalaya snowfall with dry winds and seasons of drought in India. *Procceedings of the Royal Society of London*, **37**, 3–22.

Edmon HJ, Hoskins BJ, McIntyre MF, 1980: Eliassen-Palm crosssections for the troposphere. *Journal of Atmospheric Sciences.* **37**, 2600–2616.

Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057.

Wu, R., and B.P. Kirtman, 2007: Observed relationship of spring and summer East Asian rainfall with winter and spring Eurasian snow. *J. Climate*, **20**, 1285–1304.

Yang, S., 1996: ENSO–snow–monsoon associations and seasonal–interannual predictions. *Int. J. Climatol.*, **16**, 125–134.

Zuo, Z., R. Zhang, B. Wu, and X. Rong, 2011: Decadal variability in springtime snow over Eurasia: Relation with circulation and possible influence on springtime rainfall over China. *Int.l J. Climatol.*, DOI: 10.1002/joc.2355

Zuo, Z., Y. Song, W. Wang, A. Kumar, Y. Xue, and R. Zhang, 2011: Relationship between anomalies of Eurasian snow and southern China rainfall in winter. *Environ. Res. Lett.*, 6, 6pp, DOI: 10.1088/1748-9326/6/4/045402.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Tropical-Extratropical Teleconnections in Boreal Summer: Observed Interannual Variability

Qinghua Ding[1], Bin Wang[1], John M. Wallace[2] and Grant Branstator[3]

[1]*Department of Meteorology, School of Ocean and Earth Science and Technology, University of Hawaii at Manoa, Honolulu, HI*

[2]*Department of Atmospheric Sciences, University of Washington, Seattle, WA*

[3] *National Center for Atmospheric Research (NCAR), Boulder, CO*

ABSTRACT

Maximum covariance analysis is performed on the fields of boreal summer, tropical rainfall and Northern Hemisphere (NH) 200 hPa height for the 62-year period of record 1948-2009. The leading mode, which appears preferentially in summers preceding the peak phases of the El Niño-Southern Oscillation (ENSO) cycle, involves a circumglobal teleconnection (CGT) pattern in the NH extratropical 200 hPa height field observed in association with Indian monsoon rainfall anomalies. The second mode, which tends to occur in summers following ENSO peak phases, involves a Western Pacific-North America (WPNA) teleconnection pattern in the height field observed in association with western North Pacific summer monsoon rainfall anomalies. The CGT pattern is primarily a zonally oriented wave train along the westerly waveguide, while the WPNA pattern is a wave train emanating from the western Pacific monsoon trough and following a great circle. The CGT is accompanied by a pronounced tropical-extratropical seesaw in the zonally-symmetric geopotential height and temperature fields and the WPNA is observed in association with hemispherically-uniform anomalies. These ENSO-related features modulate surface air temperature in both tropics and extratropics. ENSO also affects the wave structure of the CGT and WPNA indirectly, by modulating the strengths of the Indian and western North Pacific monsoons. Linear barotropic mechanisms, including energy propagation and barotropic instability of the basic state flow, also act to shape and maintain the CGT. The implications of these findings for seasonal prediction of the NH extratropical circulation are discussed.

It is well known that ENSO can affect the NH circulation during the boreal winter. Here we have shown that it can also excite a strong extratropical circulation response in the boreal summer preceding the peak phase of an El Niño or La Niña event and a weaker, but sometimes detectable response in the summer following the event. ENSO affects the extratropical circulation during the boreal summer by relocating the monsoonal heat sources.

Global scale co-variability between NH 200 hPa geopotential height (0-87.5ºN) and tropical rainfall (15ºS-30ºN) has been analyzed by applying Maximum Covariance Analysis (MCA) to a 62-year-long dataset consisting of JJA-means. This analysis yields a comprehensive picture of the global scale tropical-extratropical teleconnections that reveals how tropical heating anomalies force an NH extratropical response on the interannual timescale. Emphasis has been placed on assessing the role of the summer monsoons in mediating the response. In Fig. 1 we present a schematic diagram that highlights complementary roles of ENSO and monsoons in driving tropical-extratropical teleconnections during various phases of the ENSO cycle.

The leading MCA mode M1, which accounts for 47% of the squared covariance between tropical rainfall and NH 200 hPa height links the CGT pattern with precipitation anomalies over Indian monsoon regions. The

---

Correspondence to: Qinghua Ding, Department of Meteorology, School of Ocean and Earth Science and Technology, University of Hawaii at Manoa, Honolulu, HI 96822; E-mail: qinghua@u.washington.edu.

CGT is mainly confined within the wave guide associated with the westerly jet stream. The centers of action of the CGT tend to occur at preferred longitudes, which is reproduced in the left hand panel of Fig. 1. The second MCA mode M2, which explains another 25% of the squared covariance, reflects the relationship between summer rainfall anomalies in the WNPSM and the WPNA teleconnection pattern (right hand panel of Fig. 1).

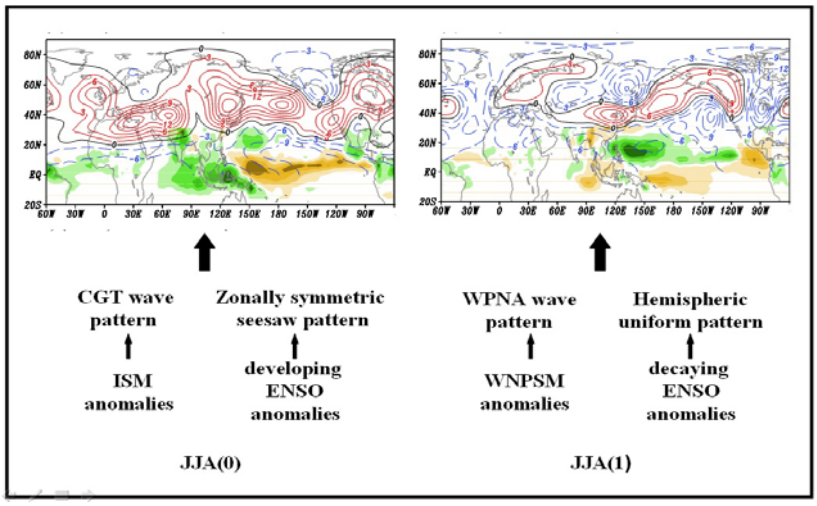


Fig. 1 Schematic showing the role of ENSO in organizing monsoon rainfall and the zonally-symmetric and wave components of the boreal summer (JJA) circulation in Year (0) preceding and Year (1) following the peak phases of the ENSO cycle. See text for further details.

The extratropical circulation patterns in M1 and M2 closely resemble the leading EOFs of the NH upper-tropospheric circulation. The CGT pattern (M1, EOF2) and the WPNA pattern (M2 and EOF1) both exert a substantial impact on surface temperature over tropical and extratropical land areas. The CGT pattern is linked to Indian summer monsoon variability, while the WPNA pattern is more connected with the WNPSM. The CGT explains most of interannual variability along the westerly jet. Thus, the CGT is the primary agent in conveying the influence of tropical thermal forcing to the NH extratropics.

M1 and M2 are associated with developing and decaying phases of extreme in the ENSO cycle, respectively. The strong tropical SST anomalies observed during the summer preceding the peak phases of the ENSO cycle excite a strong extratropical response. In the summers following the peak phase of the ENSO cycle, ENSO-related rainfall anomalies are too weak to produce a robust extratropical response. The asymmetry between the strength of the extratropical patterns in the pre- and post- peak phase summers is largely a reflection of the seasonality of the ENSO cycle itself, and, in particular, the predictability barrier in the boreal spring.

The circulation patterns associated with M1 and M2 both contain zonally symmetric and zonally asymmetric components (Fig. 1). The wave components of the CGT and WPNA patterns are mainly related to the ISM and WNPSM forcing, respectively. The zonally symmetric component exhibits a deep equivalent barotropic structure. M1 is marked by an out of phase relationship between tropospheric temperature anomalies in the tropics and extratropics while the zonally symmetric component of M2 tends to be spatially homogeneous throughout the tropical and extratropical NH. These zonally symmetric features exhibit a distinctive pattern of evolution in association with the ENSO cycle, with the tropical-extratropical seesaw prevailing in summers preceding the peak phase and the hemispheric uniform pattern prevailing in the summers following the peak phase. Tropical and extratropical land surface temperatures vary in a similar manner.

We have shown evidence suggesting that linear, barotropic mechanisms, including energy propagation and barotropic instability of the midlatitude basic flow, shape the structure of the CGT and contribute to its prominence. Such internally generated extratropical variability can exist in the absence of tropical forcing but may also be excited by tropical forcing.

Our results suggest that three different factors are instrumental in producing tropical-extratropical teleconnections during the boreal summer. ENSO forces a zonally symmetric response in the tropics and extratropics and it also modulates the rainfall in the ISM and other regional NH monsoons. The monsoons, in turn, act to excite the wave components of the CGT and WPNA patterns. Internal dynamics of basic state flow contributes to the characteristic structure of the CGT as well as to its maintenance.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Decadal Variation of Rainfall Seasonality in the North American Monsoon Region and Its Potential Causes

Paola A. Arias[1, 2], Kingtse C. Mo[3], and Rong Fu[1]

[1]*Department of Geological Sciences, The University of Texas at Austin*
[2]*Grupo de Ingeniería y Gestión Ambiental (GIGA), Universidad de Antioquia, Colombia*
[3]*Climate Prediction Center, NCEP/NWS/NOAA*

1. Introduction

The North American Monsoon System (NAMS) produces most of the annual rainfall over the southwestern (SW) United States (US) and Mexico (Douglas *et al.* 1993; Stensrud *et al.* 1995). Previous studies have suggested that the interannual variation of the NAMS is mainly controlled by the El Niño-Southern Oscillation (ENSO), whereas its decadal variability can be linked to the Pacific Decadal Oscillation (PDO) (Higgins and Shi 2000; Castro *et al.* 2001, 2007, Grantz *et al.* 2007), the Atlantic Multi-decadal Oscillation (AMO) and the Arctic Oscillation (AO, *e.g.*, Hu and Feng 2008, 2010). These studies have been focused on variability of its onset and rainfall amount, whereas variability of its retreat has received much less attention. This study aims to identify whether seasonality and strength of the NAMS have changed during the period 1948-2009. If so, what causes such a change? A recent study of Li *et al.* (2011) has shown a westward expansion of the North Atlantic Subtropical High (NASH) since late 1970s during the summer season (June-August). This change has increased rainfall variability over southeast US. We will investigate whether such a change could also influence climate variability of the NAMS.

2. Data and Methodology

We used 1-degree gridded daily precipitation over the US and Mexico during 1948-2009 from the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC, Higgins *et al.* 1999). The daily rain rate was averaged over this analysis domain (Fig. 1a) and converted to mean pentad values (*i.e.*, 5-day averages) before obtaining the onset/retreat dates. We also use the National Center for Environmental Prediction (NCEP) reanalysis and extended reconstructed monthly mean sea surface temperature (SST) from the NOAA Climate Diagnostic Center (CDC) (Reynolds 1988) to determine

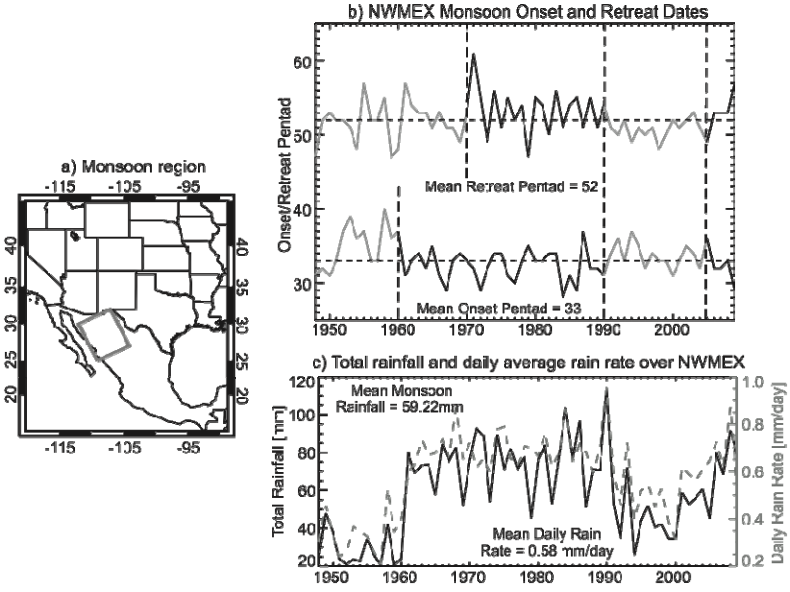


Fig. 1 a) Monsoon region (NWMEX) considered for onset and retreat computations. b) Monsoon onset (lower line) and retreat dates (upper line) over NWMEX during 1948-2009 obtained from the CPC rain rate data. Solid black lines represent periods with early onset and late retreat whereas solid grey lines represent periods with late onset and late retreat. c) Total rainfall (solid black line) and daily average rain rate (dashed grey line) during the entire monsoon season over NWMEX. In b), vertical dashed lines indicate the years when monsoon regime changed. Mean values during 1948-2009 are indicated in b) and c).

Correspondence to: Rong Fu, Department of Geological Sciences, The University of Texas at Austin, Austin, Texas 78712; E-mail: rongfu@jsg.utexas.edu.

atmospheric circulation and SST changes associated with changes of NAMS rainfall and seasonality.

Different definitions for NAMS onset have been used in previous studies (*e.g.*, Higgins *et al.* 1999; Zeng and Lu 2004). The onsets and retreats of the NAMS were identified using the Li and Fu (2004) method to identify tbecause of its objectivity and potential for unifying definitions of the North and South American monsoon onsets and retreats. Because NAMS rainfall is more variable than that of South America, the duration threshold is relaxed from 6 consecutive pentads used in the South American region to 5 consecutive pentads in the North American monsoon region. Timing of the onsets of the NAMS obtained by this definition is generally consistent with those defined by Higgins *et al.* (1997). Early (late) retreat monsoon events are defined as those when the retreat occurred one pentad before (after) the climatological retreat pentad, namely, during the period August 21-September 5 (September 25-October 30), following Gutzler (2004). Weak (strong) monsoons were identified as those when the monsoon total rainfall (*i.e.*, total amount of rainfall during the monsoon season) was 0.5 standard deviation (σ) below (above) its the climatological mean. We also use a sequential t-test analysis of regime shifts (STARS; Rodionov 2004; Rodionov and Overland 2005) to test for changes in monsoon retreat and onset dates, and composites for early and late-retreat events to identify changes of the atmospheric fields. The statistical significance of the composite difference between weak and strong monsoons was tested using a bootstrap test (Efron 1979). We performed 1000 iterations using 95% as the statistical significance threshold and used the bias corrected and accelerated percentile method to estimate the confidence interval. Rotated Empirical Orthogonal Functions (REOFs) was applied to the September global SSTAs during 1948-2009 to identify the leading modes of SSTA, following Schubert *et al.* (2009).

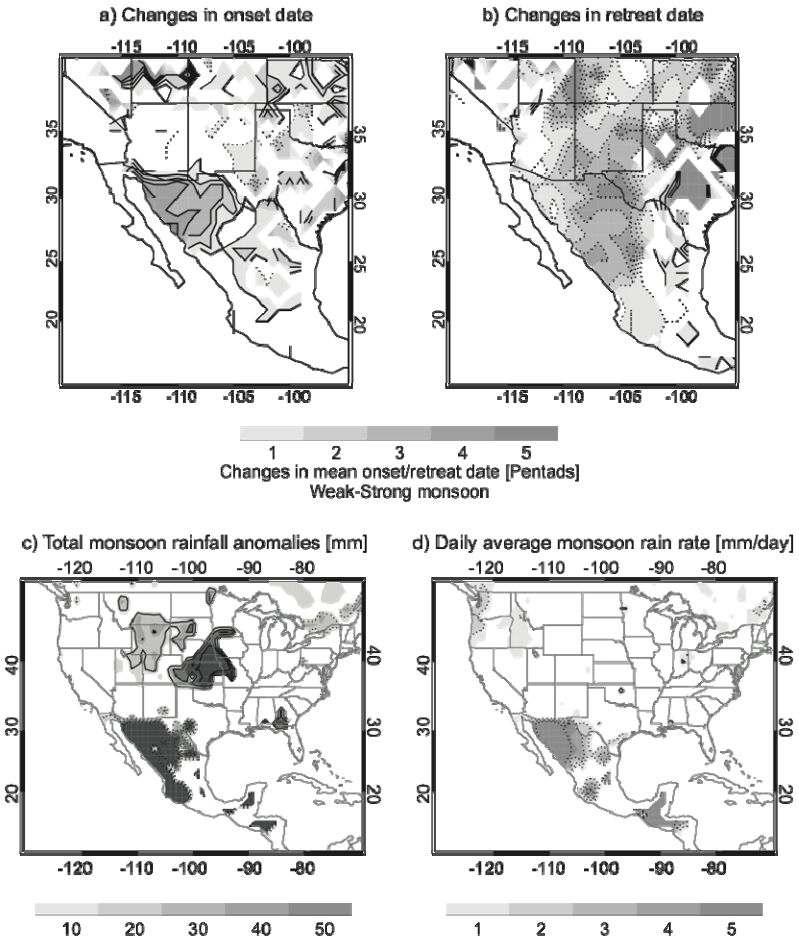


Fig. 2 Changes of a) onset and b) retreat dates over the SW US and Mexico between weak and strong monsoons. Composite difference between weak and strong monsoons for c) total monsoon precipitation anomalies and d) daily average monsoon rain rate. The differences are statistically significant according to a bootstrap test. Dotted (solid) contours indicate negative (positive) changes. The color scale indicates the magnitude of the change.

3. Results

The decadal variations of onset/retreat dates are more clear over northwestern Mexico (NWMEX, Fig. 1a) than over southwestern Mexico. The sequential analysis STARS identified shifts in monsoon retreat during 1971, 1991, and 2006 whereas shifts in monsoon onset were identified during 1960, 1991, and 2006. A simple running-mean-based analysis (not shown) supports the shifts identified using STARS. This analysis suggests that both monsoon onset and retreat exhibit multi-decadal variations: more late-onset and early-retreat monsoons occurred during the periods 1948-1970 and 1991-2005 whereas more early-onset and late-retreat monsoons occurred during 1971-1990 and 2006-2009. Early-onset events are associated with late-retreat

events while late-onset events are associated with early-retreat events. Fig. 1c shows periods of lower monsoon rainfall over NWMEX during 1948-1970 and 1991-2005, whereas periods of higher monsoon seasons occur during 1971-1990 and 2006-2009. The periods of lower and higher monsoon rainfall were generally overlap with those of late onset/early retreat and early onset/late retreat, respectively, except for the ending of the first early retreat and lower monsoon rainfall regime. Correlation coefficients between monsoon total rainfall and onset and retreat dates are -0.58 and 0.6, respectively, both significant at 1% level.

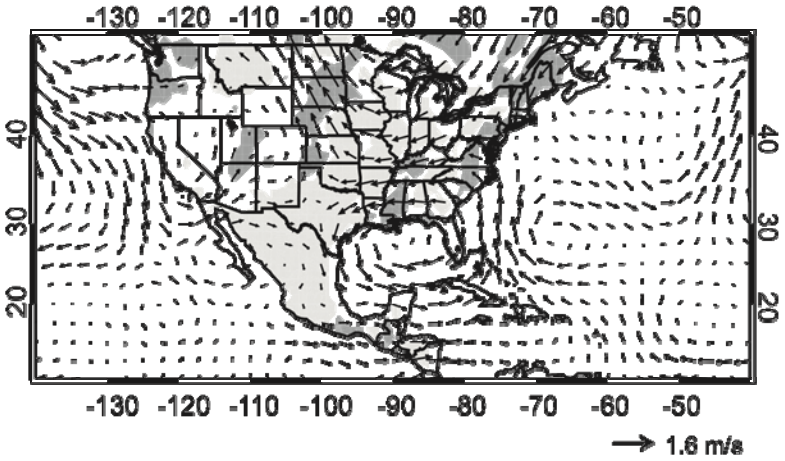


Fig. 3 September mean 850 hPa wind anomalies (arrows) and rain rate anomalies (shades) composites for early NAMS retreat events during 1948-2009. Dark (light) shades represent positive (negative) rain rate anomalies.

Figs. 2a and 2b show the spatial patterns of composite difference in onset/retreat dates between 21 weak and 25 strong monsoon events, passed the bootstrap test (Efron 1979). A weaker summer monsoon over NWMEX is associated with a later monsoon onset and an earlier retreat over the entire core area of the NAMS, namely, the western Mexico region. In addition, changes in the monsoon retreat associated with variations in the monsoon strength occur over the entire NAMS domain whereas changes in the monsoon onset are more confined to NWMEX. Monsoon total precipitation anomalies and daily average monsoon rain rate were also composited for 14 early and 17 NAMS late retreats. Early retreats are associated with reduced precipitation over the monsoon region and increased rainfall over the central United States (US), not only in terms of total monsoon rainfall (Fig. 2c) but also in terms of average daily rain rate (Fig. 2d). This out-of-phase relationship between rainfall over the central US and the monsoon region has been extensively documented (*e.g.* Douglas *et al.* 1993; Douglas and Englehart 1996; Mo *et al.* 1997; Higgins *et al.* 1997; Barlow *et al.* 1998).

To explore the causes of NAMS early retreats, we plot the anomalous circulation patterns at 850 hPa associated with early retreats in Fig. 3, and compare to those associated with positive anomalies of AMO mode, the warming SST mode and northwestern expansion of the NASH in September when NAMS retreats occurred (Figs. 4a-4c). The SST warming mode and AMO mode are represented by the first and third REOF of the global SSTA in September during 1948-2009 (Schubert *et al.* 2009). The composite anomalous circulation patterns for these two SSTA modes were based on the events with their principle component values greater than 0.5. The 2nd REOF mode represents variability of ENSO like SSTA in Pacific. This mode is not chosen because its associated anomalous circulation pattern does not resembles the pattern associate with variation of NAMS retreats.

Fig. 3 shows that the early retreats are characterized by a clearly defined lower tropospheric cyclonic center over the Gulf of Mexico and the SE US and an anticyclonic center over Baja California. These circulation anomalies are associated with the positive rainfall anomalies over the SE US and negative rain rate anomalies over the NAMS region (*i.e.*, most of Mexico and the SW US). The southwesterly lower tropospheric wind anomalies over the NAMS region associated with the positive AMO and SST warming modes (Figs. 4a and 4b) are similar to those associated with the early-retreat events (Fig. 3). However, these two modes cannot adequately explain the anomalous cyclonic circulation over the SE US and the Gulf of Mexico associated with the NAMS early retreats, which appears to be associated with northwestward expansion of the NASH (Fig. 4c). Thus, the composite anomalous lower tropospheric circulation pattern of the combined positive AMO, SST warming modes and northwestward expansion of the NASH appears to contribute to the anomalous circulation pattern associated with NAMS early retreats.

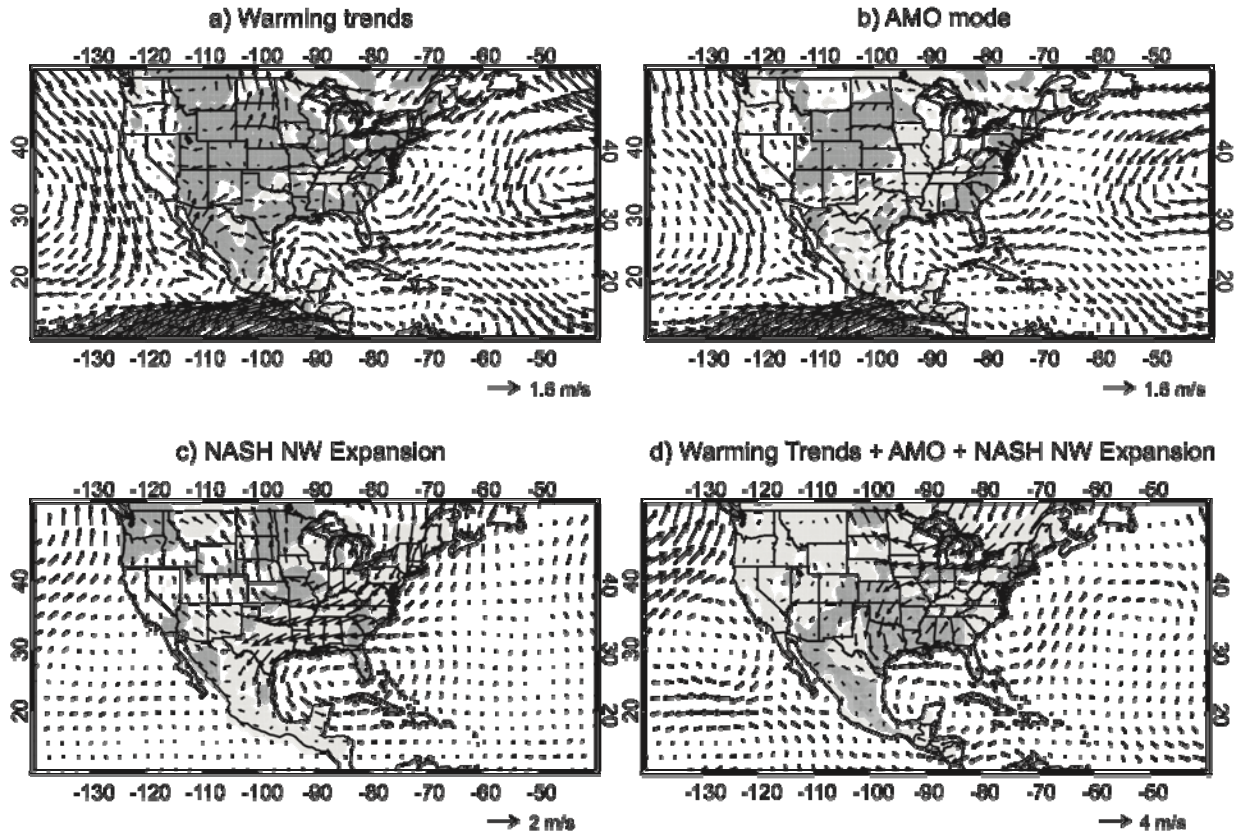


Fig. 4 September mean rain rate anomalies (shades) and 850 wind anomalies (vectors) composites for the events for (a) warming mode, b) AMO mode, c) NASH northwestern expansion, and d) warming + AMO + NASH northwestern expansion occurred during 1948-2009. The vector scale is shown in the right-bottom corner of each panel. Dark (light) shades represent positive (negative) rain rate anomalies.

Main features of the 200 hPa circulation changes associated with early-retreats are an anomalous anticyclonic circulation centered over Texas and northeastern Mexico (to the northwest of the lower tropospheric anomalous cyclonic circulation) and a northwestward shift of the subtropical jets over western US and North Mexico (not shown). These features also most closely resemble those obtained from the combined anomalous circulation patterns associated with positive AMO, SST warming modes and northwestern expansion of NASH.

4. Conclusion

This study shows a decadal scale NAMS regime change occurred during the period of 1948-2009. NAMS was weak due to its late onset, early retreats and weak rainrate during the periods of 1948-1970 and 1991-2005, and was strong due to early onset, late retreat and stronger rainrate during the period of 1971-1990. NAMS appears to have recovered its strength after 2006. This study focuses on causes of the decadal variability of the NAMS retreats because it is not as well understood as those of NAME onset and rainfall. The NAMS early-retreats are associated with an anomalous anticyclonic flow over Baja California and an anomalous cyclonic low-level flow and an increase of rainfall over the SE US and Gulf of Mexico. These changes appear to be more related to an increased atmospheric stability than to a decreased moisture transport to the monsoon region. In the upper troposphere, the early retreats are associated with anomalous divergence, increased (decreased) 200 hPa geopotential height over the SE US (the SW US and northern Mexico), and a northwestward displacement of the subtropical jets over western North America.

The comparison of the anomalous circulation patterns between the variations of NAMS, AMO and SST warming modes suggests that the early NAMS retreats are contributed by positive anomalies of the two SST modes, as represented by the first and third REOFs of SSTA during Septembers. In particular, they contribute

to the westerly anomalous wind and anticyclonic circulation in the lower troposphere over the NAMS region. They also contribute to an increase of geopotential height in the upper troposphere over the SE and the south central US, and in part to an anomalous cyclonic circulation in the lower troposphere over the SE US and the Gulf of Mexico associated with early NAMS retreats. However, these two SST modes cannot explain the anomalous cyclonic circulation in the lower troposphere over SE US and the Gulf of Mexico and the northwestward shift of the subtropical jets in the upper troposphere associated with the early NAMS retreats. The northwestward shift of the NASH appears to be primarily responsible for these two changes. The former probably drives the upper tropospheric divergence over the SE US and compensational upper tropospheric convergence and subsidence over the NAMS region, whereas the latter has been found to cause weaker NAMS by previous studies (*e.g.*, Hu and Feng, 2010). Thus, decadal variation of the NAMS retreats, especially over the NWMEX, appear to be contributed by expansion/contraction of the NASH western edge, AMO and SST warming modes.

Acknowledgments. This work was supported by the NOAA Climate Program Office Climate Prediction Program for the Americas (CPPA) grant (NA10OAR4310157). We sincerely appreciate the insightful suggestions from Dave Gochis and Renguang Wu and Robert Dickinson.

References

Barlow, M., S. Nigam, and E.H. Berbery, 1998: Evolution of the North American Monsoon System. *J. Climate*, **11**, 2238-2257.

Castro, C.L., T.B. McKee, and R.A. Pielke, 2001: The relationship of the North American monsoon to tropical and North Pacific Sea surface temperatures as revealed by observational analyses. *J. Climate*, **14**, 4449–4473.

—, R.A. Pielke, J.O. Adegoke, S.D. Schubert, and P.J. Pegion, 2007: Investigation of the summer climate of the contiguous United States and Mexico using the Regional Atmospheric Modeling System (RAMS). Part II: Model climate variability. *J. Climate*, **20**, 3866–3887.

Douglas, A.V., and P.J. Englehart, 1996: An analysis of the starting date for the summer monsoon in Western Mexico and Southeast Arizona. *Proc. Twentieth Ann. Climate Diagnostics Workshop*, U.S. Department of Commerce, NOAA 207-211.

Douglas, M.W., R.A. Maddox, K. Howard, and S. Reyes, 1993: The Mexican monsoon. *J. Climate*, **6**, 1665-1677.

Efron B., 1979: Bootstrap methods: Another look at the jackknife. *Ann. Stat.*, **7**, 1-26.

Grantz, K, B. Rajagopalan, M. Clark, and E. Zagona, 2007: Seasonal Shifts in the North American Monsoon. *J. Climate*, **20**, 1923–1935.

Gutzler, D.S., 2004: An index of interannual precipitation variability in the core of the North American monsoon region. *J. Climate*, **17**, 4473–4480.

Higgins, R.W., J.E. Janowiak, and X. Wang, 1997: Influence of the North American Monsoon System on the United States summer precipitation regime. *J. Climate*, **10**, 2600–2622.

—, Y. Chen, and A.V. Douglas, 1999: Interannual variability of the North American warm season precipitation regime. *J. Climate*, **12**, 653–680.

—, and W. Shi, 2000: Dominant factors responsible for interannual variability of the summer monsoon in the southwestern United States. *J. Climate*, **13**, 759–776.

Hu, Q., and S. Feng, 2008: Variation of the North American Summer Monsoon Regimes and the Atlantic Multidecadal Oscillation. *J. Climate*, **21**(11), 2371-2383.

—, and S. Feng, 2010: Influence of the Arctic oscillation on central United States summer rainfall. *J. Geophys. Res.*, **115**, D01102, doi:10.1029/2009JD011805.

Li, W., and R. Fu, 2004: Transition of the large-scale atmospheric and land surface conditions from the dry to the wet season over Amazonia as diagnosed by the ECMWF Reanalysis. *J. Climate*, **17**, 2637-2651.

—, L. Li, R. Fu, Y. Deng, H. Wang, 2011: Changes of the North Atlantic Subtropical High and Its Role in the Intensification of Summer Rainfall Variability in the Southeastern United States. *J. Climate*, **24**, 1499–1506.

Mo, K.C., J.N. Paegle, and R.W. Higgins, 1997: Atmospheric processes associated with summer floods and droughts in the central United States. *J. Climate*, **10**, 3028–3046.

Reynolds, R.W., 1988: A real-time global sea surface temperature analysis. *J. Climate*, **1**, 75-86.

Rodionov, S.N., 2004: A sequential algorithm for testing climate regime shifts. *Geophys. Res. Lett*., **31**, L09204.

Rodionov, S.N., and J.E. Overland, 2005: Application of a sequential regime shift detection method to the Bering Sea ecosystem. *ICES J. Mar. Sci*., **62**, 328-332.

Schubert S.D., and coauthors, 2009: A USCLIVAR project to assess and compare the responses of global climate models to drought related SST forcing patterns: Overview and Results. *J. Climate*, **22**, 5251-5272.

Stensrud, D.J., R.L. Gall, S.L. Mullen, and K.W. Howard, 1995: Model climatology of the Mexican monsoon. *J. Climate*, **8**, 1775–1794.

Zeng, X., and E. Lu, 2004: Globally unified monsoon onset and retreat indexes. *J. Climate*, **17**, 2241–2248.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

A Metrics for Boreal Summer Monsoon Intraseasonal Oscillation

June-Yi Lee[1] Bin Wang[1] Matthew Wheeler[2], Xiouhua Fu[1] and Duane Waliser[3]

[1]*IPRC and Department of Meteorology, University of Hawaii, USA*

[2]*Centre for Australia Weather and Climate Research, Bureau of Meteorology, Australia*

[3]*Joint Institute for Regional Earth System Science and Engineering, University of California, USA*

1. Introduction

The tropical intraseasonal oscillation (ISO) is one of the most prominent short-term climate variability in the tropics that has a far reaching influence worldwide (Lau and Waliser 2005). In boreal winter, ISO is characterized by the eastward propagating Julian oscillation (MJO) along the equator (Madden and Julian 1972, 1994), which influences a wide range of weather and climate phenomena and may act an important source of predictability at the subseasonal time scale (Lau and Waliser 2005). In boreal summer, northward propagating monsoon ISO (MISO) is prominent at both 10-20 day and 30-60 day periods in the Asian summer monsoon region (Yasunari 1979, 1980; Webster and Hoyos 2004; Kajikawa and Yasunari 2005 and many others). The MISO is more complex in nature than the MJO due to intrinsic monsoon variability as well as the interaction between the basic monsoon circulation and Madden-Julian Oscillation MJO (Webster *et al.* 1998; Lau and Waliser 2005; Wang 2006). The MISO is known to affect summer monsoon onsets, the active/break phases and the seasonal means of summer monsoons. The wet and dry spells of the MISO strongly influence the extreme hydro-meteorological events, which composed of about 80% of natural disaster, thus the socio-economic activities in the World's most populous monsoon region.

The Real-time Multivariate MJO (RMM) index (Wheeler and Hendon 2004) is most widely used the first two leading multi-variate EOF modes of the equatorial mean (between 15ºS and 15ºN) OLR, and zonal winds at 850 and 200 hPa. This index captures equatorial eastward propagating mode, the MJO, very well and has been applied all year around to depict MJO activity. It has been well recognized that the tropical intraseasonal variations exhibits prominent seasonal variation (Madden 1986, Wang and Rui 1990). During boreal summer, the variability centers of OLR are shifted from equatorial zone during boreal winter to off-equatorial monsoon troughs and the

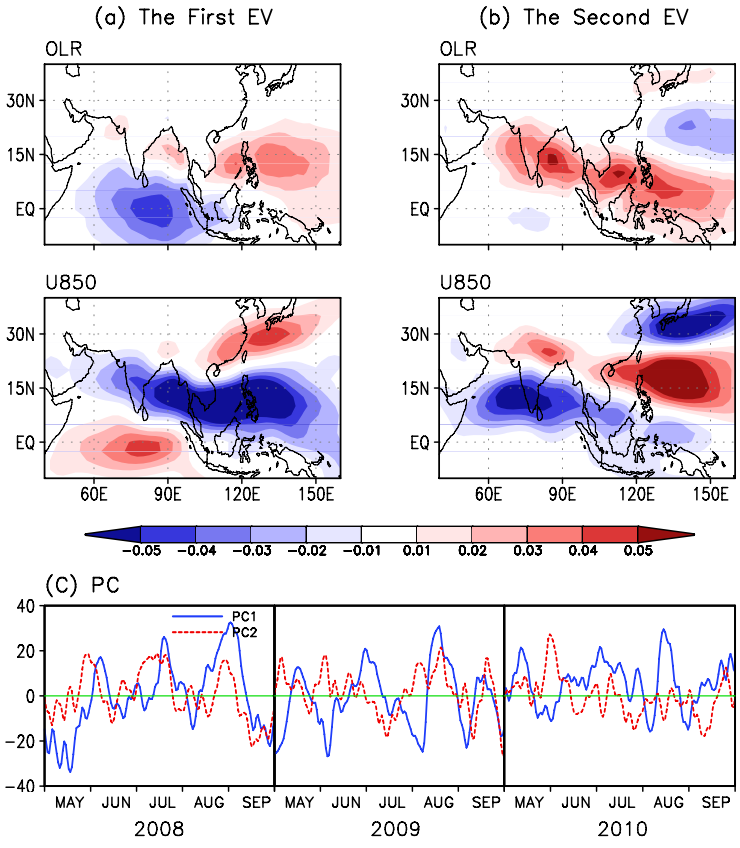


Fig. 1 Spatial pattern (a, b) and PC time series (c) of the first two leading MV-EOF modes of pentad OLR and zonal wind at 850 hPa. In (c), the blue solid line indicates the first and the red dashed line the second principal component (PC). The MV-EOF modes were obtained during MJJAS for the 33 years of 1979-2010.

Correspondence to: June-Yi Lee, Department of Meteorology and International Pacific Research Center, University of Hawaii at Manoa, Honolulu HI 96822; E-mail: juneyi@hawaii.edu.

propagation patterns changed dramatically. It is hence not clear whether RMM remains a best measure of the boreal summer monsoon intraseasonal oscillation (MISO). This study has made an effort on design a better index to describe boreal summer MISO.

2. Definition of MISO index

The data used include interpolated daily outgoing longwave radiation (OLR) with 2.5° horizontal resolution from NOAA (Liebmann and Smith 1996) and daily horizontal wind at 850 and 200 hPa from NCEP/department of Energy (DOE) reanalysis II (Kanamitsu *et al.* 2002).

The MISO metrics introduced in this study was designed to represent larger fractional variance and to better capture the observed northward propagating ISO over the Asian summer monsoon (ASM) region than the RMM index after considerable sensitivity tests. Multivariate empirical orthogonal function (MV-EOF) analysis was applied to daily mean normalized OLR and 850-hPa zonal wind (U850) anomalies over the ASM region (10°S-40°N, 40°-160°E) from May 1st to September 30th in the 30 years of 1981-2010. The OLR and U850 anomalies were obtained from removing the first three harmonics in climatological annual cycle and removing the effect of interannual variation through subtracting last 120 day mean. After that, each of two anomaly fields were normalized by area averaged temporal standard deviation over the ASM region. The standard deviation used is 33.34 W m^{-2} for OLR and 4.02 m s^{-1} for U850. We do not apply filtering to define the index for monitoring and forecast purpose except 1-2-1 filtering. After applying the MV-EOF on the normalized OLR and U850 anomalies, we identified the first four modes as important components for representing ISO propagation over the ASM region. The first four PCs were defined as components of the MISO index. Percentage variance of each mode is 8.57, 5.49, 4.43, and 3.66% for the first, second, third, and fourth mode, respectively. Thus, the first four modes can account for 18.4% of total daily variance of the OLR and U850 anomalies over the ASM region.

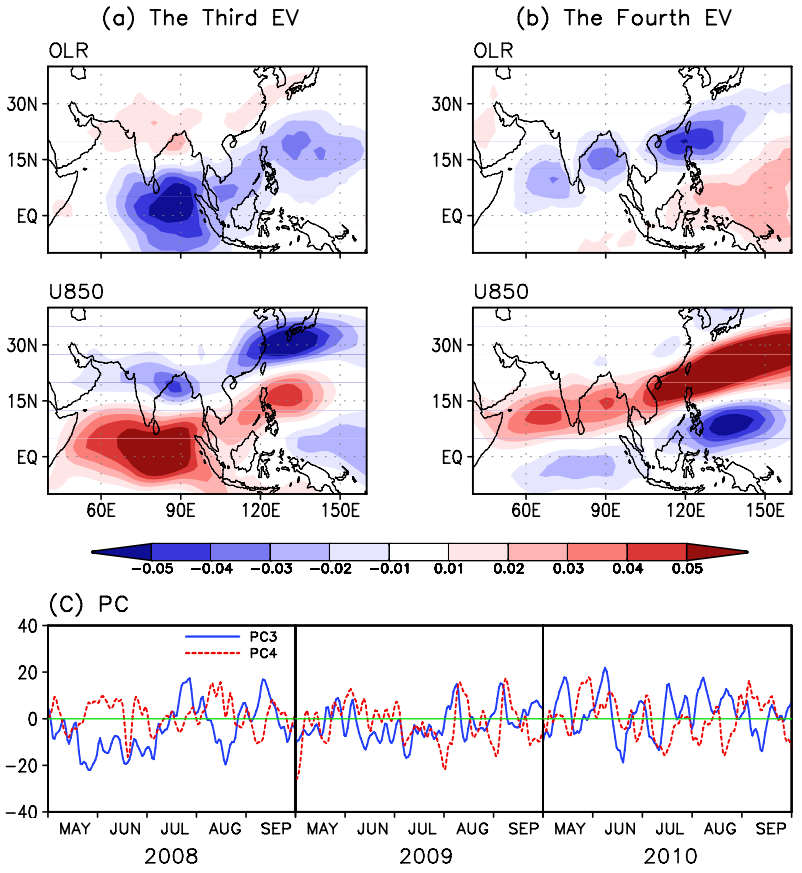


Fig.2 Same as Fig. 1 except for the third and fourth mode.

3. Characteristics of the MISO Components

Figures 1 and 2 show the spatial distribution of eigen vector (EV) and principal component (PC) time series of the first four leading MV-EOF modes of the normalized OLR and

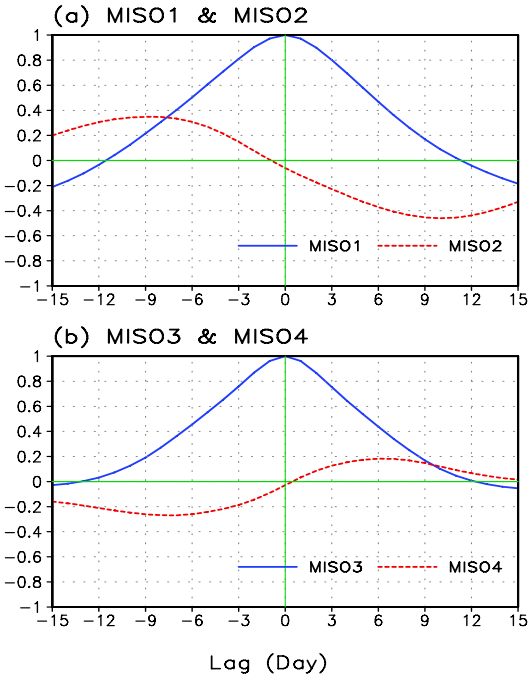


Fig. 3 Lead-lag correlations (a) between the MISO 1 and itself, and with MISO 2 and (b) between the MISO 3 and itself, and with MISO 4 during MJJAS.

U850 anomalies. The first two MV-EOF modes together represent the dominant northward propagating ISO over the ASM region during entire warm season (Fig. 1) addressed in many previous studies (Waliser *et al.* 2003) and the third and fourth modes mainly capture the grand onset mode (LinHo and Wang 2002) of the ASM system (Fig. 2).

The first two EOF modes represent prominent northward propagating ISO over the entire ASM region which is typically oriented in a northwest-southeast direction (Fig. 1a and b). Power spectra of the PCs indicate that the bulk of the variance of the PCs is concentrated at intraseasonal periods (biweekly and 20-60 days). The northward propagating ISO is obvious with PC1 (or MISO1) lagging PC2 (or MISO2) by 7 to 12 days (Fig. 3a). The maximum correlation between MISO1 and MISO2 is -0.4 at a lag of 12 days. The RMM index captures the OLR variability primarily in the equatorial region whereas the MISO 1 and 2 capture large portion of the variability in the off-equatorial region, yielding more realistic variance pattern. Figure 4 shows the life cycle composite of the normalized OLR and 850-hPa wind anomalies using PC1 and PC2 phase space. It is noted that the MISO 1 and 2 describe better ISO variability center and represent better northward as well as eastward propagating pattern in the ASM region than the RMM.

This study demonstrates that the third and fourth MV-EOF modes are also important northward propagating ISO mode, particularly during early summer in association with the grand onset of the ASM system which was described by LinHo and Wang (2002). Differently from the first and second modes, OLR and wind anomalies are in phase over the Indian monsoon and West North Pacific-East Asian monsoon with a southwest-northeast tilt (Fig. 2a and b). Power spectra of the third and fourth PCs indicate that the bulk of their variance is concentrated at intraseasonal periods (biweekly and 20-40 days). The northward propagating ISO is

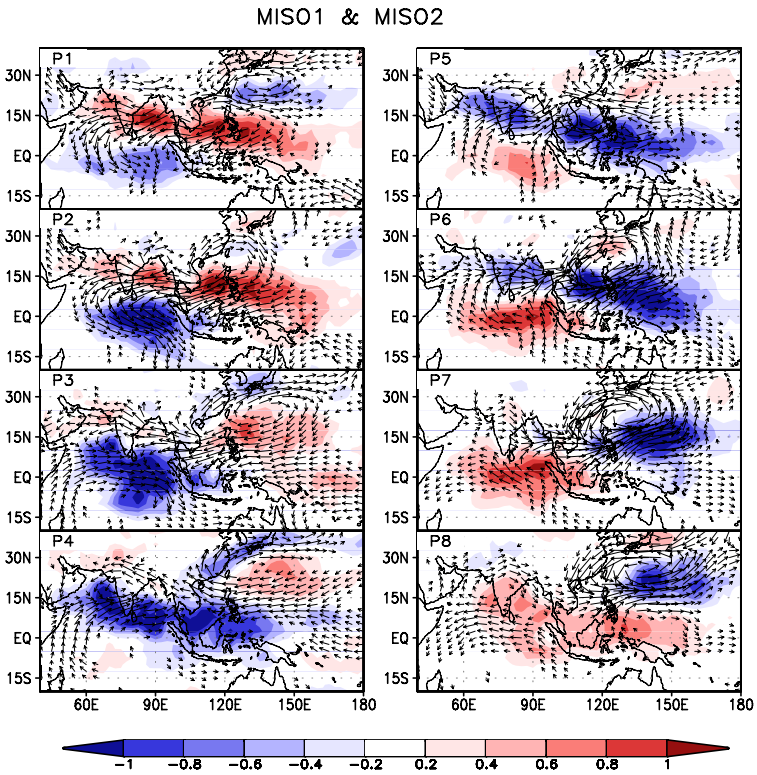


Fig. 4 The life cycle composite of pentad OLR (shading) and 850-hPa wind (vector) anomalies reconstructed based on the MISO1 and MISO2 in 8 phases.

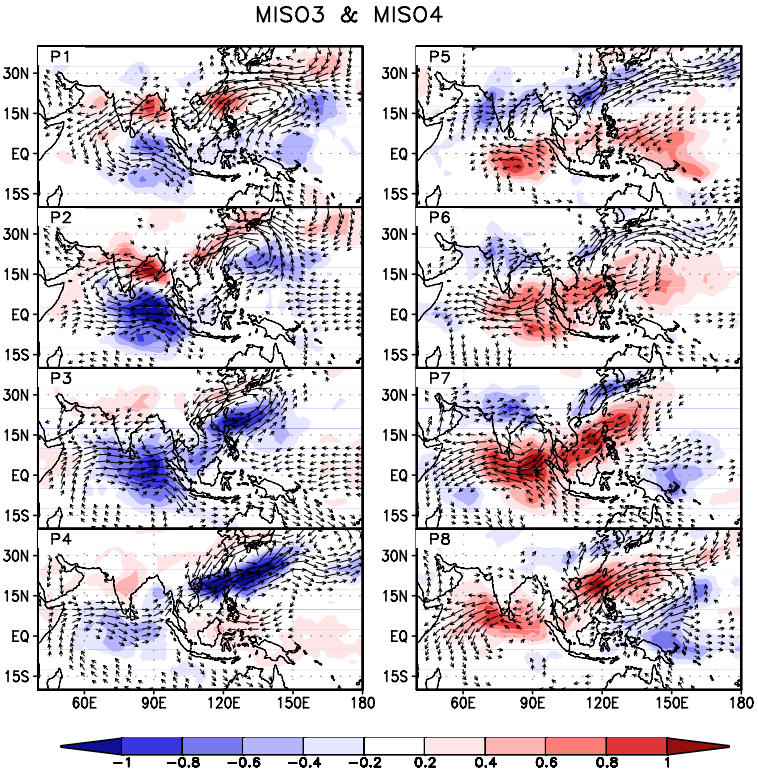


Fig. 5 Same as Fig. 4 except for MISO3 and MISO4.

obvious with PC4 (or MISO 4) lagging PC3 (or MISO 3) by 6 to 8 days (Fig. 3b). The maximum correlation between PC1 and PC2 is -0.28 at a lag of 7 days. The life cycle composite of the MISO 3 an MISO4 represents stepwise onset over the entire ASM region (Fig. 5).

4. Application to real-time monitoring

The northward-propagating MISO component can be monitored using the phase diagram between the MISO 1 and MISO 2 and between the MISO 3 and MISO 4 similar as the eastward-propagating MJO. Figure 6 show example of application to real-time monitoring. During late June to early December, the typical northward propagating ISO was dominant well represented by the points in the two-dimensional phase space defined by the MISO 2 and MISO 1 starting from June 17, 2006. Figure 6b well represents early onset and strong ISO activities over the East Asian monsoon region during early this summer of 2011.

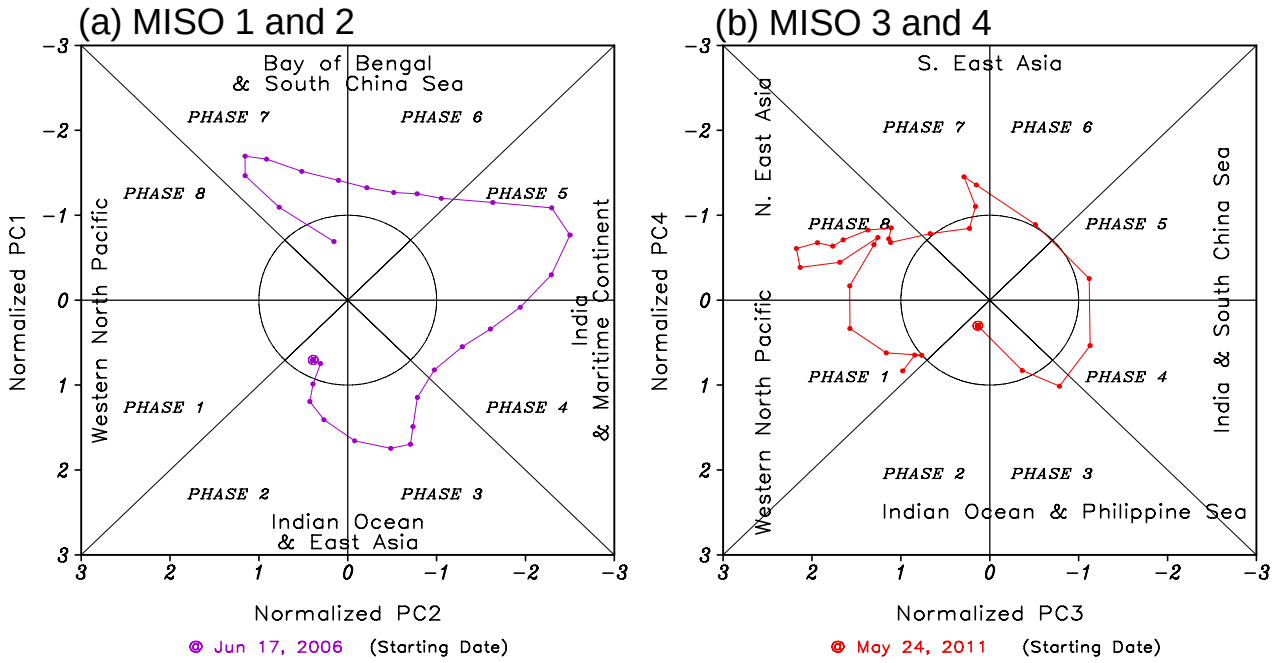


Fig. 6 (a) (MISO2, MISO1) phase space points starting from June 17, 2006. Eight defined regions of the phase space are labeled, as it the region considered to signify weak MISO activity. (b) same as (a) except for (MISO4, MISO3) phase space points starting from May 24, 2011.

5. Summary

Boreal summer Monsoon Intraseasonal Oscillation (MISO) is one of the most prominent short-term climate variability in the global monsoon system and more complex in nature than the Madden-Julian Oscillation (MJO) due to the interaction between the basic monsoon circulation and tropical ISO. To monitor and forecast MISO, we defined real-time multivariate MISO index using daily outgoing longwave radiation (OLR), zonal wind at 850 hPa (U850) over the Asian summer monsoon (ASM) region (10°S-40°N, 40°-150°E), differently from Real-time Multivariate MJO (RMM) index.

The RMM index captures the OLR variability primarily in the equatorial region whereas the new MISO index captures large portion of the variability in the off-equatorial region, yielding more realistic variance pattern. In addition, The MISO index describes ISO variability center better, captures more fractional variance and describes northward as well as eastward propagating pattern better in the ASM domain then RMM index. Sensitivity tests revealed that the MISO index with the first four modes using pentad OLR and U850 is most adequate in terms of fractional variance explained by the reconstructed field and ability to capture the northward propagating MISO.

The northward-propagating MISO component can be monitored using the phase diagram between the first and second PC similar as the eastward-propagating MJO. Taking into account distinct regional characteristics

of MISO with smaller horizontal scale than MJO, the reconstructed field from the first four modes may provide more useful information.

References

Kajikawa, Y. and T. Yasunari T, 2005: Interannual variability of the 10-25- and 30-60-day variation over the South China Sea during boreal summer. *Geophys. Res. Lett.*, **32**, L04710, doi:10.1029/2004GL021836.

Kanamitsu, M., and Coauthors, 2002: NCEP dynamical seasonal forecast system 2000. *Bull. Am. Meteor. Soc.*, **83**, 1019-1037.

Lau, W. K. M., and D. E. Waliser (Eds.), 2005: *Intraseasonal Variability of the Atmosphere-Ocean Climate System*, Springer-Verlag, 474 pp.

Liebmann, B., and C. A. Smith, 1996: Description of a complete (interpolated) outgoing longwave radiation dataset. Bull Am Meteor Soc 77:1275-1277.

LinHo, and B. Wang, 2002: The time-space structure of the Asian-Pacific summer monsoon: a fast annual cycle view. *J. Climate*, **15**, 2001-2019.

Madden, R. A., 1986: Seasonal-variations of the 40–50 day oscillation in the tropics. *J Atmos Sci.*, **43**, 3138–3158.

Madden, R. A., and P. R. Julian, 1972: Description of global-scale circulation cells in tropics with a 40-50day period. *J. Atmos Sci.*, **29**, 1109-1123.

Madden, R. A., and P. R. Julian, 1994: Observations of the 40-50-Day Tropical Oscillation - a Review. *Mon. Wea. Rev.*, **122**, 814-837.

Waliser, D.E., W. Stern, S. Schubert, and K. M. Lau, 2003: Dynamic predictability of intraseasonal variability associated with the Asian summer monsoon. *Quart. J. Roy. Meteor. Soc.*, **129**, 2897-2925.

Wang, B., 2006: *The Asian monsoon*. Springer-Verlag, 787 pp.

Wang, B., and H. Rui, 1990: Synoptic climatology of trans ient tropical intraseasonal convection anomalies: 1975-1985. *Meteor. Atmos. Phys.*, **44**, 43-61.

Webster, P. J., and C. Hoyos, 2004: Prediction of monsoon rainfall and river discharge on 15-30 day time sclaes. *Bull. Amer. Met. Soc*., **85**, 1745-1769.

Webster, P. J., and Coauthors, 1998: Monsoon: Processes, predictability and the prospects for prediction. *J. Geophys. Res.*, **103**, 13341-1451.

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.

Yasunari, T., 1979: Cloudiness fluctuations associated with the northern hemisphere summer monsoon. *J. Meteor. Soc. Japan*, **57**, 227-242.

Yasunari, T., 1980: A quasi-stationary appearance of 30 to 40 day period in the cloudiness fluctuation during the summer monsoon over India. *J. Meteorol. Soc. Japan*, **58**, 225-229.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

On the Connection Between Low-frequency Modulation of Large-scale Weather Regimes and Springtime Extreme Flooding over the Midwest of the United States

Andrew W. Robertson[1], Yochanan Kushnir[2], Upmanu Lall[3] and Jennifer Nakamura[4]

[1]*International Research Institute for Climate and Society (IRI), Columbia University, New York*
[2]*Lamont-Doherty Earth Observatory, Columbia University, New York*
[3]*Department of Earth and Environmental Engineering, Columbia University, New York*
[4]*Lamont-Doherty Earth Observatory, Columbia University, New York*

1. Introduction

The April 2011 flood event in Ohio River Basin and related lower Mississippi River floods was the latest of a set of major such flooding events recorded over the twentieth century (defined in terms of a 10-year return maximum in streamflow). Composite analysis of these events reveals an anomalous northward moisture transport in a "moist conveyor belt" from the Gulf of Mexico and the tropical Atlantic, focused by convergence associated with the "Bermuda High" and the synoptic events impinging on it (Nakamura *et al.* 2011). The questions of whether the recent 2011 event heralds a return of more frequent flooding, and the degree of potential climate predictability of such events both require a better understanding of how the frequency and intensity of the synoptic events responsible for the floods vary on interannual to interdecadal timescales, and are thus potentially influenced by large-scale modes of low frequency climate variability.

2. Data and methodology

We employ an analysis of daily weather regimes over North America [30ºN–50ºN, 105ºW–75ºW] derived from NCEP-NCAR reanalysis 700hPa geopotential height data using a K-means cluster analysis (*e.g.* Robertson and Ghil 1999) for the March–May (MAM) season, 1961–2011 (http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP-NCAR/.CDAS-1/.DAILY/), together with a complementary analysis of daily rainfall gridded gauge-based data over the Ohio River Basin, [88ºW–84ºW, 36–40ºN] for the same period using a hidden Markov model (HMM; *e.g.* Greene *et al.* 2008), based on the CPC Unified Precipitation dataset 1979-2011 (http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP/.CPC/.UNIFIED_PRCP/). Ten-year flood events were estimated from daily

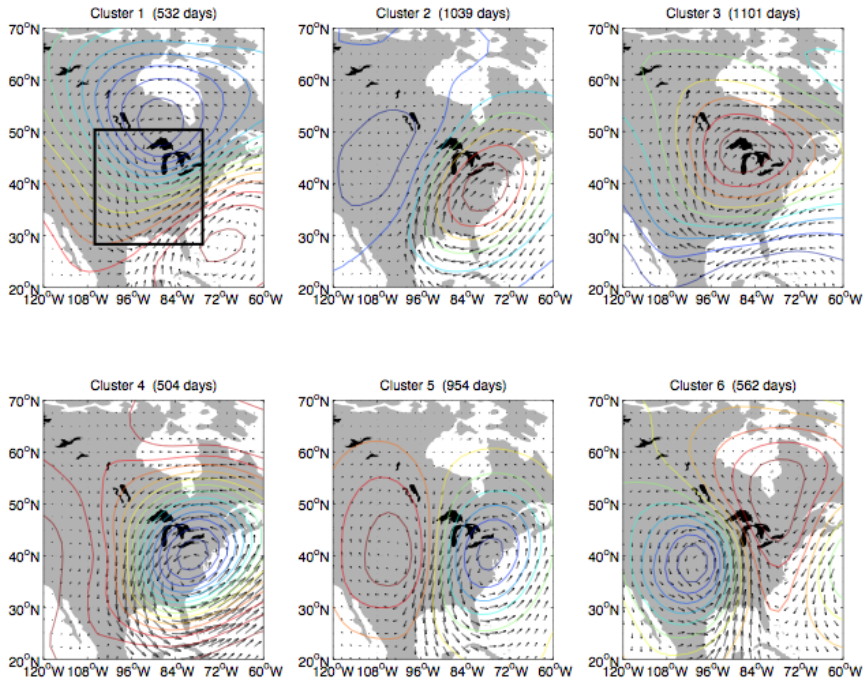


Fig. 1 Six-cluster K-means solution, showing 700hPa geopotential height anomalies (CI 20 gpm), together with anomaly composites of vertically integrated moisture fluxes. Panel titles give the number of MAM days assigned to each cluster. Box in first panel indicates the region used in the K-means analysis.

Correspondence to: Andrew W. Robertson, International Research Institute for Climate and Society (IRI), Columbia University, New York; E-mail: awr@iri.columbia.edu.

river discharge data at 7 gauging stations in sub-basins of the Ohio River based on the Hydro-Climatic Data Network (HCDN) of the U.S. Geological Survey (Nakamura *et al.* 2011).

3. Results

The K-means six-cluster solution was found to yield a near-maximum classifiability index (Michelangeli *et al.* 1995), within the range of K = 5 – 10, and was selected for further analysis; it is depicted in Fig. 1 in terms of geopotential height anomalies from the long-term MAM average, with vertically-integrated moisture flux anomaly composites superimposed. Clusters 1, 2 and 6 are each associated with southerly moisture advection over the eastern U.S.

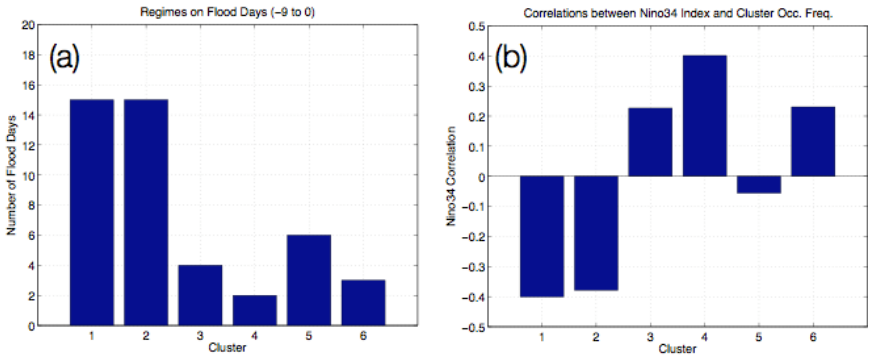


Fig. 2 (a) Frequency of occurrence of each cluster during the 10-day period preceding five 10-year MAM flood events: 13 May 1961, 5 March 1963, 10 March 1964, 25 May 1968, and 4 May 1996. (b) Anomaly correlation between the number of days in each cluster in each MAM season, and the value of the Nino34 index.

To determine whether these flow regimes were active during past extreme flooding events in the Ohio River basin, the frequency of occurrence of each regime (*i.e.* cluster) during the 10-day period preceding five 10-year MAM flood events during the 1961–1996 period (Nakamura *et al.* 2011) is plotted in Fig. 2a; there is a clear preference for clusters 1 and 2 during the lead up to these five events.

The association between cluster frequency and the El Niño-Southern Oscillation is plotted in Fig. 2b, in terms of the anomaly correlation between the Nino34 index, averaged over each MAM season, 1961–2011. There is a tendency for clusters 1 and 2 to be preferentially associated with La Niña events, statistically significant at the 99% confidence level.

Figure 3 shows the daily evolution of rainfall and cluster membership during April 2011, when extreme floods were recorded on the Ohio River, peaking on 27 April. Much of the month was characterized by cluster 2, with clusters 1 and 6 also playing role. All three of these circulation types are characterized by strong northward moisture fluxes from the Gulf of Mexico.

4. Discussion and concluding remarks

The work reported here demonstrates clear associations between synoptic circulation types and historical flood events on the Ohio River. Anomalous southerly fluxes of moisture from the Gulf of Mexico are pronounced in weather types that occurred in connection with these floods. Two of these circulation types are preferentially associated with La Niña, providing one causal mechanism for the recent flooding during April of 2011. Daily rainfall states identified using a rainfall-based Hidden Markov Model indicate a clear eastward propagating synoptic scale wave (not shown). Further work is underway to isolate in more detail the pathways between climate anomalies and extreme flood events through the intermediaries of large-scale atmospheric circulation patterns and synoptic-scale waves.

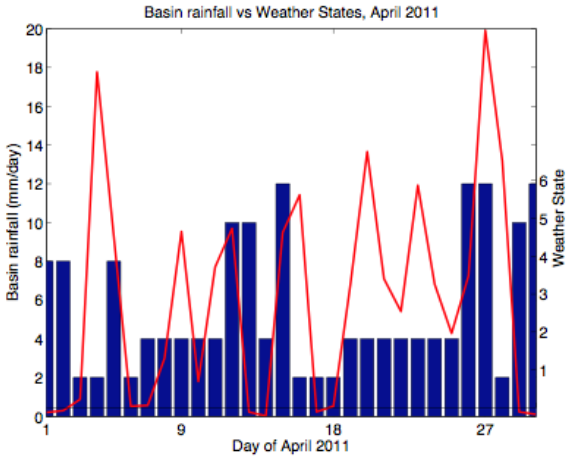


Fig. 3 Daily rainfall averaged over the Ohio Basin (red curve), together with cluster membership (bars), during April 2011.

Acknowledgements. This work was supported by a NOAA Climate Prediction Program for the Americas (CPPA) grant.

5. References

Greene, A. M., A. W. Robertson and S. Kirshner, 2008: Analysis of Indian monsoon daily rainfall on subseasonal to multidecadal time scales using a hidden Markov model. *Quart. J. Royal Meteor. Soc.*, **134**, 875-887.

Michelangeli, P.-A., R. Vautard, and B. Legras, 1995: Weather regimes: Recurrence and quasi stationarity. *J. Atmos. Sci.*, **52**, 1237–1256.

Nakamura, J., U. Lall, Y. Kushnir, A. Robertson, and R. Seager, 2011: An American Floodway: The climatic fingerprint of major regional floods in the Ohio River basin. *Geophys. Res. Letts.*, submitted.

Robertson, A. W., and M. Ghil, 1999: Large-Scale Weather Regimes and Local Climate Over the Western United States. *J. Climate,* **12**, 1796-1813.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

The Limits of Detecting Forced Responses on Seasonal and Continental Scales

Liwei Jia[1] and Timothy DelSole[1,2]

[1]*Center for Ocean-Land-Atmosphere Studies, Calverton, MD*
[2]*George Mason University, Fairfax, VA*

1. Introduction

A wide range of studies have concluded that anthropogenic greenhouse gas increases have very likely caused most of the warming on global and continental scales since the middle of the twentieth century, and that further warming over the next century is expected (Hegerl *et al.* 2007). This conclusion is not necessarily "actionable" by local governments and policy makers without more precise predictions on smaller spatial and temporal scales. The question arises as to what are the shortest space and time scales for which detection, attribution, and prediction are possible. One may also question whether the role of separate forcings, such as the role of greenhouse gases, aerosols, solar variability, can be investigated in specific climate events. Another complication is whether the indices for climate events have been selected specifically for their extreme nature, leading to selection bias. Also, pre-selecting indices (*e.g.*, based on spatial average) may lead us to overlook certain kinds of important events. This study proposes an objective framework for addressing the above questions by identifying components that maximize the signal-to-noise ratio of an externally forced event.

2. Method

Assuming forced climate variability (*i.e.*, response to external natural and anthropogenic forcing, including anthropogenic, volcanic and solar forcing) is an independent and additive perturbation to internal unforced variability, the total variance equals the sum of the variances due to forced and unforced components. Therefore, the variance of forced runs ought to be larger than the variance of unforced runs, since the forced runs contain an "extra" component of variability relative to the unforced runs. Moreover, components whose forced variance differs as much as possible from the unforced variance define components in which the forced response is most easily distinguished from unforced variability. Therefore, we seek the component that maximizes the ratio of forced variance to unforced variance. It can be shown that maximizing the variance ratio leads to the generalized eigenvalue problem (Noble and Daniel 1988; DelSole and Tippett 2009)

$$\hat{\Sigma}_F q = \lambda \hat{\Sigma}_U q, \qquad (1)$$

where $\hat{\Sigma}_F$ and $\hat{\Sigma}_U$ are sample covariance matrices of forced and unforced runs respectively. The eigenvalue λ turns out to be the variance ratio corresponding to eigenvector q. Equation (1) has more than one eigenvalue and eigenvector. Each eigenvector corresponds to a discriminant component. It is convention to order eigenvectors in decreasing order of their eigenvalues, such that the first eigenvector maximizes the variance ratio, the second eigenvector maximizes the variance ratio subject to being uncorrelated with the first, and so on. The time series associated with forced and unforced runs are

$$\mathrm{r}_F = \mathrm{F}q \quad \text{and} \quad \mathrm{r}_U = \mathrm{U}q, \qquad (2)$$

where F and U are matrices containing the (centered) time series for the forced and unforced simulations, respectively. It is shown in Jia and DelSole (2011) that the pattern given by

Correspondence to: Liwei Jia, Center for Ocean-Land-Atmosphere Studies, 4041 Powder Mill Road, Suite 302, Calverton, MD 20705; E-mail: liwei@cola.iges.org.

$$p = \hat{\Sigma}_U q \qquad (3)$$

maximizes the mean statistic used to perform detection analysis in optimal fingerprinting analysis, and therefore maximizes detectability in the models. It follows that if no significant pattern can be found, the role of external forcing cannot be distinguished from internal variability.

3. Models and data

The data set used in this study is from the Coupled Model Intercomparison Project phase 3 (CMIP3) multimodel dataset. The 3-month means of surface air temperature and precipitation from the twentieth century runs (*i.e.*, forced runs) and pre-industrial control runs (*i.e.*, unforced runs) were analyzed. All fields were interpolated to a common 72 x 36 grid.

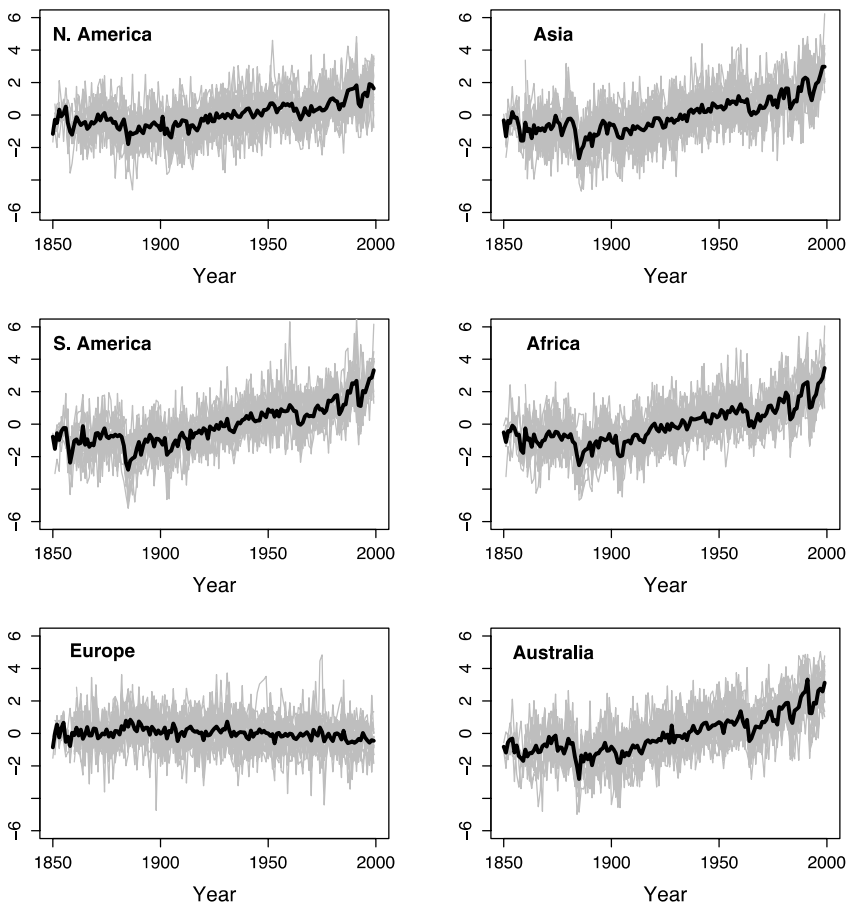


Fig. 1 Time series of the leading component of JFM mean surface air temperature in each forced ensemble member (thin grey curves) over six continents in independent verification data. The thick black curve in each panel shows the multi-model mean time series.

All statistical quantities are estimated in a multi-model sense; more precisely, the covariance matrices from different model simulations are averaged together. Only the last 300 years of unforced runs were used. The first half of the 300-year data was used as training data to maximize the variance ratio, and the second half was reserved for verification. Only one unforced run from each model was used as training. Models with significant trends in unforced runs, and significantly different variances compared to other models, were omitted. This screening procedure leads to a selection of eight models (GFDL-CM2.0, GFDL-CM2.1, IPSL-CM4, MIROC3.2 (medres), ECHO-G, MRI-CGCM2.3.2, CCSM3, UKMO-HadCM3). The selected unforced runs from each model were first centered with respect to each model's mean, and then lined up in temporal dimension to generate multi-model training and verification datasets.

For the forced runs, we used a maximum of five ensemble members in each model, and if the ensemble members are less than five in a model, we used all available members. One member of each model was used as training data to maximize the variance ratio, and the remaining members were used as verification data. Each member was centered with respect to the mean of the run. Similarly, members of eight models were lined up to form multi-model training and verification datasets.

To mitigate overfitting, we reduced the dimension of the data by projecting the data onto the leading 30 principal components. This study shows results only from independent verification data.

4. Results

4.1 Identifying forced response of continental surface air temperature in JFM

We first identify the component that maximizes the ratio of forced-to-unforced variance of JFM mean surface air temperature in training data over six continents. The resulting components were then projected onto the verification data to determine the variance ratios. According to the standard F-test, only the ratio of the leading component is well above the 5% significance level for all six continents except Europe (not shown). Thus, except for Europe, the forced response is distinguishable from unforced variability. The fact that Europe has no significant variance ratio implies that it is impossible to detect a forced response in these models over Europe based on JFM mean surface air temperature. Furthermore, since the response is not detectable in this "perfect model scenario", there is no reason to expect it to be detectable with real observations. Our conclusion pertains to seasonal mean response, whereas most previous studies, which claim that a forced response is detectable over Europe, employ longer-term means (for instance, Fig. 1 of FAQ 9.2 in Hegerl *et al.* (2007) is based on ten-year means).

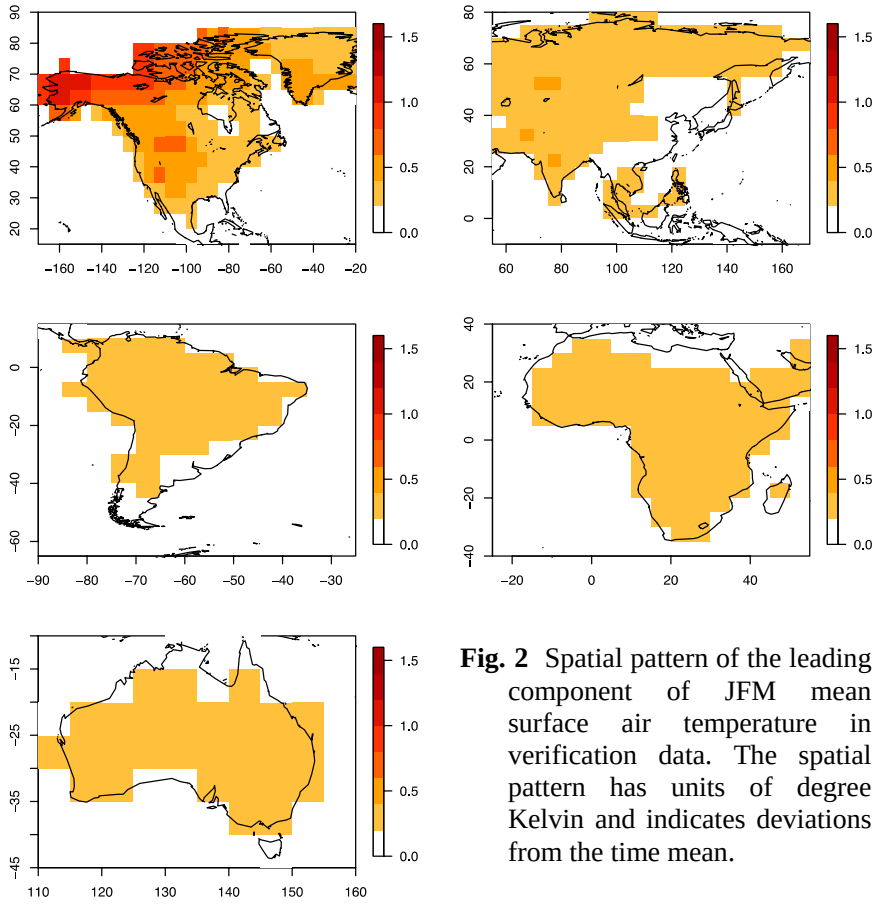


Fig. 2 Spatial pattern of the leading component of JFM mean surface air temperature in verification data. The spatial pattern has units of degree Kelvin and indicates deviations from the time mean.

The fact that only one significant variance ratio can be found in other continents implies that 1) detection of a forced seasonal response pattern in the other continents is possible, and 2) separating the response to different forcings using JFM mean surface air temperature pattern alone will prove difficult, because the similarity of the responses is so great that the responses can be compressed into a single pattern. Therefore, if there are more than one response patterns, they all project on the leading component, and will be collinear and hence difficult to separate.

We emphasize that our analysis is based only on JFM mean spatial structure, *i.e.*, no time lag information is taken into account. This allows us to apply detection and attribution on seasonal scale, but it limits our ability to attribute anomalies to specific forcings. Previous studies that attribute temperature changes to distinct forcings were based on both spatial and temporal information (Stott 2003; Zwiers and Zhang 2003).

The time series of the leading component for forced runs are shown in Fig. 1. The time series of the ensemble mean (thick black curve) shows an increasing trend in each continent except Europe. No significant trend in Europe is consistent with the fact that the corresponding variance ratio is insignificant. The spatial patterns of the leading component (Fig. 2) are of single sign. The positive sign associated with the increasing trend in each continent indicates warming on continental scales. Largest amplitudes are concentrated in high latitudes of North America.

4.2 Identifying forced response of continental surface air temperature in JAS

We repeated the above analysis except this time for JAS mean surface air temperature. Only the ratio of the leading component is well separated from the others and is statistically significant in each continent except Europe (not shown). The time series of the leading component for forced runs (Fig. 3) reveal increasing trends in all continents except Europe. The trends are generally larger in JAS than in JFM. The lack of obvious trend in Europe is consistent with the fact that Europe has no significant variance ratio. However, the forced time series in Europe does show an increasing trend in the last two decades of the twentieth century. It is possible that the trend is real, but that the short time for which the forced response is distinguishable from unforced variability leads to small (and statistically insignificant) variance ratio.

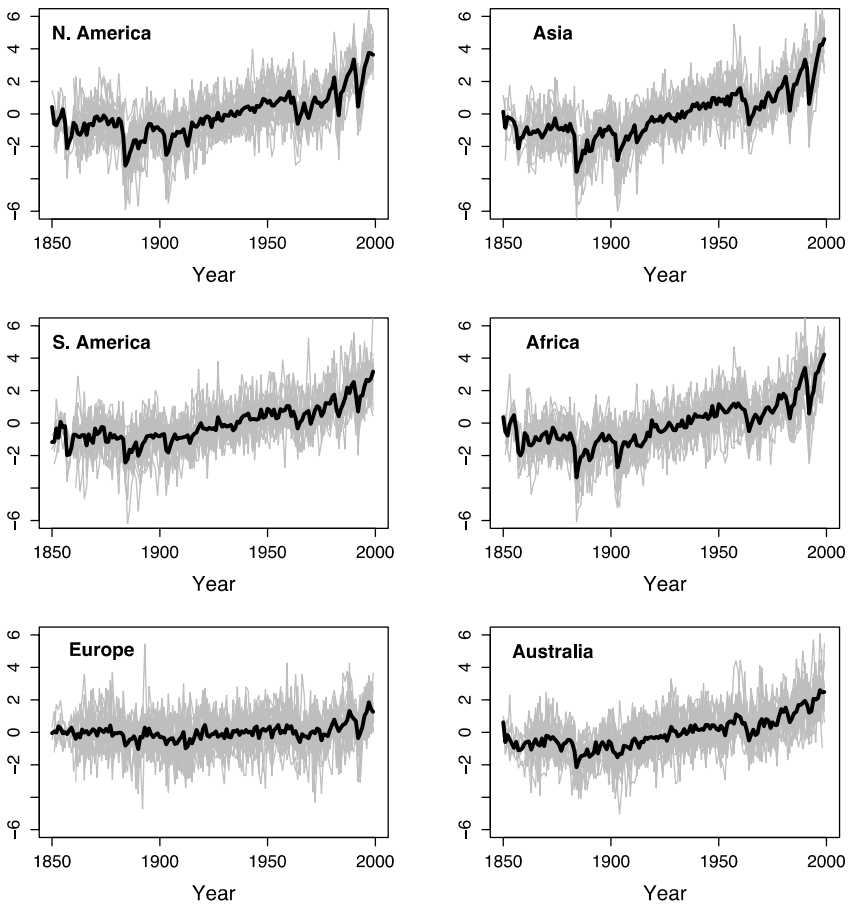


Fig. 3 Time series of the leading component in forced runs over six continents in independent verification data, as in Fig. 1, but for JAS mean surface air temperature.

The spatial patterns of the leading component, shown in Fig. 4, are of positive sign, as those in JFM. The positive sign in spatial pattern associated with the increasing trend in forced time series indicates warming on continental scales. The largest amplitudes are concentrated in high latitudes of North America in JFM, but in the interior of the continent in JAS.

We have tested that the variance ratios determined by projecting vector *q* onto verification data are larger than the ratios of continental averages for all seasons and all continents, except for Europe (not shown).

4.3 Identifying forced response of seasonal precipitation

As for precipitation, none of the variance ratios of JFM and JAS mean precipitation are significant at a 5% level in any continent, implying that the forced response of seasonal mean precipitation is not detectable. This conclusion is somewhat at odds with Zhang *et al.* (2007), who claim that anthropogenic forcing has had a detectable influence on observed changes in precipitation. Although Zhang *et al.* (2007) use land averages in zonal bands while we use continental patterns, we have repeated our analysis for global domains and zonal bands and still find only marginal-to-no significant component. A key difference between the two studies is that Zhang *et al.* (2007) test a trend pattern, which includes decadal scale information of the response, whereas here we test a seasonal mean pattern. Nevertheless, the fact that no significant forced response for precipitation can be found suggests that the precipitation trend must be weak, if it exists at all. This is in fact the case, as Zhang et al. (2007) use scaling factors around 5-10 to match modeled trends with observed trends. It is not surprising that different statistical procedures produce different conclusions for weak signals.

5. Summary

This study addresses the limit to which the response to anthropogenic and natural forcing can be distinguished from unforced variability on seasonal and continental scales. Only one statistically significant forced pattern of seasonal mean surface air temperature can be identified in each season and continent (except Europe, which has no significant forced response), implying that detection of anthropogenic and natural forcing of temperature on seasonal and continental scales is possible. The pattern in each continent is of single sign and consistent with long-term warming, but varies with season. However, the fact that only one significant pattern was obtained implies that different forcings produce similar patterns that may be difficult to separate in an attribution analysis on seasonal and continental scales. No significant forced pattern of seasonal mean precipitation could be identified, implying that detection of anthropogenic and natural forcing of precipitation is not generally possible on seasonal and continental scales. The forced response identified in this study provides the basis for detection and attribution studies on seasonal scales, for instance, in the detection and attribution of observed extreme events.

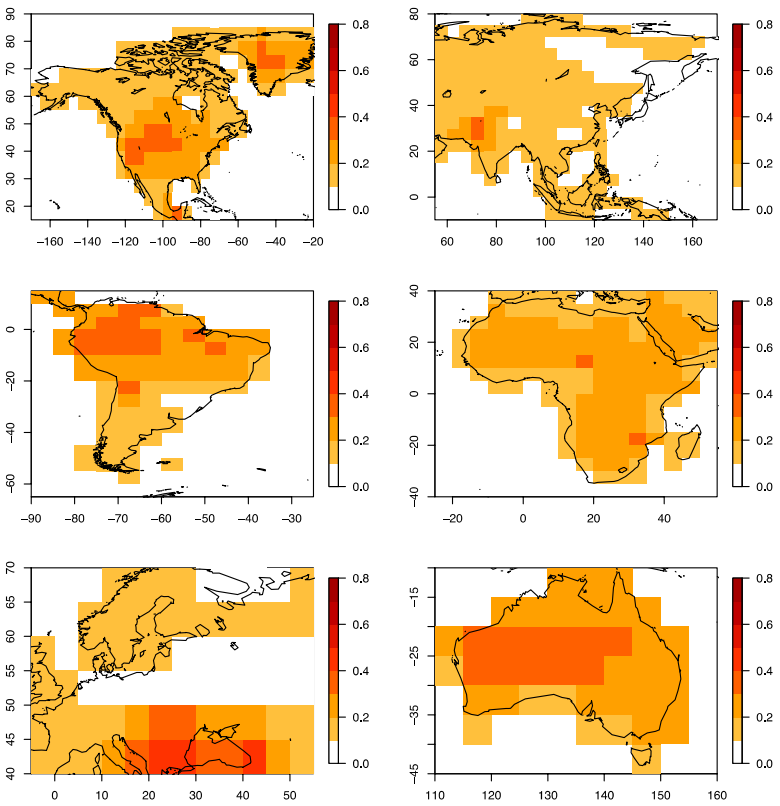


Fig. 4 Spatial pattern of the leading component of JAS mean surface air temperature in verification data. The spatial pattern has units of degree Kelvin and indicates deviations from the time mean.

Reference

DelSole, T., and M. K. Tippett, 2009: Average predictability time. Part II: Seamless diagnoses of predictability on multiple time scales. *J. Atmos. Sci.*, **66**, 1188–1204.

Hegerl, G. C., and Coauthors, 2007: Understanding and attributing climate change. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jia, L. and T. DelSole, 2011: Diagnosis of multiyear predictability on continental scales. *J. Climate*, **24**, 5108-5124.

Noble, B., and J. W. Daniel, 1988: *Applied Linear Algebra*. 3rd ed., Prentice-Hall, 521 pp.

Stott, P. A., 2003: Attribution of regional-scale temperature changes to anthropogenic and natural causes. *Geophys. Res. Lett.*, **30**, 1728, doi:10.1029/2003GL017 324.

Zhang, X., F. W. Zwiers, G. C. Hegerl, F. H. Lambert, N. P. Gillett, S. Solomon, P. A. Stott, and T. Nozawa, 2007: Detection of human influence on twentieth-century precipitation trends. *Nature*, **448**, 461–465.

Zwiers, F. W., and X. Zhang, 2003: Toward regional-scale climate change detection. *J. Climate*, **16**, 793–797.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

The Role of Sub-Seasonal Tropical Convective Variability for the Midlatitude Response to ENSO

Erik Swenson[1] and David Straus[1, 2]

[1]*George Mason University, Fairfax, VA*
[2]*Center for Ocean-Land-Atmosphere Studies, Calverton, MD*

1. Introduction

The boreal wintertime midlatitude response to a significant El Niño Southern Oscillation (ENSO) event has been studied extensively for over 30 years. As well as the forcing itself, the response is primarily understood in terms of an average across the season. However, its development and variability in terms of the sub-seasonally varying forcing is still unclear. For instance, which is more important, short-lived strong convective events or persistent moderate convective events? Can knowledge of the sub-seasonal forcing yield any additional seasonal predictability?

In this study we present model results using large ensembles forced with observed 1982/83 sea surface temperature (SST). The model is modified such that diabatic heating (Q) is prescribed across the tropical Indo-Pacific for which sub-seasonal and ensemble Q variability is neglected. Results indicate a significant impact of low frequency sub-seasonal Q variability for the response in 1982/83.

2. Experimental framework

We use the NCAR Community Atmospheric Model v4.0 (CAM4.0) with a finite volume dynamical core and a resolution of 1.9° x 2.5° x L26. 1982/83 DJFM monthly interpolated SST and sea ice is prescribed for boundary conditions taken from the Hadley Centre observed SST and sea ice dataset. DJFM climatological conditions are generated similarly by forcing the model with climatological SST and sea ice. 50 ensemble members are used for each set of experiments and are generated by slightly perturbing Nov. 1st initial conditions.

Taken from a control set of integrations (CTL), Q is decomposed into the evolving climatological and SST-forced signal (Q_0), the seasonal mean deviation about that signal (Q_1), the 30-120 day low frequency variability (Q_{LOW}), and the < 30 day high frequency variability (Q_{HIGH}). Q_0 is constructed in a way roughly equivalent to a least squares fit to Legendre polynomials (T=8). Q_{LOW} & Q_{HIGH} are computed by partitioning $Q-Q_0-Q_1$ using spherical harmonics. We perform experimental integrations by prescribing varying subsets of Q (Table 1, Fig. 1) at every vertical level across the tropical Indo-Pacific (62.5°E - 87.5°W, 24°S - 15°N) such that the seasonal/ensemble mean is maintained in all experiments. In order to reduce potential boundary issues, Q is relaxed along the boundaries of the domain at a zonal and meridional width of 10^0 and 7°, respectively. Initially, prescription is turned on through gradual relaxation for a period of 10 days. The local effect of Q-circulation decoupling is assessed in separate integrations and confirmed to not have any systematic influence (not shown).

FIX	$Q_0(t)$
EFIX	$Q_0(t) + Q_1(e)$
ESUBFIX	$Q_0(t) + Q_1(e) + Q_{LOW}(t,e)$

Table 1 Experimental integrations and prescribed Q varying in time (t) and ensemble member (e) at every grid point across the tropical Indo-Pacific.

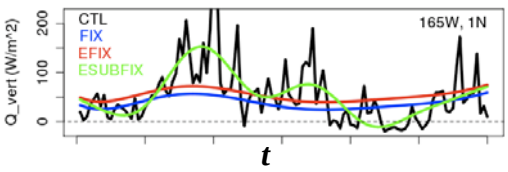


Fig. 1 800-200 hPa vertically integrated Q (Q_{VERT}, W/m^2) for all integrations at an individual grid point in equatorial Pacific (165° W, 1° N).

3. Seasonal/ensemble mean response

CAM4.0 realistically simulates the tropical upper-level divergence/convergence and the midlatitude wave train (Fig. 2) and is quite strong compared to typical model responses. Results

Correspondence to: Erik Swenson, George Mason University, Fairfax, VA; E-mail: swenson@cola.iges.org.

from FIX and EFIX are very similar with the local tropical response is amplified in the equatorial Pacific (maximum div. 9e-6 s^{-1} + 3e-6 s^{-1}). The extratropical response is amplified significantly, particularly with a deeper upper-level low in the North Pacific (minimum height -200 m - 40 m) as well as lower heights over the southern U.S. and Mexico, and over western Asia. For ESUBFIX, only minor differences with CTL occur suggesting that primarily the low frequency Q variability has a tendency to weaken the response for 1982/83.

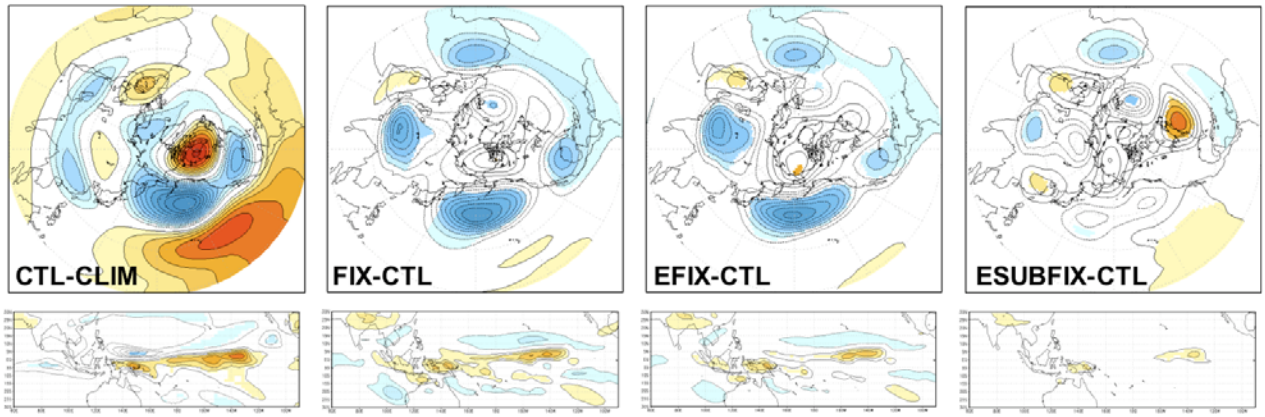


Fig. 2 Seasonal/ensemble mean 200 hPa Z (upper, Z_{200}) and tropical 200-150 hPa divergence (lower, DIV) for CTL-CLIM (left, Z200 in 20 m, DIV in 2e-6 s^{-1}), FIX-CTL, EFIX-CTL, and ESUBFIX-CTL (middle left, middle right, and right, respectively, Z200 in 5 m, DIV in 1e-6 s^{-1}). Significance shaded at 5% level according to Student's t-test.

We find similar differences for the response when examining other ENSO events using different SST (as well as climatological SST), although a similar attribution to low frequency Q variability cannot be made in the extratropics (not shown). This suggests that under other ENSO events, higher frequency variability (< 30 days) may also be playing a role.

Examining the distribution of these seasonal mean differences during the season, we find that tropical upper-level divergence is consistently stronger evenly throughout the season whereas the deepening of the low in the North Pacific has a clear peak in mid-winter (not shown). Consistent with this, additional mid-winter suppression of baroclinic activity occurs along with a peak in eddy feedback onto the mean flow and in the jet extension.

4. Mixing vertical Q profiles

Significantly enhanced upper-level divergence occurs in the equatorial Pacific for FIX and EFIX despite having identical seasonal/ensemble mean forcing. We examine the daily grid point vertical column relationship where the seasonal mean differences occur (120°E - 120°W, 0 - 5°N) and find that the differences are evident for all positive values of Q_{VERT} (Fig. 3, left panel). In terms of the probability density function (PDF) of Q_{VERT} (Fig. 3, middle panel), values corresponding to moderate convection (0 W/m^2 < Q_{VERT} < 70 W/m^2) occur much more often for FIX and EFIX, compensating for the lack of values that correspond to the absence of convection and deep convection (an effect of averaging). Across the same range of values for moderate convection, the corresponding vertical profiles reveal a level of peak Q that is elevated compared to that in CTL. The column redistribution of Q occurs from averaging which consequently absorbs profiles associated with deep convection that have a much higher level of peak heating. This effect is shown for FIX-CTL (Fig. 3, right panel) but also evident in all other prescribed Q experiments. For FIX and EFIX, we estimate that this effect linearly accounts for at least half of the enhanced divergence. For ESUBFIX, the same issues are evident but lessened due to a PDF more similar to that of CTL – a result of including low frequency variability.

5. Conclusion and discussion

We find that for the CAM4.0 simulation of the 1982/83 DJFM El Niño, low frequency (30-120 day) Q variability primarily weakens the response both locally in tropical upper-level divergence, and in the height

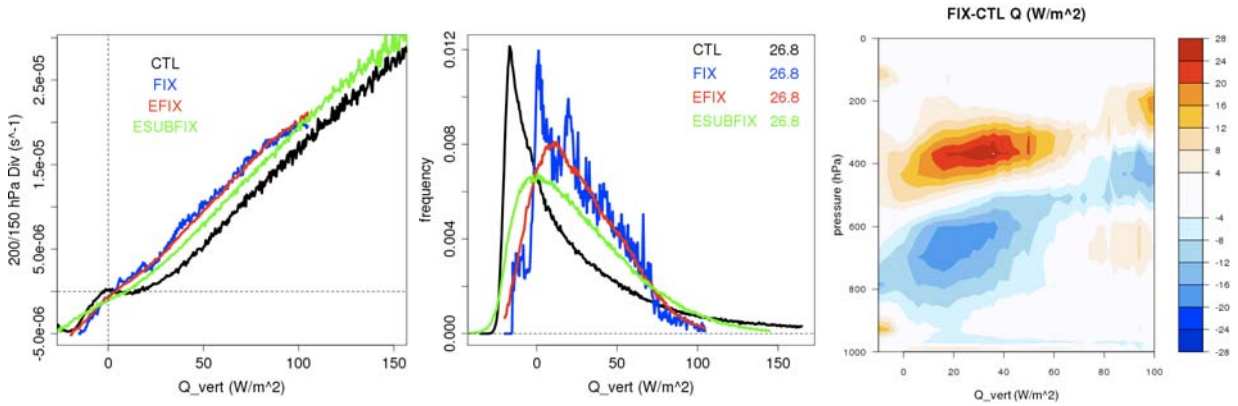


Fig. 3 Average DIV (s^{-1}) as a function of Q_{VERT} (left) and PDF for Q_{VERT} (middle, total integral value in upper right) for CTL, FIX, EFIX and ESUBFIX. Vertical profile of Q for FIX-CTL as a function of Q_{VERT} (right) across the Eq. Pacific (right, for 120°E - 120°W, 0 - 5°N in W/m^2). Binning is used with bins of width 1 W/m^2

response across the Pacific and North America, whereas the seasonal mean ensemble spread of Q interestingly does not seem have much of a role. Results suggest that the local enhancement of tropical divergence may be the primary mechanism forcing the midlatitude differences in the North Pacific, though the potential existence of a dynamical mechanism dependent on the temporal characteristics of the forcing is not ruled out. For instance, Li and Nathan (1997) show preferential jet stream interaction with tropically forced waves of periods > 30 days (in contrast to < 30 day periods) in a barotropic model. Could such behavior have an accumulated effect on the seasonal mean?

For other ENSO events, differences are qualitatively similar although the consistency between differences in tropical divergence and North Pacific upper-level height for ESUBFIX is not found (not shown) suggesting that high frequency variability may be important. When interpreting different events, we keep in mind that characteristics of Q variability have some dependence on the mean and underlying SST, and variability of the extratropical response is dependent on the basic state circulation modified by ENSO as well as changes Q variability. This was examined by Schubert *et al.* (2001) who also investigated the role of sub-seasonal variability during the 1982/83 El Niño, though any potential impact on the ensemble mean signal itself was not addressed. Another thing to keep in mind is that Q variability is highly dependent on the model's convective parameterization, *e.g.* much more ensemble variability is generally simulated during El Niño using CAM4.0 compared to CAM3.1 (not shown). Nevertheless, the use of a state-of-the-art nonlinear moist GCM for this particular work is a strength and an advantage over past studies.

In order to more adequately address the role of sub-seasonal and intra-ensemble Q variability, more work is currently under way examining relationships with the extratropical circulation at sub-seasonal time scales with emphasis on its consequences on the seasonal mean response. This is not trivial provided the large degree of chaotic internal variability in the extratropics that is completely independent of tropical Q, however an adequately large ensemble size may allow potential relationships to be distinguishable.

References

Li, L., and T. R. Nathan, 1997: Effects of low-frequency tropical forcing on intraseasonal tropical–extratropical interactions. *J. Atmos. Sci.*, **54**, 332–346.

Schubert, S. D., M. J. Suarez, Y. Chang, and G. Branstator, 2001: The impact of ENSO on extratropical low-frequency noise in seasonal forecasts. *J. Climate*, **14**, 2351–2365.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Precipitation Characteristics of the South American Monsoon System Derived from Multiple Data Sets

Leila M. V. Carvalho[1, 2], Charles Jones[2], Adolfo N. D. Posadas[3],
Roberto Quiroz[3], Bodo Bookhagen[1, 2] and Brant Liebmann[4]

[1]*Department of Geography, University of California, Santa Barbara, CA, USA*
[2]*Earth Research Institute, University of California, Santa Barbara, CA, USA*
[3]*International Potato Center (CIP), Lima, Peru*
[4]*CIRES Climate Diagnostics Center, Boulder, CO, USA*

1. Introduction

The monsoon (hereafter, South American Monsoon System, SAMS) is the most important climatic feature in South America (Zhou and Lau 1998; Vera *et al.* 2006; Marengo *et al.* 2010). The main feature of the SAMS is the enhanced convective activity and heavy precipitation in tropical South America, which typically starts in October-November, is fully developed during December-February and retreats in late April or early May (Kousky 1988; Horel *et al.* 1989; Marengo *et al.* 2001; Grimm *et al.* 2005; Gan *et al.* 2006; Liebmann *et al.* 2007).

Although the variability of precipitation in the SAMS has been extensively investigated over the years, one of the main challenges has been the availability of data sets with suitable spatial and temporal resolutions able to resolve the large range of meteorological systems observed during the monsoon. While some stations in South America have precipitation records going back several decades, the sparseness of stations is not adequate to characterize mesoscale precipitation systems. To overcome this difficulty, some studies have developed considerable efforts to collect precipitation records from stations and develop quality-controlled gridded precipitation data sets (Legates and Willmott 1990; Liebmann and Allured 2005; Silva *et al.* 2007).

Recently, new generation of reanalysis products have been completed (Saha *et al.* 2010; Dee *et al.* 2011; Rienecker *et al.* 2011). The new reanalyses, which are derived from state-of-the-art data assimilation systems and high resolution climate models, provide substantial improvements in the spatiotemporal variability of precipitation relative to the first generation of reanalyses (Higgins *et al.* 2010; Saha *et al.* 2010; Rienecker *et al.* 2011; Silva *et al.* 2011).

This paper evaluates and compares statistical properties of daily precipitation in three types of data sets: gridded station data, satellite-derived precipitation and reanalyses. This study employs several analyses to determine consistencies and disagreements in the representation of precipitation over SAMS. The period 1998-2008 is selected in order to minimize missing data and develop a consistent comparison among the data sets. In addition, since the data sets are available with different horizontal resolutions, the comparison is performed in two ways: 1) all data sets regridded to a common resolution and 2) data sets with their original resolutions.

2. Data

The statistical properties of precipitation in the SAMS region are investigated with daily gridded data from multiple sources during 1 Jan-31 Dec 1998-2008. The following data sets are used:

i) Physical Sciences Division, Earth System Research Laboratory (PSD): This data set is formed from observed precipitation collected at stations distributed over South America (Liebmann and Allured 2005, 2006). The daily gridded precipitation is constructed by averaging all observations available within a specified radius of each grid point. Two grid resolutions (1° and 2.5° lat/lon) are used in this study.

Correspondence to: Leila M. V. Carvalho, Department of Geography, University of California, Santa Barbara, CA 93106, USA; Email: leila@eri.ucsb.edu.

ii) Global Precipitation Climatology Project (GPCP): The daily GPCP combines Special Sensor/Microwave Imager (SSM/I), GPCP Version 2.1 Satellite-Gauge, geosynchronous-orbit Infrared (IR), (geo-IR) Tb histograms (1°x1° grid in the band 40°N-40°S, 3-hourly), low-orbit IR GOES Precipitation Index (GPI), TIROS Operational Vertical Sounder (TOVS) and Atmospheric Infrared Sounder (AIRS) data (Huffman *et al.* 2001). The GPCP data used in this study have 1° lat/lon grid spacing.

iii) Climate Prediction Center unified gauge (CPC-uni): The NOAA Climate Prediction Center (CPC) unified gauge uses an optimal interpolation technique to re-project precipitation reports to a grid (Higgins *et al.* 2000; Silva *et al.* 2007; Chen *et al.* 2008; Silva *et al.* 2011). This study uses data with 0.5° lat/lon grid spacing. Although the PSD and CPC-uni data sets share some of the same station observations, it is worth noting that the quality control and gridding methods are distinct. In addition, it is likely that the number and origin of station data in both data sets are different.

iv) Climate Forecast System Reanalysis (CFSR): Daily precipitation from the NCEP CFSR (Saha *et al.* 2010) is used at 0.5° lat/lon grid spacing. It is also important to note that precipitation is not assimilated in the CFSR production and is a forecast (first-guess) product.

v) Modern-Era Retrospective Analysis for Research and Applications (MERRA): Daily precipitation from MERRA at 0.5° latitude/0.3° longitude is used (Rienecker *et al.* 2011). As in the CFSR, precipitation is a forecast product.

vi) Tropical Rainfall Measurement Mission (TRMM 3B42 V6): Daily precipitation from TRMM is used with 0.25° lat/lon (Bookhagen and Strecker 2008; Bookhagen and Strecker 2010; Bookhagen and Burbank 2011).

3. Results

Several statistical analysis have been applied to the daily precipitation and the reader is referred to Carvalho *et al.* (2011) for additional details. The annual evolution of SAMS is examined to determine consistencies and disagreements among the data sets. The large-scale features of interest are: the dominant spatial precipitation pattern, dates of onset and demise, duration and amplitude of the monsoon. These characteristics are determined with empirical orthogonal functional (EOF) analysis applied to the daily precipitation (only land grid points) from each data set separately. Before computation of EOF analysis, the time series of precipitation in each grid point are scaled by the square-root of the cosine of the latitude and the long-term mean removed (1 Jan-31 Dec, 1998-2008). The first mode (EOF1) and associated temporal coefficient (PC1) explain the largest fraction of the total variance of precipitation over land and are used to describe the annual evolution of SAMS.

To determine dates of onset, demise and duration of SAMS, the daily PC1 is smoothed with ten passes of a 15-day moving average. This smoothing procedure is obtained empirically and used to decrease the influence of high frequency variations during the transition phases of SAMS. The large-scale onset of SAMS is defined as the date when the smoothed PC1 changes from negative to positive values. This implies that positive precipitation anomalies during that time become dominant over the SAMS domain. Likewise, the demise of SAMS is defined as the date when the smoothed PC1 changes from positive to negative values. The duration of the monsoon is defined as the period between onset and demise dates. The seasonal amplitude of the monsoon is defined as the integral of positive unsmoothed PC1 values from onset to demise. Therefore, the seasonal amplitude index represents the sum of positive precipitation anomalies and minimizes the effect of "break" periods in the monsoon especially near the onset and demise. Active/break periods in SAMS are particularly frequent on intraseasonal time scales (Jones and Carvalho 2002).

Figure 1 shows the spatial patterns of EOF1 derived from each data set with 2.5° lat/lon grid spacing and expressed as correlations between PC1 and precipitation anomalies. Positive correlations are interpreted as positive precipitation anomalies and indicative of active SAMS. In general, all data sets show similar features such as positive precipitation anomalies over central South America and negative anomalies over the northern parts of the continent. The region of negative anomalies over northern South America is substantially smaller in the PSD due to missing data (the "bull's eye" at ~60°W, 10°S is a grid point with missing data). The

magnitude of positive correlations varies slightly and is highest for PSD. The largest positive correlation in MERRA is slightest to the west relative to the other data sets.

The percentages of explained variance by EOF1 are: 20.5% (PSD), 11.6% (GPCP), 8.4% (CPC-uni), 10% (CFSR), 17.9% (MERRA) and 6.9% (TRMM). EOF1 captures the largest fraction of the total variance, which includes subseasonal, seasonal and interannual variations, since the EOF analysis is performed removing only the long-term mean. Main differences in explained variance are associated with how much each PC1 represents the distribution of subseasonal, seasonal and interannual variations. These percentages are comparable to the percentages obtained with the data sets at their original resolutions, which suggests that spatial resolution of the data sets is not the main issue, but rather how each data set represents temporal variations.

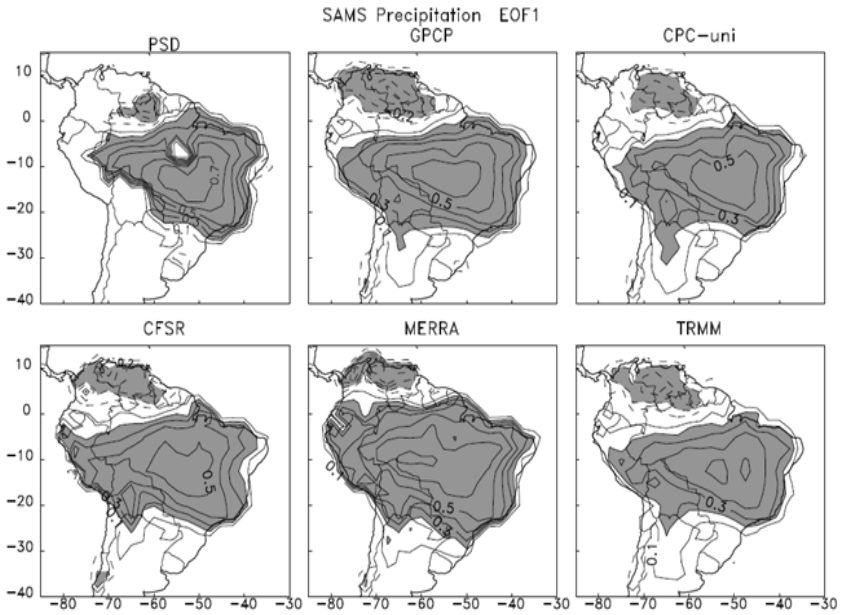


Fig. 1 First EOF patterns described as correlations between the first temporal coefficient (PC1) and precipitation anomalies. Solid (dashed) contours indicate positive (negative) correlations at 0.1 intervals (zero contours omitted). Shadings indicate correlations ≥ 0.2 (≤ -0.2) and are significant at 5%. Data grid spacing: 2.5° lat/lon.

Dates of onset, demise and duration of SAMS derived from each data set with the original resolution are shown in Fig. 2. The mean onset date (Fig.2 top) is highly coherent among PSD, GPCP, CPC-uni and TRMM (~21 October) including the ranges of minimum and maximum onset dates. In contrast, the mean onset dates in CFSR and MERRA are off by several weeks. The variability in dates of mean demise (Fig. 2 middle) indicates agreements among PSD, GPCP, CPC-uni and TRMM and some differences in CFSR and large disagreement in MERRA. Consequently, the mean durations of SAMS (~180 days) agree reasonably well among PSD, GPCP, CPC-uni and TRMM data and is shorter and more variable in the CFSR and MERRA reanalyses (Fig. 2 bottom). These results indicate that differences in data resolution do not explain disagreements in the annual evolution of SAMS especially between CFSR and MERRA and the other data sets.

5. Conclusions

Carvalho *et al.* (2011) compares some statistical properties of daily gridded precipitation from different data (1998-2008): 1) Physical Sciences Division, Earth System Research Laboratory (PSD) (1.0^o and 2.5^o lat/lon), 2) Global Precipitation Climatology Project (GPCP at 1^o lat/lon), 3) Climate Prediction Center unified gauge (CPC-uni) (0.5^o lat/lon), 4) NCEP CFSR reanalysis (0.5^o lat/lon), 5) NASA MERRA reanalysis (0.5^o lat/0.3^o lon) and 6) TRMM 3B42 V6 data (0.25^o lat/lon). The same statistical analyses are applied to data in: 1) a common 2.5^o lat/lon grid and 2) in the original resolutions of the data sets.

All data sets consistently represent the large-scale patterns of the SAMS. The onset, demise and duration of SAMS are consistent among PSD, GPCP, CPC-uni and TRMM data sets, whereas CFSR and MERRA seem to have problems in capturing the correct timing of SAMS. Power spectrum analysis shows that intraseasonal variance is somewhat similar in the six data sets. Moreover, differences in spatial patterns of mean precipitation are small among PSD, GPCP, CPC-uni and TRMM data and some discrepancies are found CFSR and MERRA. Fitting of gamma frequency distributions to daily precipitation shows differences in the parameters that characterize the shape, scale and tails of the frequency distributions. This suggests that

significant uncertainties exist in the characterization of extreme precipitation, an issue that is highly important in the context of climate variability and change in South America.

Acknowledgements. L. M. V. Carvalho, C. Jones and B. Liebmann thank the support of NOAA's Climate Program Office (NA07OAR4310211 and NA10OAR4310170). L.M.V Carvalho, C. Jones, A. Posadas and R. Quiroz thank USAID-CIP (Sub-Contract SB100085). NCEP/NCAR Reanalysis and OLR data were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA (www.esrl.noaa.gov). TRMM data was acquired by an international joint project sponsored by the Japan National Space Development Agency (NASDA) and the US National Aeronautics Space Administration NASA) Office of Earth Science. The help from Bob Dattore, NCAR-CISL, in providing the NCEP CFSR data is greatly appreciated.

References

Bookhagen, B., and M. R. Strecker, 2008: Orographic barriers, high-resolution TRMM rainfall, and relief variations along the eastern Andes. *Geophys Res Lett*, **35,** -.

Bookhagen, B., and M. R. Strecker, 2010: Modern Andean rainfall variation during ENSO cycles and its impact on the Amazon Basin. Neogene history of Western Amazonia and its significance for modern diversity. C. Hoorn, H. Vonhof and F. Wesselingh. Oxford, U.K., Blackwell Publishing.

Bookhagen, B., and D. W. Burbank, 2011: Towards a complete Himalayan hydrologic budget: The spatiotemporal distribution of snow melt and rainfall and their impact on river discharge. *Journal of Geophysical Research-Earth Surface, doi:10.1029/2009JF001426, J Geophys Res-Earth, 115*.

Carvalho, L. M. V., C. Jones, A. N. Posadas, R. Quiroz, B. Bookhagen, and B. Liebmann, 2011: Precipitation characteristics of the South Monsoon System derived from multiple data sets. Journal of Climate (in review).

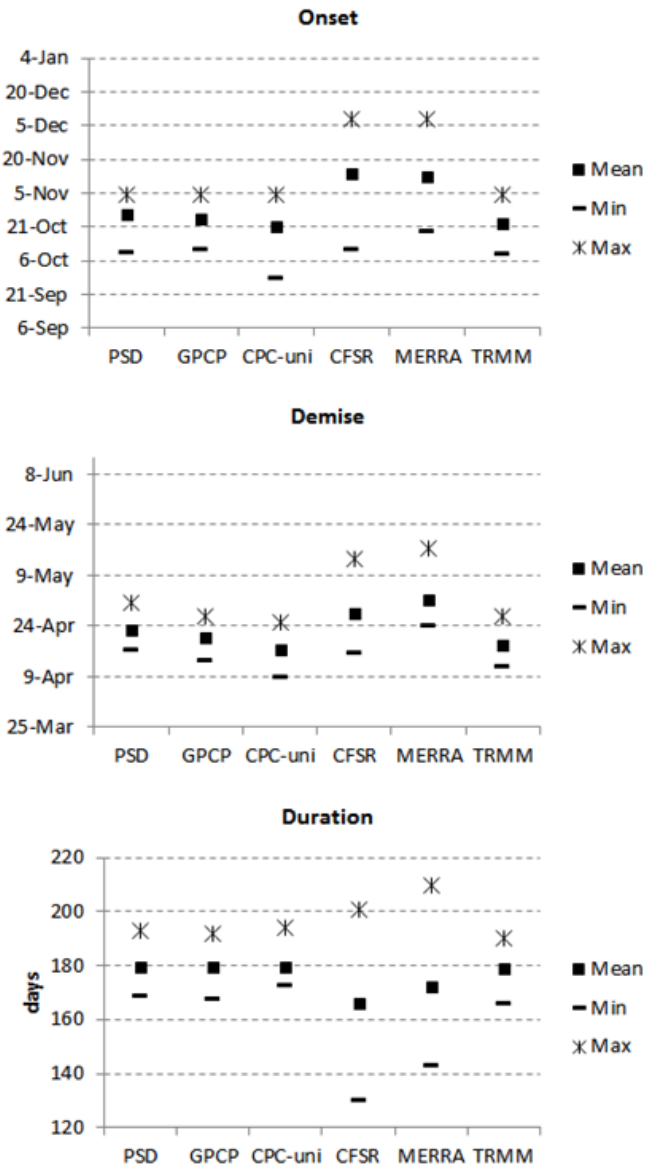


Fig. 2 Top: mean (squares), minimum (dash) and maximum (asterisk) dates of SAMS onset. Middle: mean (squares), minimum (dash) and maximum (asterisk) dates of SAMS demise. Bottom: mean (squares), minimum (dash) and maximum (asterisk) durations of SAMS. Data sets are indicated in the horizontal axis. Data sets have different grid spacings.

Chen, M. Y., W. Shi, P. 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.-Atmos.*, **113,** D04110, 04111-04113.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society. *Quart. J. Roy. Meteor. Soc.*, **137,** 553–597, doi: 510.1002/qj.1828.

Gan, M. A., V. B. Rao, and M. C. L. Moscati, 2006: South American monsoon indices. *Atmos. Sci. Let.*, **6,** 219-223.

Grimm, A. M., C. S. Vera, and C. R. Mechoso, 2005: *The South American Monsoon System.* In: The American Monsoon Systems: An Introduction. C.-P. Chang, B. Wang and N.-C. G. Lau, Eds WMO/TD No. 1266 (TMRP Report No. 70), 197-206.

Higgins, R. W., W. Shi, E. Yarosh, and R. Joyce, 2000: Improved United States precipitation quality control system and analysis. *NCEP/Climate Prediction Center ATLAS No. 7,* National Oceanic and Atmospheric Administration, 40 pp.

Higgins, R. W., V. E. Kousky, V. B. S. Silva, E. Becker, and P. Xie, 2010: Intercomparison of Daily Precipitation Statistics over the United States in Observations and in NCEP Reanalysis Products. *J. Climate*, **23,** 4637-4650.

Horel, J. D., A. N. Hahmann, and J. E. Geisler, 1989: An investigation of the Annual Cycle of Convective Activity over the Tropical Americas. *J. Climate*, **2,** 1388-1403.

Huffman, G. J., and Coauthors, 2001: Global Precipitation at One-Degree Daily Resolution from Multisatellite Observations. *J. Hydrometeorol.*, **2,** 36-50.

Jones, C., and L. M. V. Carvalho, 2002: Active and break phases in the South American Monsoon System. *J. Climate*, **15**, 905-914.

Kousky, V. E., 1988: Pentad outgoing longwave radiation climatology for the South American sector. *Rev. Bras. Met.*, **3,** 217-231.

Legates, D. R., and C. J. Willmott, 1990: Mean Seasonal and Spatial Variability in Gauge-Corrected, Global Precipitation. *Int. J. Climatol.*, **10,** 111-127.

Liebmann, B., and D. Allured, 2005: Daily precipitation grids for South America. *Bull. Amer. Meteor. Soc.*, **86,** 1567-1570.

——, 2006: Daily precipitation grids for South America - Reply. *Bull. Amer. Meteor. Soc.*, **87,** 1096-1096.

Liebmann, B., and Coauthors, 2007: Onset and end of the rainy season in South America in observations and the ECHAM 4.5 atmospheric general circulation model. *J. Climate*, **20,** 2037-2050.

Marengo, J. A., B. Liebmann, V. E. Kousky, N. P. Filizola, and I. C. Wainer, 2001: Onset and End of the Rainy Season in the Brazilian Amazon Basin. *J. Climate*, **14,** 833-852.

Marengo, J. A., and Coauthors, 2010: Review: Recent developments on the South American monsoon system. *Int. J. Climatol.,* DOI: 10.1002/joc.2254.

Rienecker, M. M., and Coauthors, 2011: MERRA - NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate,* (in press).

Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91,** 1015-1057.

Silva, V. B. S., V. E. Kousky, and R. W. Higgins, 2011: Daily Precipitation Statistics for South America: An Intercomparison between NCEP Reanalyses and Observations. *J. Hydrometeorol.*, **12,** 101-117.

Silva, V. B. S., V. E. Kousky, W. Shi, and R. W. Higgins, 2007: An Improved Gridded Historical Daily Precipitation Analysis for Brazil. *J. Hydrometeorol.*, **8,** 847-861.

Vera, C., and Coauthors, 2006: Toward a unified view of the American Monsoon Systems. *J. Climate*, **19,** 4977-5000.

Zhou, J. Y., and K. M. Lau, 1998: Does a monsoon climate exist over South America? *J. Climate*, **11,** 1020-1040.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Dynamic Linkage between the Sahel Greening and Intense Atlantic Hurricanes

Shih-Yu Wang[1, 2] and Robert R. Gillies[1, 2]

[1]*Utah Climate Center, Utah State University, Logan, UT*

[2]*Department of Plants, Soils, and Climate, Utah State University, Logan, UT*

1. Introduction

Beginning around the early 1980s, the Sahel belt of Africa has undergone a continual increase in both precipitation and vegetation greening (Herrmann *et al.* 2005; Olsson *et al.* 2005). This so-called Sahel greening has been a signal of a gradual recovery from very dry conditions in the 1980s, and from the days of prolonged droughts and famines during the late 1960s–mid-1990s (Nicholson and Yin 2001; Dai *et al.* 2004). The increase in Sahel rainfall has been attributed to a combination of factors such as global warming (Lu and Delworth 2005; Hagos and Cook 2008), decadal-scale variability in the global sea surface temperature (SST) (Giannini *et al.* 2003; Hoerling *et al.* 2006; Cook 2008; Latif *et al.* 2010) and associated changes in the carbon cycle (Schimel *et al.* 2001). Previous studies have also pointed to a robust connection between the low-frequency rainfall variability in the Sahel and the Atlantic Multi-decadal Oscillation (AMO) – a basin-scale pattern of SST variability driven by Atlantic meridional overturning (Knight *et al.* 2005).

As shown in Fig. 1, the summer precipitation differences between 1995-2009 and 1979-1994 illustrates the northward displacement of the oceanic rainband that is connected to the wetness domain along the Sahel. Precipitation in the Sahel is produced through a set of complex interactions between different circulation systems – in particular, the Tropical Easterly Jet (TEJ) in the upper troposphere that extends from South Asia, the African Easterly Jet (AEJ) in the middle troposphere, and the tropical monsoon westerly in the lower troposphere that lies beneath the TEJ; these circulation features are depicted in Fig. 1. African Easterly Waves (AEWs) contribute to about 50% of the June-September rainfall in West Africa (Chen and Wang 2007). Moreover, AEWs are important precursors to intense hurricanes of Category 3 or above (Landsea and Gray 1992; Landsea 1993; Chen *et al.* 2008) and recent studies (Hopsch *et al.* 2007; Fink et al. 2010) have found a positively correlated, yet fluctuating relationship between the tropical cyclone activity, the West African monsoon, and AEWs.

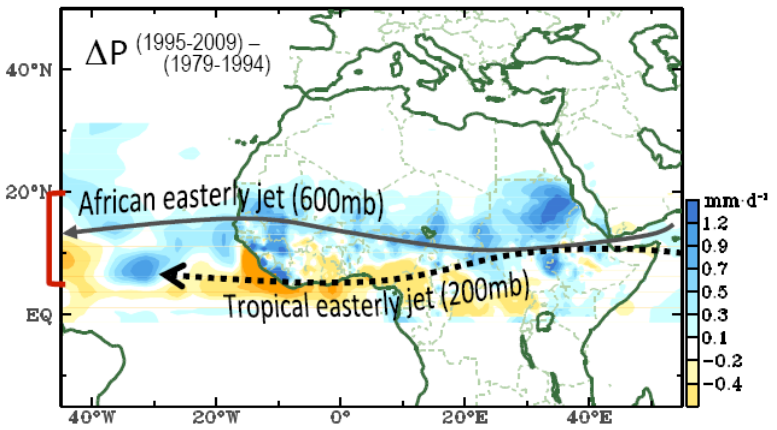


Fig. 1 July-August precipitation differences between the periods 1995-2009 and 1979-1994, from the University of Delaware data over land and the GPCP data over ocean. Jet cores along the African Easterly Jet and the Tropical Easterly Jet are indicated. The tropical monsoon westerly lies underneath the TEJ. The left boundary in red outlines the latitude zone used in Figs. 2-4.

Here we report a recent study by Wang and Gillies (2011) that examined the synoptic environment associated with the increasing precipitation in the Sahel. We also investigated the extent of any downstream effects the Sahel climate change has had on the increase in tropical cyclone threat (Emanuel 2005, 2007). Precursors to the analysis were first the adoption of several precipitation datasets and different generation global reanalyses as described in Section 2. The analysis in Section 3 is followed by a discussion of the

Correspondence to: Shih-Yu (Simon) Wang, Utah Climate Center & Department of Plants, Soils, and Climate, Utah State University, Logan, UT; E-mail: simon.wang@usu.edu.

temporal-spatial evolutions of prescribed meteorological variables and their links to Atlantic tropical cyclones. Concluding remarks are presented in Section 4.

2. Data sources

Our study utilized three sets of precipitation data: the Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP; Xie and Arkin 1997), the Global Precipitation Climatology Project (GPCP) version 2 (Adler *et al.* 2003), and gauge-based precipitation compiled by the University of Delaware (UDel) (Legates and Willmott 1990). For verification purposes, we adopted the un-interpolated version of outgoing longwave radiation (OLR) measured from NOAA's polar orbiting satellites. For sea surface temperatures we utilized the NOAA Extended Reconstructed SST V3b (Smith and Reynolds 2003); for surface temperatures over land we used the CRU Air Temperature Anomalies Version 3 (*Brohan et al.* 2006) with the time period up to present. In the later part of the study we examined the occurrences of Atlantic tropical cyclones. Information of tropical cyclones was obtained from two sources: 1) the Atlantic Hurricane Dataset Re-analysis Project (Landsea *et al.* 2004) for the period 1979-2009, and 2) the National Hurricane Center's Tropical Prediction Center for 2010.

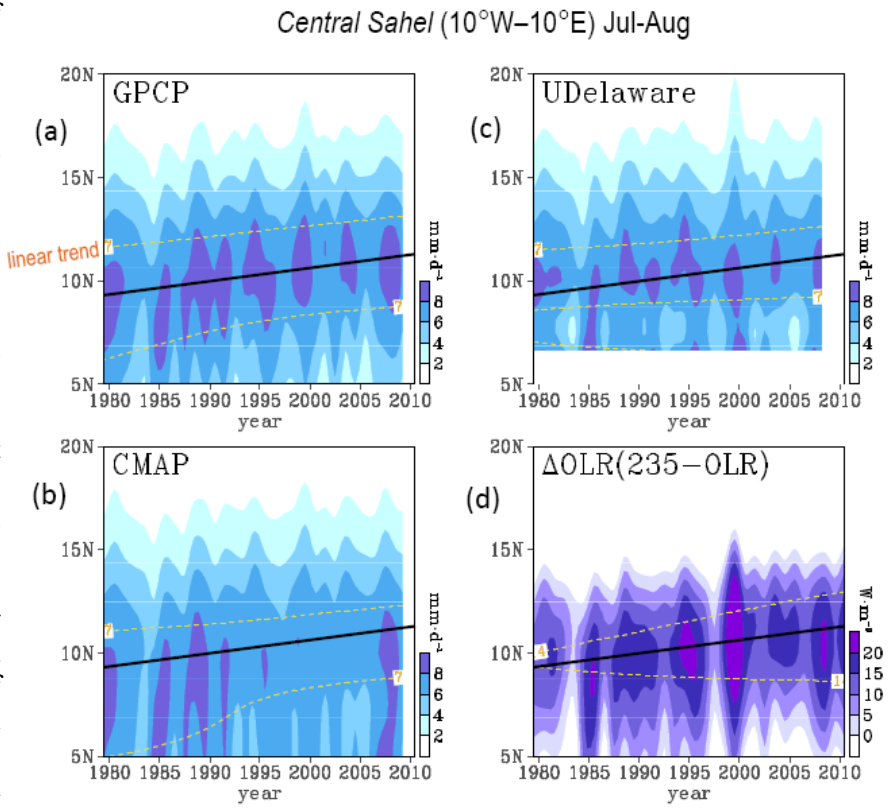


Fig. 2 Latitude-time profiles of precipitation averaged between 10°W and 10°E for the July-August season, derived from (a) GPCP, (b) CMAP, (c) University of Delaware, and (d) ΔOLR (= 235 Wm^{-2} – OLR). Yellow dashed lines are linear trends of the designated value. The solid black lines indicate the average latitudinal position of maximum precipitation for the depiction of the Sahel rainband.

Five global reanalysis datasets (hereafter referred to as reanalyses) were utilized: (1) the National Centers for Environmental Prediction/National Center for Atmospheric Research Global Reanalysis (NCEP1; Kalnay *et al.* 1996), (2) the NCEP/Department of Energy Global Reanalysis II (NCEP2) – the improved version of NCEP1 that included additional satellite-derived atmospheric information and newer physics schemes (Kanamitsu *et al.* 2002), (3) the European Centre for Medium-Range Weather Forecasts (ECMWF) 40-yr Reanalysis (ERA-40; Uppala *et al.* 2005), (4) the ERA-Interim reanalysis (Simmons *et al.* 2007), and (5) the Modern Era Retrospective-analysis for Research and Applications (MERRA) developed by NASA (Rienecker *et al.* 2011). All of the reanalyses are daily means. Because ERA-40 is only available up to 2002, the period 1989-2010 was merged with ERA-Interim by interpolating ERA-Interim's higher resolution onto ERA-40's 2.5° resolution through a bilinear approach. This merged reanalysis is referred to as ERA40/I.

3. Results

a. Changes in the Sahel climate

The change of precipitation during July-August (*i.e.* Sahel's rainy season) of 1979-2010 is illustrated in Fig. 2 as latitude-time diagrams of GPCP, CMAP, UDel and ΔOLR (=235 Wm^{-2} – OLR), averaged between 10°W and 10°E representing the central Sahel. A clear northward shift is revealed in these precipitation data

and proxy, without a discernable change in the amount (as inferred from the parallel trends). This feature is indicative of the fact that the increasing Sahel rainfall is not an expansion of the seasonal rainband but rather is a result of a positional shift. On the other hand, ΔOLR reveals a noticeable increase in magnitude in the precipitation condition (as noted from the diverging trends), suggestive of an intensification in the convective activity within the migrating rainbands. The latitudinal shift of the rainfall maximum – averaged from all four datasets – amounts to about +1.1° over the 32 year period and is indicated by the solid black line in Fig. 2. However, an examination of the model-generated precipitation from the four reanalyses (not shown) revealed a systematic bias that consisted of a southward displacement of the mean rainband (*i.e.* positioned at 7.5°N instead of 11°N as is the case in Fig. 2), without a discernable positional shift. Such reanalyses biases highlight a primary challenge of global climate models when it comes to producing reliable climate simulations and projections. For this reason we decided to use only the kinetic fields of the reanalyses.

Figure 3 shows changes in the TEJ, AEJ, and the tropical monsoon westerly in terms of zonal winds at (a) 200mb, (b) 600mb, and (c) 850mb, respectively, from the four reanalyses over the central Sahel. The point to note here is that there is a substantial increase in the TEJ and a slight southward shift of the jet core, although MERRA depicts less of a positional shift but does show a much stronger jet speed compared to the other reanalyses. NCEP2 and ERA40/I depict a similar structure of the TEJ but are somewhat different from the other reanalyses. Even though the TEJ has intensified, its southward expansion does not seem to correspond to the decreased tropical rainfall (south of 10°N; Fig. 2). For the AEJ (Fig. 3b), all the reanalyses, with the exception of MERRA, depict a northward shift of the jet consistent with the migrating rainband. Both NCEP2 and ERA40/I depict a clear intensification of the AEJ, whereas the intensification is weaker in NCEP1 and MERRA. For the tropical monsoon westerly, the reanalyses are somewhat divergent – while NCEP1 points to an increase in the westerly wind speed, ERA40/I indicates a sharp decrease. Nevertheless, all reanalyses agree upon a quasi-stationary behavior of the monsoon westerly. It appears that the four reanalyses exhibit somewhat different characteristics of the African

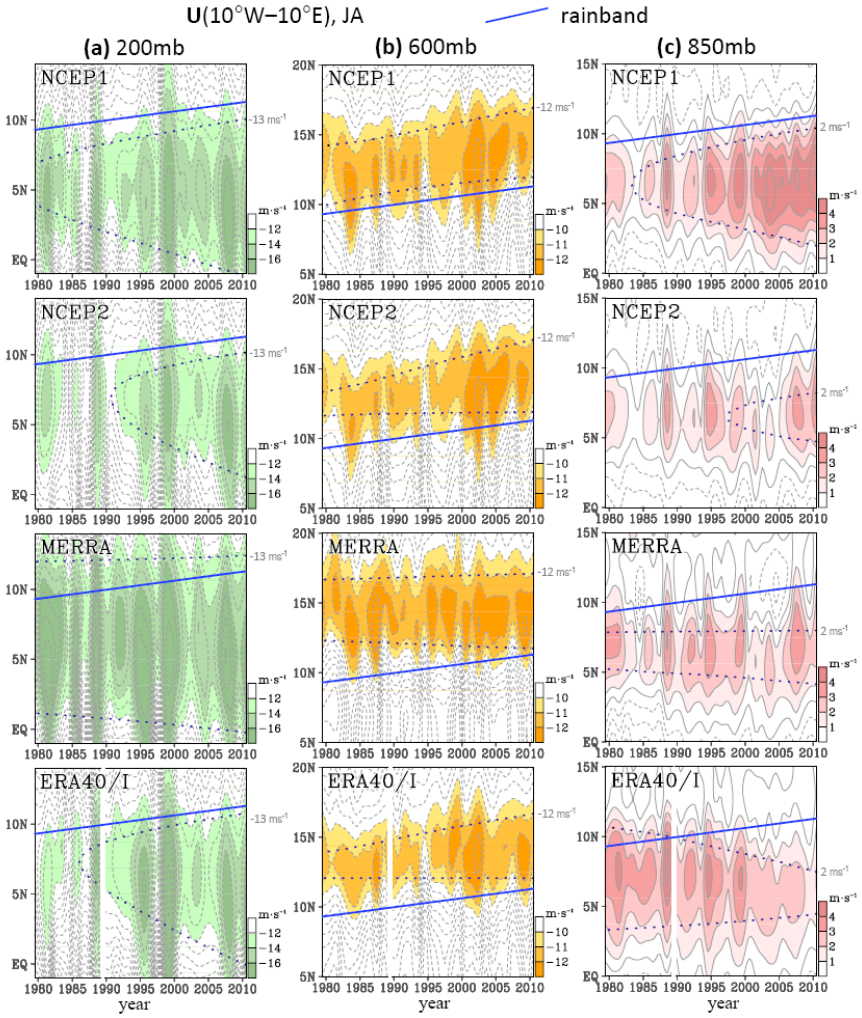


Fig. 3 Same as Fig. 2 but for zonal wind speeds at levels of (a) 200 mb, (b) 600 mb, and (c) 850 mb derived from (top-down) NCEP1, NCEP2, MERRA, and ERA40/I. The solid blue lines indicate the Sahel rainband migration. The dotted lines indicate linear trends of the designated value. Note the different latitudinal scales for each level. In ERA40/I, year 1989 is masked out to highlight the transition from ERA40 to ERA-Interim.

circulations, while newer reanalyses do not appear to correspond with each other, such as the AEJ's trend between MERRA and ERA40/I. Therefore, in the ensuing analysis we adopted an ensemble approach by averaging the four reanalyses with an equal weight, denoted as the *ensemble reanalyses*.

Recall from Fig. 2 that, although the precipitation amounts have not increased significantly, ΔOLR has intensified. The root-mean-square (RMS) of daily OLR over the central Sahel (Fig. 4a) supports this observation as it reveals an enhanced fluctuation along the northern edge of the migrating rainband (~15°N). This likely is due to the rainband moving towards the Saharan boundary where stronger potential temperature gradients and lower static stabilities are present; this subsequently leads to greater thermal and convective instabilities (discussed later). To examine the change in moist convection, we computed the frequency from which the daily OLR values at each grid point were lower than 200 Wm^{-2}, which is an empirical threshold obtained for deep convection that usually occurs over tropical oceans (Zhang 1993). The result (Fig. 4b) portrays a northward migration in the frequency of intense convection that is coupled to the migrating rainband (*i.e.* the blue line). The eastern Sahel (not shown) undergoes consistent changes in convective activity.

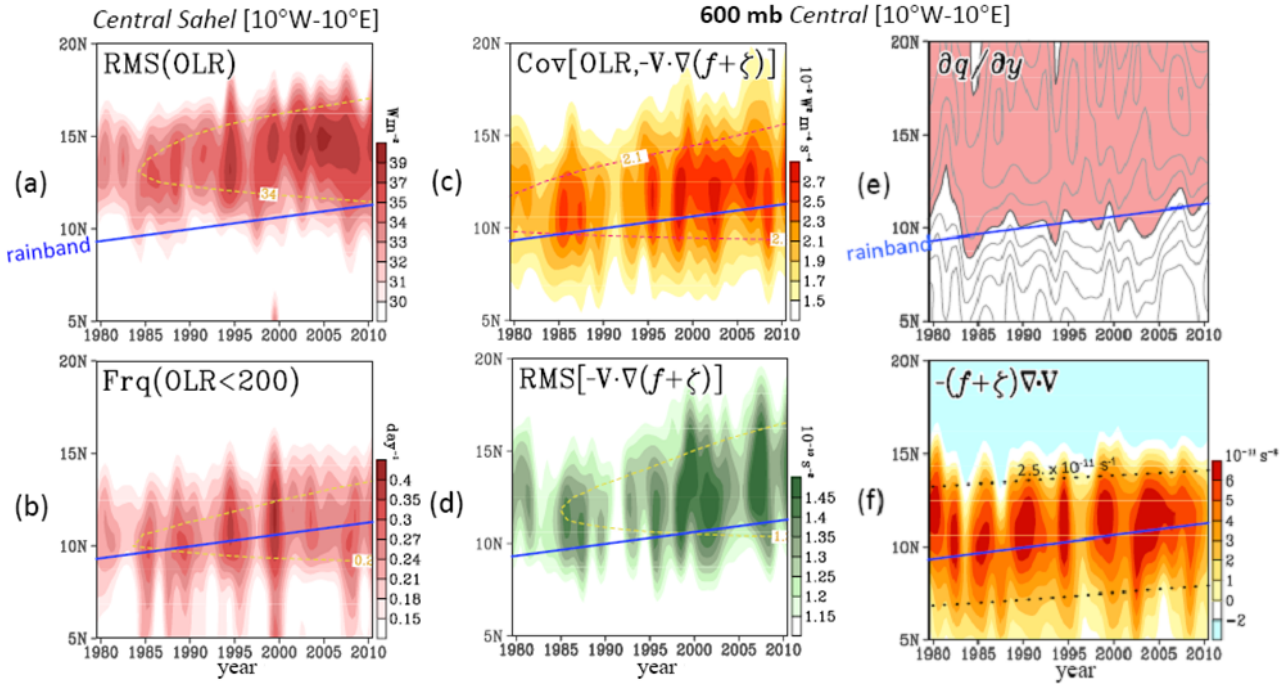


Fig. 4 Same as Fig. 3 but for (a) root-mean-square (RMS) of the 2-8 day bandpass filtered OLR, (b) frequency of the OLR values lower than 200 Wm-2, (c) covariance of the filtered OLR with horizontal vorticity advection at 600 mb, (d) RMS of the filtered horizontal vorticity advection at 600 mb, (e) meridional gradient of Ertel potential vorticity at 600 mb, and (f) vorticity source due to vortex stretching at 600 mb. Dashed lines are linear trends of the designated value. The Sahel rainband is indicated by solid blue lines.

It is known that, within the latitude zone of 10°-15°N, AEWs contribute a significant portion to the seasonal precipitation (Grist *et al.* 2002; Chen and Wang 2007). The association of the changing convective activity with AEWs was examined by calculating the covariance of the ΔOLR in conjunction with the vorticity advection forcing of vorticity tendency at 600 mb, $-\mathbf{V}\cdot\nabla(f+\zeta)$, where **V** indicates horizontal wind vectors, f is the planetary vorticity, and ζ is the relative vorticity. Both variables were bandpass filtered with 2-8 days in order to isolate the signal of AEWs. As is shown in Fig. 4c, a substantial increase and a northward position shift are observed in the covariance of ΔOLR and vorticity advection. This suggests an expansion and northward shift of the wave-convection interaction. The RMS of the filtered vorticity advection (Fig. 4d) shows a similar shift with an amplifying tendency that may be interpreted as a northward displacement of the AEW track accompanied by an amplification of those waves. The wave activity in the eastern Sahel – *i.e.* one

of the breeding zones of AEWs – is generally weaker than in the central Sahel but shows a similar tendency: a northward shift and a latitudinal expansion. Chen and Wang (2007) have noted a 20% increase in the number of "moist" AEWs that occur south of the AEJ (the so-called southern track). These results indicate that the migrating and intensifying convection, AEJ, and AEWs are connected with the shifting Sahel rainband and, that convection coupled to the AEJ-AEW system has become stronger over the northern part of the Sahel.

b. Changes in the dynamic structure

It is well established that the formation mechanism of AEWs, particularly the southern track, relies upon the Charney-Stern instability (Charney and Stern 1962) in which the meridional gradient of potential vorticity changes sign south of the AEJ and is negative underneath the AEJ (Burpee 1972; Thorncroft and Hoskins 1994a, b). The Charney-Stern instability occurs in an environment of combined thermal gradient and vertical shear that promotes the release of available potential energy (from the mean circulation) towards any pressure perturbations. Therefore, we examined the change in this dynamics through the meridional gradient of potential vorticity, dq/dy, at 600 mb (Fig. 4e). Compared with the AEJ as depicted in Fig. 3b, the sign change of dq/dy indeed occurs south of the jet while exhibits a northward progression corresponding to the AEJ's migration. Noteworthy is that the intensified AEJ may also induce stronger shear instability south of the jet; this being favorable for the development of AEWs. As has been shown previously (Chen 2006), mid-tropospheric vortex stretching is a dominant forcing source in the vorticity budget over the AEW genesis region. An examination of the seasonal mean vortex stretching at 600 mb reveals an amplifying tendency associated with a northward migration (Fig. 4f), with a more pronounced amplification in the eastern Sahel. These features correspond well with the intensification and positional shift of AEWs revealed from Figs. 4c and 4d. Since the AEW activity is positively correlated with the Sahel rainfall (on the interannual timescale; Thorncroft and Rowell 1998), these also seem to support the position shift of the Sahel rainband.

During the last 30 years, surface temperature over the Saharan desert has warmed within the range of 0.5°-1.0°C (not shown); a warming like this strengthens the heat low and lowers the static stability. The change in warming and static stability [from the combination of low static stability and a stronger meridional temperature gradient; Chang 1993] likely reinforces the heat low's interaction with the northward migrating AEJ, leading to an enhancement in baroclinic instability. When coupled with the increased moisture supply and enhanced moist convection conditions, mixed barotropic and baroclinic instabilities likely result in a stronger reversal of the mid-tropospheric potential vorticity gradient, thus enhancing dq/dy. This appears to be the case as is evident in Fig. 4e. Such processes further explain the intensified AEW-OLR activity suggested from Fig. 4c.

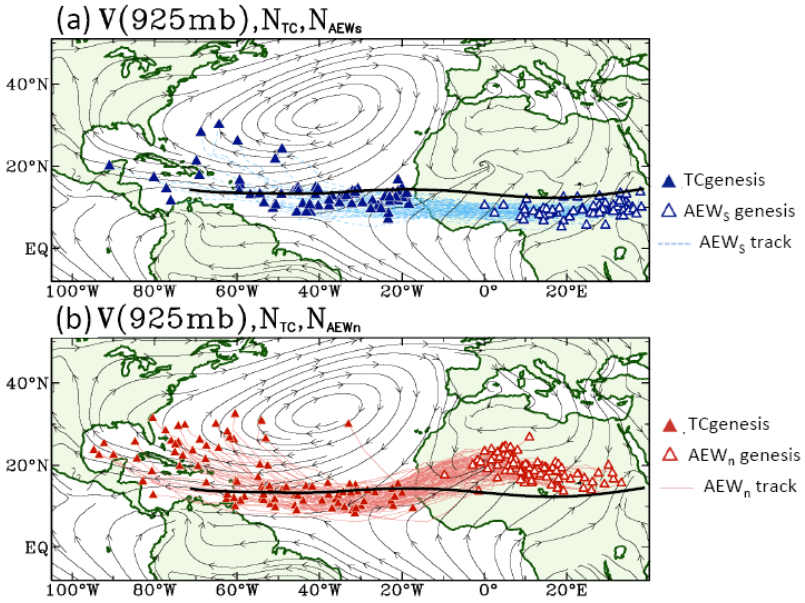


Fig. 5 July-September climatological streamlines at 925mb overlaid with tropical cyclogeneses of AEW origin and the trajectory and genesis of AEWs for (a) the southern track (AEWS) and (b) the northern track (AEWn). Symbols are explained in the right. The AEJ core is indicated by the black solid line. Data period: 1979-2006. (Modified from Chen *et al.* 2008)

c. Link with Atlantic tropical cyclones

During the past 30 years, both the frequency and the intensity of Atlantic tropical cyclones have increased significantly. A growing number of studies have suggested that tropical cyclone activity and Sahel rainfall anomalies are linked to the uptrend phase of the AMO [as reviewed by Latif *et al.* 2010], but such a linkage has not yet been substantiated. Landsea

(1993) has estimated that about 60% of tropical storms and moderate tropical cyclones in the Atlantic basin, and over 80% of intense tropical cyclones (*i.e.* Category 3 and above) originate from AEWs. Later studies (Thorncroft and Hodges 2001; Hopsch *et al.* 2007; Chen *et al.* 2008) found that southern-track AEWs contribute the most to tropical cyclogenesis, because their nature of moist convection facilitates the conversion of cold-core waves into warm-core tropical cyclones (Pytharoulis and Thorncroft 1999).

The Atlantic tropical cyclogeneses and those that originate from AEWs was examined by performing a back-tracking method of AEWs as was used in Chen *et al.* (2008). The back-tracking procedure begins with the genesis location of a tropical cyclone, then tracks any pre-existing perturbation associated with the tropical cyclogenesis *back* to its origin of perturbation by using daily-mean wind and vorticity fields at 925 mb and 600 mb in conjunction with daily OLR. Here, if the perturbation originated over the African continent, then its related tropical cyclogenesis is regarded as being of AEW origin. The tracking was performed manually and only those tropical cyclones during the period 2005-2010 were tracked, while cases prior to 2005 were adopted from Chen et al. (2008). Illustrated in Fig. 5 are tropical cyclogeneses with an AEW origin, the initial locations of those AEWs, and the trajectory between the two. Southern-track AEWs tend to form tropical cyclones further to the east and closer to West Africa than northern-track AEWs, likely due to stronger latent heat release in the West African monsoon region (Thorncroft and Hodges 2001).

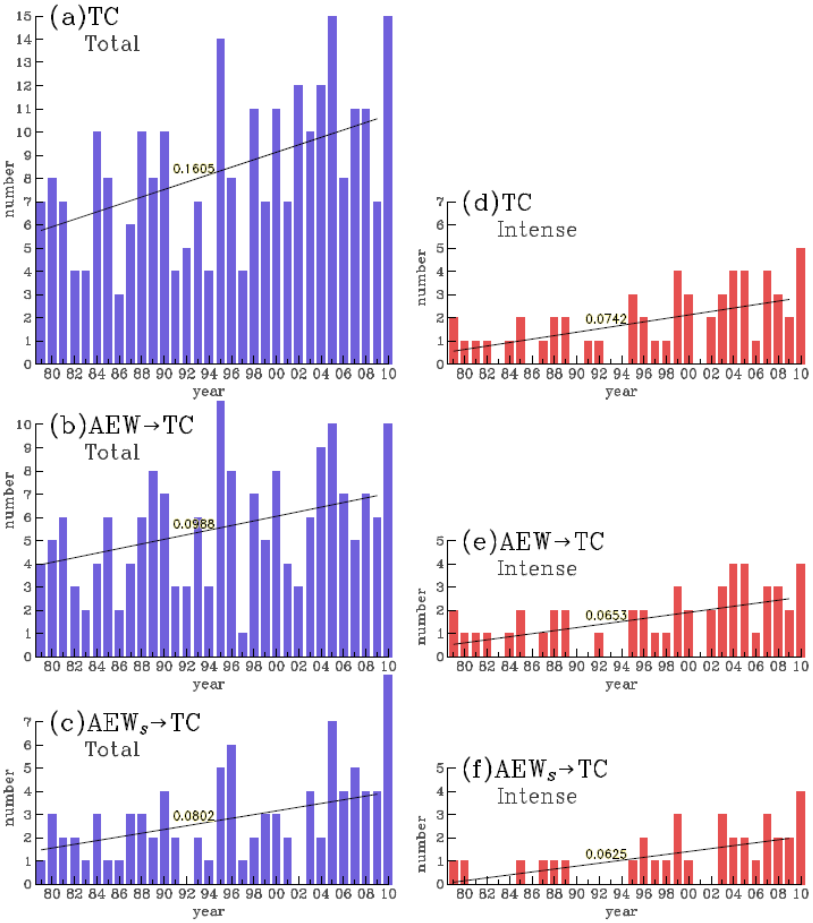


Fig. 6 Number of Atlantic tropical cyclones (TCs) during July-September, for (a) total TCs, (b) AEW-induced TCs, and (c) southern-track AEW-induced TCs. (d)-(f) Same as (a)-(c) but for intense TCs of Category 3 and above. The linear trends and slopes are given. All trends are significant at the 95% confidence interval per *t*-test.

Figure 6a shows the number of tropical cyclones initiated in the July-September season and their linear trend. Year 2010 was not included in the trend analysis because the track records at the time were provisional. Of the 86% *increase* in the number of tropical cyclones since 1979, 62% originated from AEWs (Fig. 6b) while 50% are linked to southern track AEWs (Fig. 6c). The connection between AEWs and intense tropical cyclones is more pronounced: Given a dramatic 148% *increase* in the number of intense tropical cyclones, 88% are of AEW origin (Fig. 6e), while 84% (out of the 88%) originated from the southern track of AEWs (Fig. 6f). It has been found that the large-scale atmospheric conditions leading to the increase in tropical cyclone frequency and power involve net surface radiation, thermodynamic efficiency, surface wind speed, and vertical wind shear (Emanuel 2007), whose variations turn to the favorable side for tropical cyclogenesis during the recent AMO uptrend (Hoerling *et al.* 2006; Latif *et al.* 2010). The results presented in Fig. 6 are a further substantiation of the synoptic condition that exists to "fuel" the observation of increasing rainfall in the Sahel and the Atlantic tropical cyclones (especially the intense ones).

5. Concluding remarks

Increased Sahel rainfall over the 1979-2010 period and associated synoptic conditions was found to result from the northward migration of the seasonal rainband. Convective activity associated with the migrating rainband also intensified, mostly along the northern boundary of the rainband and south of the AEJ. We also found a positional shift and intensification of the AEJ, consistent with the northward migration of the Sahel rainband. These features are accompanied by stationary tropical monsoon westerlies, an expanding TEJ, and an increase in moisture flux convergence (not shown). Furthermore, the poleward shift and amplification of AEW activity associated with the AEJ changes play a crucial role in the change of tropical cyclogenesis over the North Atlantic. Through a manual back-tracking method that connects tropical cyclones to their AEW origin, 88% of the dramatic (148%) increase in intense tropical cyclones were found to originate from AEWs. These results provide a synoptic linkage between the documented increase in tropical cyclone threat and the increasing Sahel rainfall, over the past 30 years. In view of the severe spreads in GCM projections of the Sahel climate, simulation quality in terms of the AEJ-AEW system may serve as an additional criterion for the assessment of projection uncertainties. Moreover, projected AEW activities may also provide an indication for future tropical cyclone threat.

References

Adler, R. F., and Coauthors, 2003: The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979,ÄìPresent). *Journal of Hydrometeorology*, **4,** 1147-1167.

Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones, 2006: Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. *J. Geophys. Res.*, **111,** D12106.

Burpee, R. W., 1972: The Origin and Structure of Easterly Waves in the Lower Troposphere of North Africa. *J. Atmos. Sci.*, **29,** 77-90.

Chang, C.-B., 1993: Impact of Desert Environment on the Genesis of African Wave Disturbances. *J. Atmos. Sci.*, **50,** 2137-2145.

Charney, J. G., and M. E. Stern, 1962: On the Stability of Internal Baroclinic Jets in a Rotating Atmosphere. *J. Atmos. Sci.*, **19,** 159-172.

Chen, T.-C., 2006: Characteristics of African Easterly Waves Depicted by ECMWF Reanalyses for 1991,Äì2000. *Mon. Wea. Rev.*, **134,** 3539-3566.

Chen, T.-C., S.-Y. Wang, and A. J. Clark, 2008: North Atlantic Hurricanes Contributed by African Easterly Waves North and South of the African Easterly Jet. *J. Climate*, **21,** 6767-6776.

Chen, T. C., and S.-Y. Wang, 2007: Interannual variation of the Sahel rainfall. Geophysical Research Abstracts, Vol. 9, 11206. *European Geosciences Union Symposium of Precipitation Science*.

Cook, K. H., 2008: Climate science: The mysteries of Sahel droughts. *Nature Geosci*, **1,** 647-648.

Dai, A., P. J. Lamb, K. E. Trenberth, M. Hulme, P. D. Jones, and P. Xie, 2004: The recent Sahel drought is real. *Inter. J. Climatol.*, **24,** 1323-1331.

Emanuel, K., 2005: Increasing destructiveness of tropical cyclones over the past 30 years. *Nature*, **436,** 686-688.

——, 2007: Environmental Factors Affecting Tropical Cyclone Power Dissipation. *J. Climate*, **20,** 5497-5509.

Fink, A. H., J. M. Schrage, and S. Kotthaus, 2010: On the Potential Causes of the Nonstationary Correlations between West African Precipitation and Atlantic Hurricane Activity. *J. Climate*, **23,** 5437-5456.

Giannini, A., R. Saravanan, and P. Chang, 2003: Oceanic Forcing of Sahel Rainfall on Interannual to Interdecadal Time Scales. *Science*, **302,** 1027-1030.

Grist, J. P., S. E. Nicholson, and A. I. Barcilon, 2002: Easterly Waves over Africa. Part II: Observed and Modeled Contrasts between Wet and Dry Years. *Mon. Wea. Rev.*, **130,** 212-225.

Hagos, S. M., and K. H. Cook, 2008: Ocean Warming and Late-Twentieth-Century Sahel Drought and Recovery. *J. Climate*, **21,** 3797-3814.

Herrmann, S. M., A. Anyamba, and C. J. Tucker, 2005: Recent trends in vegetation dynamics in the African Sahel and their relationship to climate. *Global Environmental Change Part A*, **15,** 394-404.

Hoerling, M., J. Hurrell, J. Eischeid, and A. Phillips, 2006: Detection and Attribution of Twentieth-Century Northern and Southern African Rainfall Change. *J. Climate*, **19,** 3989-4008.

Hopsch, S. B., C. D. Thorncroft, K. Hodges, and A. Aiyyer, 2007: West African Storm Tracks and Their Relationship to Atlantic Tropical Cyclones. *J. Climate*, **20,** 2468-2483.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77,** 437-471.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83,** 1631-1643.

Knight, J. R., R. J. Allan, C. K. Folland, M. Vellinga, and M. E. Mann, 2005: A signature of persistent natural thermohaline circulation cycles in observed climate. *Geophys. Res. Lett.*, **32,** L20708.

Landsea, C. W., 1993: A Climatology of Intense (or Major) Atlantic Hurricanes. *Mon. Wea. Rev.*, **121,** 1703-1713.

Landsea, C. W., and W. M. Gray, 1992: The Strong Association between Western Sahelian Monsoon Rainfall and Intense Atlantic Hurricanes. *J. Climate*, **5,** 435-453.

Landsea, C. W., N. C. C. Anderson, G. Clark, J. Dunion, J. Fernandez-Partagas, P. Hungerford, C., and a. M. Z. Neumann, 2004: The Atlantic hurricane database re-analysis project: Documentation for the 1851-1910 alterations and additions to the HURDAT database. *Hurricanes and Typhoons: Past, Present and Future*, R. J. M. a. K.-B. Liu, Ed., Columbia University Press, 177-221.

Latif, M., and Coauthors, 2010: Dynamics of decadal climate variability and implications for its prediction. *OceanObs'09: Sustained Ocean Observations and Information for Society*, Venice, Italy.

Legates, D. R., and C. J. Willmott, 1990: Mean seasonal and spatial variability in gauge-corrected, global precipitation. *Int. J. Climatol.*, **10,** 111-127.

Lu, J., and T. L. Delworth, 2005: Oceanic forcing of the late 20th century Sahel drought. *Geophys. Res. Lett.*, **32,** L22706.

Nicholson, S., and X. Yin, 2001: Rainfall Conditions in Equatorial East Africa during the Nineteenth Century as Inferred from the Record of Lake Victoria. *Climatic Change*, **48,** 387-398.

Olsson, L., L. Eklundh, and J. Ardˆ, 2005: A recent greening of the Sahel--trends, patterns and potential causes. *Journal of Arid Environments*, **63,** 556-566.

Pytharoulis, I., and C. Thorncroft, 1999: The Low-Level Structure of African Easterly Waves in 1995. *Mon. Wea. Rev.*, **127,** 2266-2280.

Rienecker, M. M., and Coauthors, 2011: MERRA - NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate,* in press.

Schimel, D. S., and Coauthors, 2001: Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems. *Nature*, **414,** 169-172.

Simmons, A. S., D. D. Uppala, and S. Kobayashi, 2007: ERA-interim: new ECMWF reanalysis products from 1989 onwards. *CMWF Newsl* **110,** 29-35.

Smith, T. M., and R. W. Reynolds, 2003: Extended Reconstruction of Global Sea Surface Temperatures Based on COADS Data (1854,Äì1997). *J. Climate*, **16,** 1495-1510.

Thorncroft, C., and K. Hodges, 2001: African Easterly Wave Variability and Its Relationship to Atlantic Tropical Cyclone Activity. *J. Climate*, **14,** 1166-1179.

Thorncroft, C. D., and B. J. Hoskins, 1994a: An idealized study of African easterly waves. I: A linear view. *Quarterly Journal of the Royal Meteorological Society*, **120,** 953-982.

——, 1994b: An idealized study of African easterly waves. II: A nonlinear view. *Quarterly Journal of the Royal Meteorological Society*, **120,** 983-1015.

Thorncroft, C. D., and D. P. Rowell, 1998: Interannual variability of African wave activity in a general circulation model. *Inter. J. Climatol.*, **18,** 1305-1323.

Uppala, S. M., and Coauthors, 2005: The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, **131,** 2961-3012.

Wang, S.-Y., and R. R. Gillies, 2011: Observed change in Sahel rainfall, circulations, African easterly waves, and Atlantic hurricanes since 1979. *International Journal of Geophysics***,** (in press) doi:10.1155/2011/259529.

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.

Zhang, C., 1993: Large-Scale Variability of Atmospheric Deep Convection in Relation to Sea Surface Temperature in the Tropics. *J. Climate*, **6,** 1898-1913.

6. APPLICATIONS OF CLIMATE INFORMATION AND IMPROVED CLIMATE SERVICES

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

An Emerging Protocol for Research-to-Operations (R2O) at CPC

E. O'Lenic[1], H. Hartmann[2], M. Ou[1], K. Pelman[1], S. Handel[1]

[1]*Climate Prediction Center, NOAA/NWS/NCEP, Camp Springs, MD*

[2]*Climate Assessment for the Southwest, The University of Arizona, Tucson, AZ*

1. Introduction

Over the last several years CPC has created products and adopted protocols, which are new to the organization. The net result is an emerging R2O protocol, which promises to radically improve our capability to implement new products, while continuing to maintain our existing ones.

We identify 5 critical elements: (1) Project Planning, (2) Software Version Control, (3) Issue-Tracking, (4) Wiki usage, and (5) Collaborative Software Development; and provide 2 example projects: (1) a Verification web tool (VWT), and (2) a Dynamic Probability of Exceedance web tool (dPOE), both of which are nearing completion.

Section 2 gives a comparison of research ("R") and operations ("O"). Section 3 describes the 5 elements of a new R2O. Section 4 describes how CPC's interactions with CLIMAS inspired development of two web-based tools to allow users to interact with databases of forecasts and observations. The Verification Web Tool and the Dynamic POE Web Tool are also described. Section 5 gives Lessons Learned.

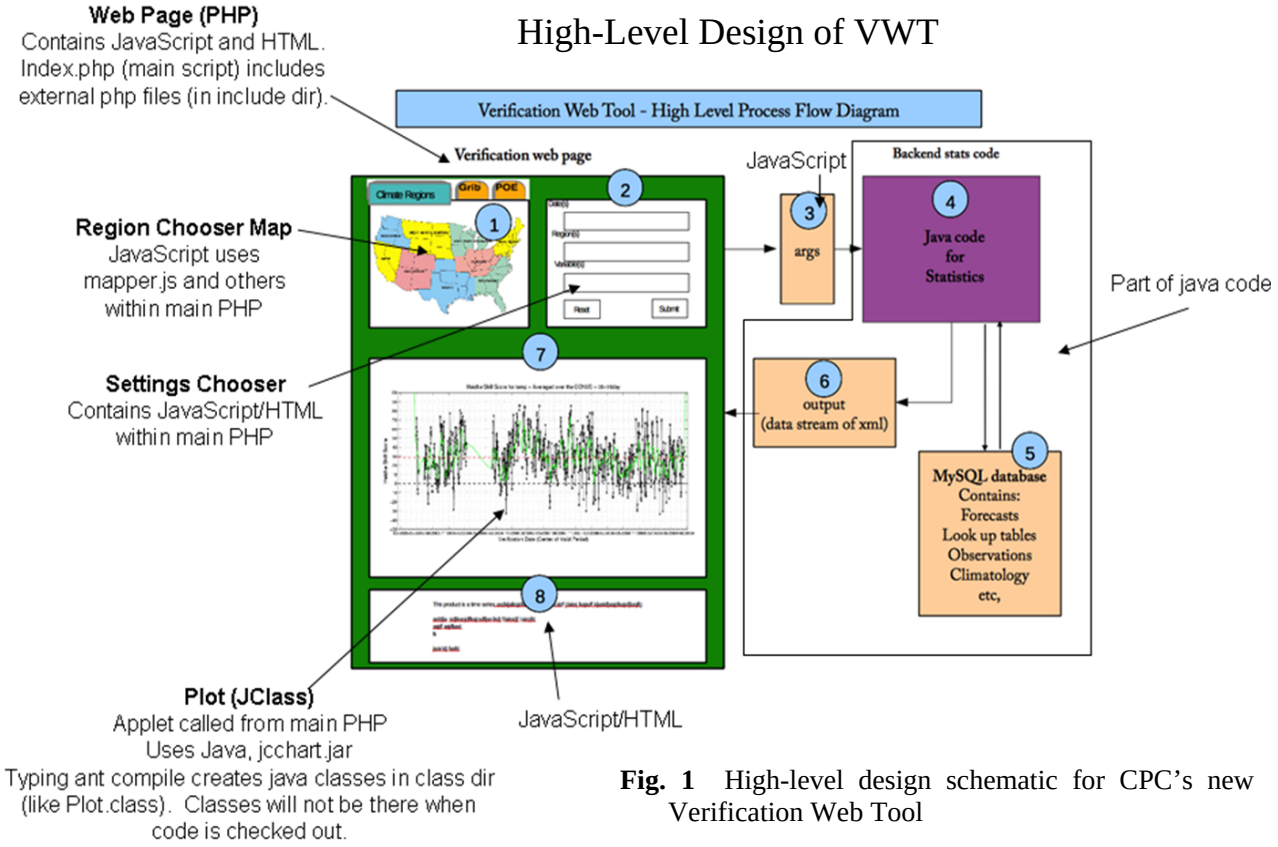


Fig. 1 High-level design schematic for CPC's new Verification Web Tool

Correspondence to: Ed O'Lenic, Climate Prediction Center, NOAA/NWS/NCEP, Camp Springs, MD 20746; E-mail: Ed.Olenic@noaa.gov.

2. Research versus Operations.

Research: Untested technologies, new knowledge of uncertain or distant future application valued

1. Funding supports entirely new topics
2. Products are papers, published episodically
3. Code standards, documentation often of secondary importance
4. Driven by funding and quest for new results
5. Research personnel is largest cost
6. Narrow, highly-trained user community

Operations: Robust technologies, continuity, practicality valued

1. Funding limits ability to do new things
2. Routine/rigid delivery of products
3. Code standards, documentation essential
4. Driven by system *security considerations*
5. Software maintenance is largest cost
6. Broad user base, often untrained

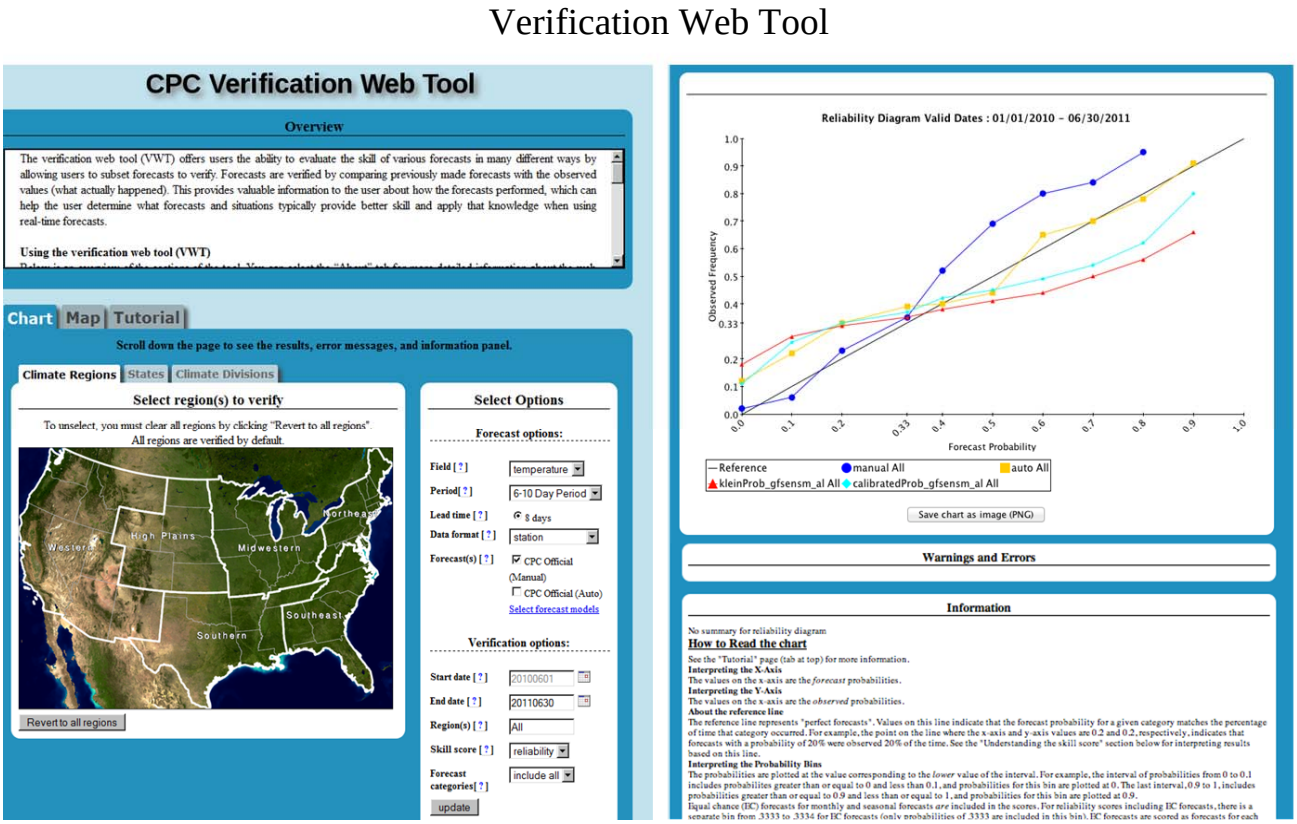


Fig. 2 The VWT is a Java-based interactive web tool which allows a user to request skill graphs or maps, for temperature and precipitation forecasts, for a variety of forecast periods, geographical regions, and historical periods.

3. The 5 essential R2O project elements

1. Collaborative Development: Leverages the talents and resources of developers inside and outside of the organization. This is facilitated by software which allows simultaneous development of code by multiple people at the same time, such as CVS or Subversion.

2. Project Planning : A formal process, starting with a Charter, which lists purpose, in and out of scope activities, costs & benefits, deliverables, resources, risks, participants.
3. Software Version Control: This software facilitates orderly code development and is essential to collaborative development. Facilitates collaborative development of code (see item 1.).
4. Issue-Tracking : Facilitates management of project through accountability. Trac issue-trackiing software is used at CPC.
5. Wiki-Usage: Enables shared documentation of information.

4. How the CLIMAS-CPC collaboration led to innovation at CPC

CLIMAS introduced issue-tracking, version control, and collaborative development to CPC through a series of lectures, culminating in a project to collaboratively develop a web tool. In the process of working together, CPC and CLIMAS successfully developed and implemented web services, dynamic process initiation, and user interfaces at CPC. They also overcame security issues involving non-NWS access to NWS computers.

As a direct result of this work, CPC embarked on an effort to internally develop the Verification Web Tool. CPC and CLIMAS also continued their joint collaboration by developing the dynamic Probability of Exeedance (dPOE) web tool. These are illustrated by Figures 1-3.

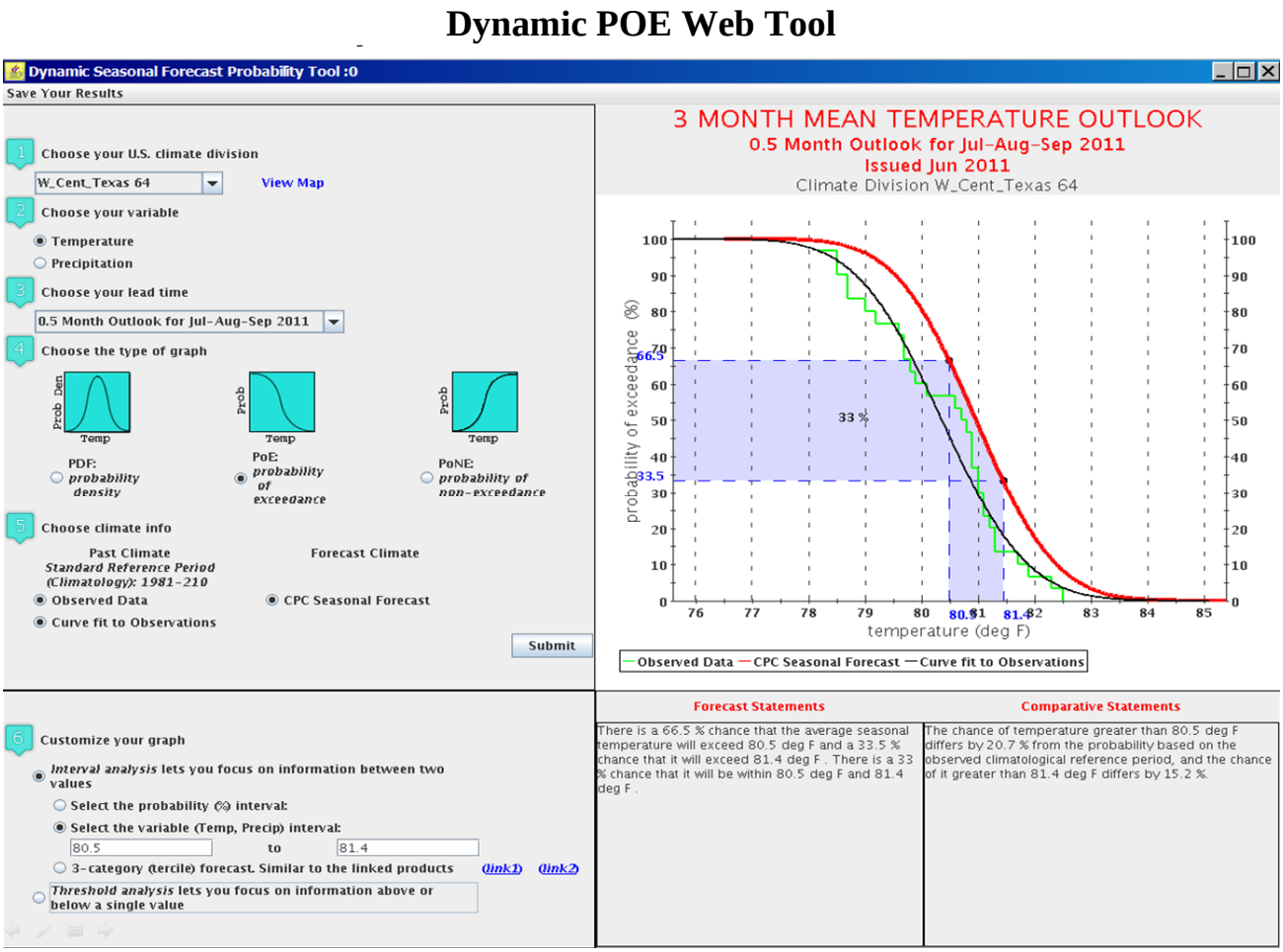


Fig. 3 The dPOE is a Java-based interactive web tool which allows a user to request graphs which place a temperature or precipitation forecast into the context of a reference distribution. These distributions may be viewed in 3 different formats controlled by the user, who may also vary the thresholds which bracket the range of possible values of a forecast.

5. Lessons learned

1. R2O is hard, partly due to natural differences in culture and communication between the research and operations side of an organization.
2. CLIMAS introduced issue-tracking, version control, and collaborative development to CPC. These powerful tools provide accountability and leverage the talents of coders inside and outside of the organization in a way never-before attempted at CPC.
3. These tools are being increasingly incorporated into CPC's Operations Branch, and have permanently changed our mode of day-to-day operations.
4. VWT and Dynamic POE were successfully developed using these, along with wiki-usage and project planning (5 elements).
5. CPC/CLIMAS Successes: web services, dynamic process initiation, user interfaces, security issues with non-NWS access to NWS computers, specific software development tools, dPOE.
6. Other parts of NCEP are adopting these software development techniques (NCO, EMC).

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Seasonal Forecasting Using the Climate Predictability Tool (CPT)

S. J. Mason

International Research Institute for Climate and Society, The Earth Institute of Columbia University

The CPT is a software package developed by the International Research Institute for Climate and Society (IRI) designed for making seasonal climate forecasts. The CPT is an easy-to-use tool that runs on Windows 95+, and a source-code version that can be compiled under other types of operating systems is available for running batch jobs. It is available free of charge from the IRI's web page: http://iri.columbia.edu/climate/tools/cpt/. At the time of writing, the latest version of CPT (11.10) consists of about 50,000 lines of Fortran 95 code, and uses compiler extensions for the user-interface and graphics, and so is not portable to other platforms. The source-code version consists of about 33,000 lines of Fortran 90 code, and provides no graphics functionality. This version has been tested with gfortran, ifort, nagfor, pgf90, and pgf95, but up-to-date versions of the compilers are required to avoid issues with compiler bugs.

The underlying goal in developing the CPT has been to promote the widespread creation and communication of quality-controlled seasonal climate forecasts that address specific needs of different user groups. The software was initially developed to enable forecasters at National Meteorological Services (NMSs) in Africa to produce updated seasonal forecasts for their country, and to provide greater consistency in inputs to the Regional Climate Outlook Forums (RCOFs) to facilitate consensus building, but the CPT has been used widely beyond the RCOFs (Fig. 1).

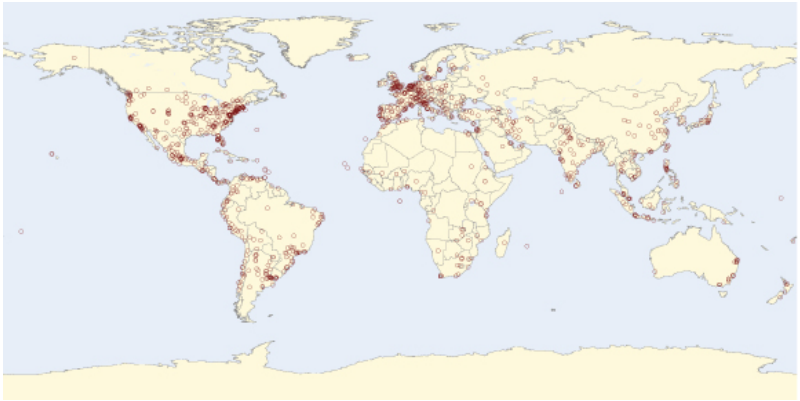


Fig. 1 Locations of downloads of the CPT software, 2004 – 2011.

There are two main approaches to generating seasonal forecasts: using large-scale models of the global atmosphere, known as general circulation models (GCMs), or using a statistical approach to relate seasonal climate to changes in sea-surface temperatures, such as those associated with El Niño, or to other predictors. In the former case, predictions are made for large-areas, and are often not very relevant for specific locations. In addition, because of the coarse scale at which the GCMs operate, the geography in the models is often distorted, and so geographical locations can be displaced. These GCM outputs therefore need to be adjusted so that they can be applied at the local level. The CPT tool is designed to perform both forms of prediction, namely downscaling of GCM output, and purely statistical predictions.

The statistical approach to making seasonal forecasts from sea-surface temperatures has been used for a number of years at many National Meteorological Services. Since the late 1990s, these statistical forecasts have been combined to produce a consensus forecast, representing a patchwork of nationally-based forecasts for subcontinental areas, in Regional Climate Outlook Forums (RCOFs; Ogallo *et al.* 2008). While such forums have been very successful in building the capacity to produce seasonal climate forecasts, a number of problems have emerged, and some systematic errors in the forecasts have been identified (Mason and Chidzambwa 2009). The CPT tool was designed in response to these problems, and, specifically, was developed to address the following issues:

Correspondence to: S. J. Mason, International Research Institute for Climate and Society, The Earth Institute of Columbia University, Palisades, NY; E-mail: simon@iri.columbia.edu.

a. *Slow production time*: the time taken to construct the statistical forecasts at the RCOFs was requiring long and expensive pre-forum workshops. By using CPT, forecasts can be produced in just a few hours or less. The quick production time makes it possible to hold shorter, and thus cheaper, pre-forum workshops, and provides more time at these workshops for advanced training. It also makes it viable for forecasters to produce updated forecasts on a monthly basis back in their home countries.

b. *Artificial skill*: an important step in producing a reliable seasonal forecast is to obtain a realistic estimate of how good the model predictions are. CPT performs rigorous cross-validation and retroactive tests for estimating skill levels, and adjusts the forecasts accordingly. Extensive diagnostic statistics are provided, including most of the scores and procedures recommended by the WMO CBS SVSLRF (including the calculation of significance levels and error bars).

c. *Dependence on one model*: it has been demonstrated extensively that the best seasonal forecasts are produced by generating a number of predictions and then combining these, perhaps by simply taking the average of the predictions. In the past, however, there has been a strong reliance on the prediction from a single model, simply because the effort invested in constructing the statistical model tended to encourage an over-confidence in the prediction from this model. Inputs from other sources (most notably the GCM predictions) were largely down-played because of the lack of a sense of ownership of these additional products. By making it much easier to generate a set of predictions by using CPT, the official forecasts from the National Meteorological Services and from the RCOFs now consider a broader range of inputs than was possible in the early years of seasonal climate forecasting efforts.

d. *Unreliable probabilities*: seasonal climate forecasts are expressed probabilistically because of the large uncertainties involved in forecasting the next few months. However, most statistical forecasting methodologies do not explicitly indicate the uncertainty in the prediction, and so this uncertainty has to be estimated. Some simple and intuitive ways of estimating forecast uncertainty are unfortunately not very reliable (Mason and Mimmack 2002). A more reliable system has been implemented in CPT based upon the error variance of the cross-validated forecasts over the training period.

e. *Forecast format*: seasonal forecasts are typically presented as the probability that the seasonal rainfall total, for example, will fall within pre-defined ranges. These ranges are most commonly set as the upper and lower terciles of the historical rainfall totals. This format tends to be unpopular with users of the forecasts, partly because the forecast is too abstract (it is not clear how much rainfall is meant by the upper and lower terciles), and partly because the forecast is too unspecific (the upper and lower terciles are often not very interesting thresholds). Within CPT there are options that provide considerable flexibility to tailor forecasts for specific user requirements, including options to redefine the ranges. Forecasts can be produced in a variety of formats, and detailed information is provided so that the forecast can be communicated to the end users in easy-to-understand terms.

Although the software is designed to implement best-practices, a second underlying principle has been to respond to user requests for improved functionality and ease of use. Major enhancements to the CPT have been released approximately once per year. The major new versions are indicated in Table 1.

The software has been introduced to most of the RCOFs and is now used fairly extensively: there are currently over 650 registered users of the Windows versions from all over the globe. Visitors to the CPC Africa Desk are provided extensive training in the software. Numerous training workshops in the use of the software have already been held, including under the auspices of the WMO CLIPS activities.

References

Mason, S.J., and S. Chidzambwa, 2009: Verification of RCOF Forecasts. *WMO RCOF Review 2008 Position Paper*, 26 pp.

Mason, S.J., and G.M. Mimmack, 2002: Comparison of some statistical methods of probabilistic forecasting of ENSO. *Journal of Climate*, **15**, 8–29.

Ogallo L.J., P. Bessemoulin, J.P. Ceron, S.J. Mason, and S.J. Connor, 2008: Adapting to climate variability and change: the Climate Outlook Forum process. *J. World Meteor. Org.*, **57**, 93–102.

CPT Versions		
Version	**Date**	**New features**
CPT 0		MATLAB code for performing canonical correlation analysis
CPT 1	Dec 2002	Translation into about 650 lines of Fortran 77 as interface to LAPACK SVD routines (requires recompilation each time) Principal component regression
CPT 2	Aug 2003	Conversion to Fortran 95 Graphical user interface Validation statistics Option to calculate a forecast using updated predictors
CPT 3	Feb 2004	Mapping of station data Missing value estimation
CPT 4	Feb 2005	Improved graphics Bootstrap confidence intervals and p-values
CPT 5	Aug 2005	Forecast uncertainty estimates based on prediction intervals WMO SVSLRF verification procedures Optional transformation of predictand to Gaussian distribution Options to define categories using climatological probabilities other than 33%, or using absolute values
CPT 6	Nov 2005	Multiple concurrent users Exceedance probability curves Options to calculate forecasts as, and to define categories in terms of, anomalies or standardized anomalies
CPT 7	Aug 2006	Option to set model update interval in retroactive procedure Option to set a zero-bound (useful for precipitation forecasts) Option to calculate forecasts as, and to define categories in terms of, % of average
CPT 8	May 2007	Retroactive forecast probabilities and verification including attributes diagrams Multiple linear regression option Permitted input of daily data
CPT 9	Mar 2008	Break up of executable to reduce executable size by using DLLs to simplify download of updates Major internal restructuring of code for closer parallel between Windows and Linux versions Additional verification procedures
CPT 10	Oct 2009	Multiple predictor fields, including extended EOFs New user-interface “Ensemble” forecasts based on error variance Additional verification procedures New input data formats for improved identification of forecast lags Forecasts as odds Option to define categories in terms of analogue years Calculation of p-values for skill maps Improved graphics functions
CPT 11	Jan 2011	Multi-lingual user-interface GCM verification option Probabilistic verification scores for retroactive forecasts Category thresholds, averages, and correlation maps Improved missing value estimation Improved handling of daily data Prevention of over-fitting of CCA models Decreased sensitivity to different compilers

Table 1 History of CPT versions indicating the major new features.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Wildland Fire Climate Needs Roundtable

Peter B. Roohr

Office of Science and Technology, NOAA's National Weather Service

In March 2011, Dr. Jack Hayes, NOAA Assistant Administrator for Weather Services and Director of the NWS, met with senior level representatives of the National Wildfire Coordinating Group to discuss NWCG requirements for weather and climate information that would improve firefighting operations and planning. The number one requirement brought up by the NWCG was climate information that could help them arrange resources and allocate funding effectively around the country in preparation for major wildfires. This presentation brings up the overall need for climate information for firefighting and the specific NWCG needs.

The overall need for information on projected long-term weather patterns is addressed in many high level documents, to include the 2008 NOAA Science Advisory Board "Fire Weather Research: A Burning Agenda for NOAA" report and the 2007 Office of the Federal Coordinator for Meteorology "National Wildland Fire Weather Needs Assessment". The requirements brought forth by the NWCG in 2011 validate and specify the needs from these documents.

The NWCG uses mid-range fore weather information for pre-positioning of initial attack resources (2-3 days in advance of an event) and movement of resources (typically 1-2 days to mobilize). They use long-range fire climate information to determine the start date of seasonal crews (which can only be used for 90 days), determine start date of air resources under contract, and to ascertain budget planning for fuel treatments. The NWCG specifically brought up needs for seasonal prediction of: (1) mean upper level flow patterns (500 mb level), (2) frequency of precipitation, (3) wind events, (4) relative humidity trends, and (5) predictions out to 120 days (at weekly or biweekly time-steps to show trends that could impact planning and operartions). The NWCG also brought up need for climate information out to 2-3 years to help determine basis for fuel treatment and land management objectives, since planning that is impacted by climate change ties directly to fiscal year budget planning, and ensure firefighters know how climate variability and complexity interacts with fires; future potential impacts of climate on a decadal scale are also desired.

After the presentation the conference attendees discussed climate products and information that could address firefighting needs, especially those that could increase the ability of firefighting agencies and groups, like the NWCG, to assess the range of outcomes and narrow down range of plans.

Correspondence to: Peter B. Roohr, Office of Science and Technology, NOAA's National Weather Service, Silver Spring, MD 20910; E-mail: Peter.Roohr@noaa.gov.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Scientific Prospects for Weather to Climate Prediction and Services

Jiayu Zhou, S&TI Climate Mission, Office of Science and Technology
Wayne Higgins and Jin Huang, Climate Prediction Center/NCEP
Fiona Horsfall, Climate Service Division/OCWWS

NOAA's National Weather Service

1. Introduction

The National Weather Service (NWS) Mission for Science & Technology Infusion is an agency-wide collaborative effort among planning, operations and services, with support by strategic community partners. A prominent focus is to enhance the connection between weather and climate for the advancement and provision of integrated services. This is a summary based on presentations for NOAA CTB Joint Seminar Series, Joint Center for Satellite Data Assimilation Workshop, Climate Prediction Application Science Workshop, Ensemble User Workshop, NCEP/EMC Seminars and COLA Seminars in 2010-2011. It highlights recent progresses and challenges in three key development areas: 1) unified modeling, 2) seamless prediction, and 3) integrated services.

2. Unified modeling

Modeling of weather and climate fluctuations and their interactions with the Earth system are integral to the simulation/prediction problem. A unified modeling approach can be used to address common processes in both classes of models, such that the progress in short-range weather forecasts will translate into improvements in long-range climate predictions and vice-versa.

NASA/Goddard multi-scale modeling system (MMS) with unified physics

The Goddard MMS shown in Figure 1 consists of the Goddard Cumulus Ensemble Model (GCE, a cloud-resolving model), the NASA unified Weather Research and Forecasting Model (WRF, a regional-scale model), and the coupled fvGCM-GCE (the GCE coupled to a general circulation model). The same cloud microphysical processes, long- and short-wave radiative transfer and land-surface processes are applied in all of the models to study explicit cloud-radiation and cloud-surface interactive processes and identify the optimal grid size and physical process. By incorporating the physical packages originally developed for high-resolution process model into both NWP and GCM, it demonstrated significant reduction of model biases and more realistic simulation of weather and climate, such as diurnal variation of precipitation systems, Typhoon, etc.

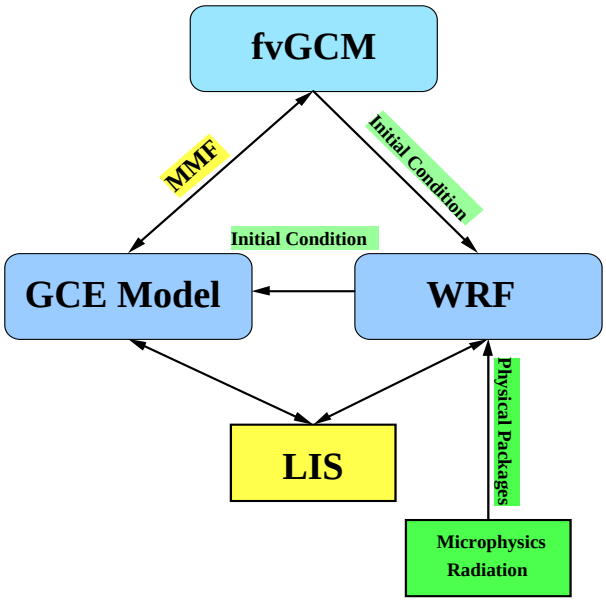


Fig. 1 Schematic diagram of the Goddard Multi-scale Modeling System with unified physics. (Tao 2010)

UMD/ESSIC CFS-CWRF nested system with optimized physics

The Climate extension of the WRF model (CWRF) developed in the Earth System Science Interdisciplinary Center (ESSIC), the University of Maryland inherits all WRF functionalities for numerical

Correspondence to: Jiayu Zhou, Office of Science and Technology, NOAA's National Weather Service, Silver Spring, MD 20910; E-mail: jiayu.zhou@noaa.gov.

weather prediction while enhancing the capability to predict climate. The CWRF incorporates a grand set of physics schemes, which contain more than 10^{24} of alternative configurations, representing interactions among surface, planetary boundary layer, cloud, aerosol, and radiation. An optimized physics (geographically dependent) ensemble approach is applied to improve weather forecasts and climate prediction along with reliable uncertainty estimates.

The CWRF model has been nested in the National Centers for Environmental Prediction (NCEP) operational Climate Forecast System (CFS). The Figure 2 shows significant improvement of cold season precipitation prediction over the United States by downscaling.

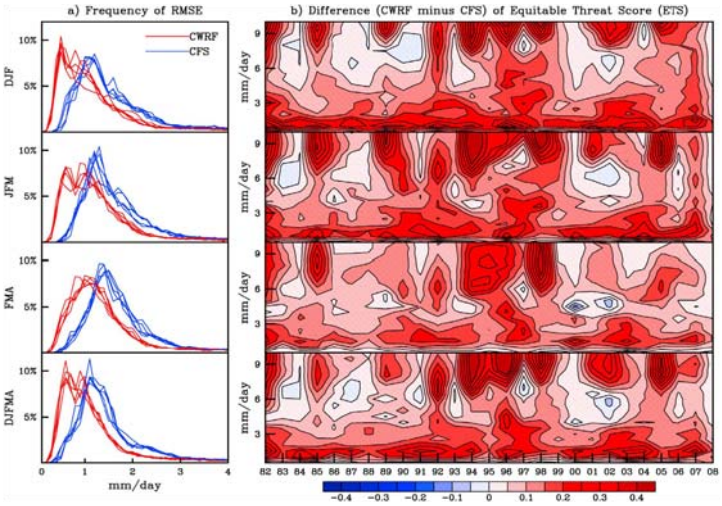


Fig. 2 (a) Frequency distributions of root mean square errors (mm/day) predicted by the CFS and downscaled by the CWRF and (b) CWRF minus CFS differences in the equitable threat score for seasonal mean precipitation interannual variations. The statistics are based on all land grids over the entire inner domain for DJF, JFM, FMA, and DJFMA from the 5 realizations during 1982–2008. (Liang 2011)

3. Seamless prediction

The seamless prediction concept emphasizes the importance of scale interconnectivities and puts the stress on the weakest link of the prediction chain. In forecast practice, reducing biases and better representing uncertainties are the common foci of both weather and climate predictions for improvement. Enhanced cooperation and exchange of experiences between the two communities would accelerate progress for both.

Progress in satellite data assimilation - Land surface data improvement

The amount of satellite data assimilated over land in the Gridpoint Statistical Interpolation (GSI) was found to be far less than over ocean. Figure 3 shows a reduction of errors in simulated brightness temperature and an increase in the number of observations (NOAA-18 AMSU-A) assimilated in GSI, especially over forest regions, by using new roughness lengths formulations in the Global Forecast System (GFS) and updated microwave land emissivity model in the Community Radiative Transfer Model (CRTM).

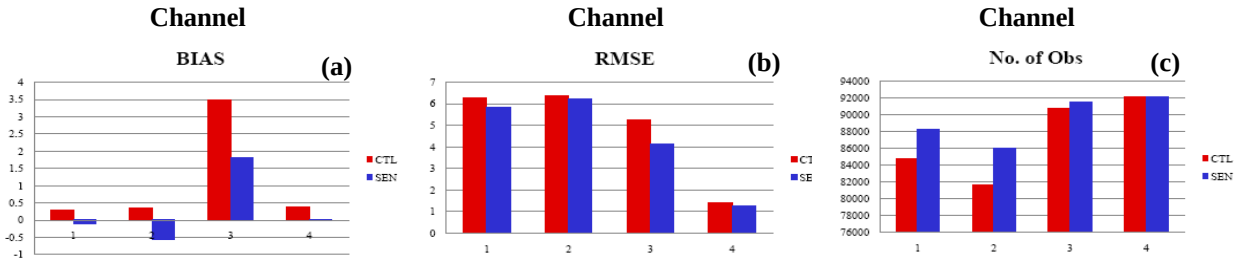


Fig. 3 Comparisons of brightness temperature (Tb) over land at 12Z assimilated by GSI averaged over the period of 1-31 July 2010, (a) bias, (b) RMSE and (c) the number of observations assimilated. (Zheng 2011)

Challenges of decadal simulation for CFS

Based on the analysis of NCEP Climate Forecast System (CFS v2) decadal runs for CMIP5, Figure 4 reveals that the simulation of Atlantic Meridional Overturning Circulation (AMOC) is weak and its 10-year trend is larger than that in the GODAS assimilation. Since the AMOC is related to the Atlantic Multidecadal

Oscillation, which can modulate ENSO activity (Te Raa *et al.* 2009), it has become a focus in CFS development to improve ENSO prediction.

Shared focus areas with weather ensemble forecast development community

How does weather forecast handle uncertainty? Why seasonal and decadal forecasting is possible, despite the limits to atmospheric predictability suggested by Lorenz? How should uncertainty be addressed in climate prediction and projection? These are issues fundamental for forecasters to understand to make reliable prediction with confidence. Here are some common focus areas learned from the 5th Ensemble User Workshop.

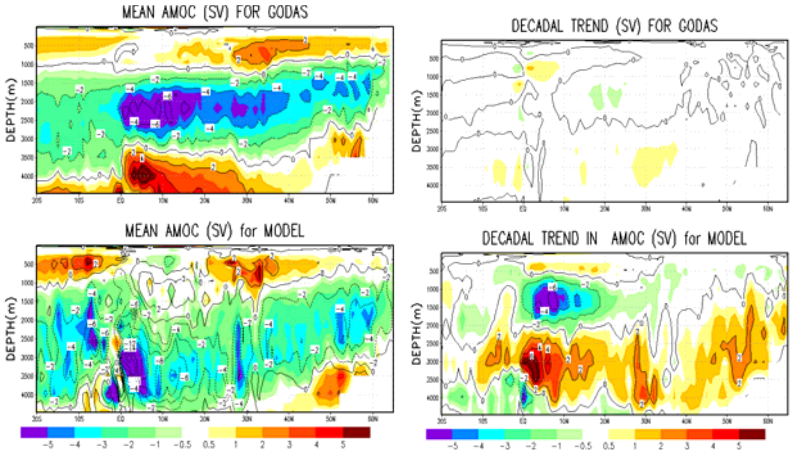


Fig. 4 Comparisons of NCEP CFSv2 decadal runs for CMIP5 (bottom) with GODAS (top). The left shows the mean AMOC; and the right: the decadal trend. (Nadiga and Wang 2011)

1. Ensemble configuration and ensemble forecast
   - Ensemble initializations
   - Multi-model, multi-physics and stochastic physics for ensemble perturbation
   - True coupled initialization
2. Statistical post-processing
   - Lead-time dependent systematic errors
   - Non-stationarity of real climate and changes of model-error statistics
   - Relating model variables to sensible weather
   - Choice of proxy of truth - error variance and ensemble of analyses to represent uncertainty
3. Issues to consider
   - Forecast of joint probability, *e.g.* "What is joint probability of heavy precipitation and strong winds?"
   - Better ways to make members (representativeness vs. ensemble size and resolution), not only for simple statistics but also for higher order statistical moments
   - Resolution change in mid-run
   - Resource allocation for real-time fore- vs. hind-casts

4. Integrated services

Regional services need reliable climate predictions integrated with weather and water information as long term adaptation is inseparable from near term decisions. Figure 5 shows that short-term variability becomes more significant compared to the long-term trends at smaller spatial scales. This information is critical when attributing anthropogenic causes to local climate variations.

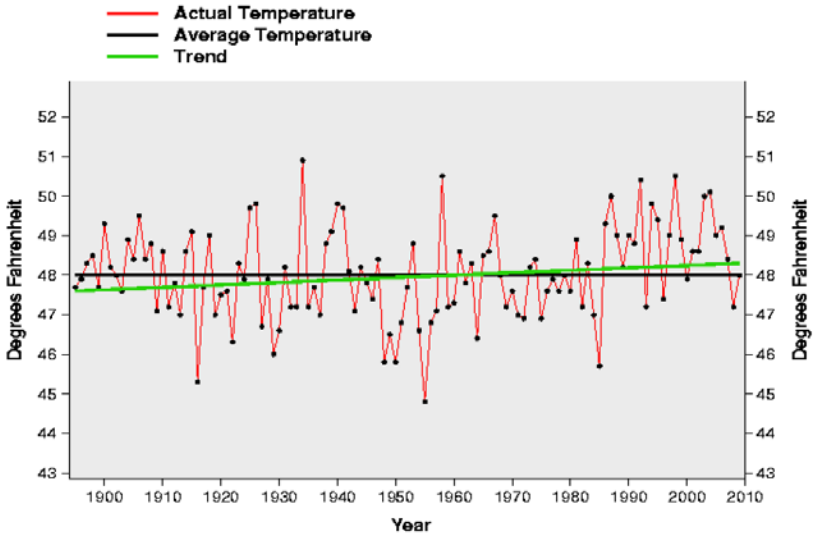


Fig. 5 Washington state temperature trend and annual variation (1895-2009). (Sarachik 2011)

Water supply forecasting for the Colorado

To meet stakeholder demands for climate information, the Colorado Basin River Forecast Center (CBRFC) supported by collaborative research with academic partners has launched a seasonal to year-two streamflow forecast intercomparison effort that has the most impact on Colorado River management (Werner 2011). CBRFC has also implemented an ensemble forecast technique developed at the NWS Office of Hydrologic Development to create a streamflow forecast ensemble based on GFS and CFS output. It is a centerpiece of the nascent NWS Hydrologic Ensemble Forecast Service.

Climate Change Impact Assessments for International Market Systems (CLIMARK) project

With broader spatial perspective and greater incorporation of temporal dynamics, new researches in climate impact assessment employed complex integrated models, which contained system components and included feedbacks. A conceptual framework for a dynamic and statistical hybrid modeling system was developed by the CLIMARK project (Winkler 2011) to make industry-wide assessments for market systems with multiple production regions. Its application to assess the impact of climate change on the tart cherry international market demonstrated impressive improvement.

Questions and concerns from local users

Local users are asking:

1. What reliable decision-support tools could be designed and deployed at current skill levels and uncertainty of climate predictions?
2. How well could climate projections predict local trends for the next 10 – 20 years? How could best practices be developed for use of projections in decision making and preparation for change?

Our regional climate service advancement is awaiting significant research progresses.

Acknowledgements. We thank Wei-Kuo Tao of NASA/GSFC, Xin-Zhong Liang of UMD/ESSIC, Weizhong Zheng, Nadiga and Jiande Wang of NWS/NCEP/EMC, Edward S. Sarachik of University of Washington, Julie Winkler of Michigan State University, and Kevin Werner of NWS/CBRFC for their contributions.

References

Liang, X.-Z., 2011: Regional Climate-Weather Research and Forecasting (CWRF) Model development & application. *NCEP Environmental Modeling Center Seminar*, Camp Springs, 17 May 2011

Nadiga, S. and J. Wang, 2011: Ocean data analyzed from the NCEP CFSv2 decadal runs for CMIP5. 13 July 2011.

Sarachik, E., 2011: The science in adaptation. *Center for Ocean-Land-Atmosphere Studies Seminar*, Calverton, Maryland, 2 May 2011

Tao, W.-K., 2010: The Goddard multi-scale modeling system with unified physics. Extended summary, *NOAA Climate Test Bed Joint Seminar Series*, Camp Springs, Maryland, 16 November 2010.

Te Raa, L. A., G. J. van Oldenborgh, H. A. Dijkstra1, and S. Y. Philip, 2009: Frequency-dependent effects of the Atlantic meridional overturning on the tropical Pacific Ocean. *Ocean Sci. Discuss.*, **6**, 477–490, 2009.

Werner, K., 2001: NOAA's Colorado Basin River Forecast Center: "Climate services on the Colorado River: capabilities, gaps, and chasms". Extended summary, *NOAA Climate Test Bed Joint Seminar Series*, Camp Springs, Maryland, 12 August 2011.

Winkler, J., 2011: Climate change impact assessments for international market systems (CLIMARK) - Moving from the local/regional to the global. *9th NOAA Climate Prediction Applications Science Workshop*, Des Moines, Iowa, 1-4 March 2011.

Zheng, W., and Coauthors, 2011: Improvement of microwave land emissivity calculation and its impact on satellite data assimilation. *9th Joint Center for Satellite Data Assimilation Workshop on Satellite Data Assimilation*, College Park, Maryland, 24-25 May 2011.

7. CLIMATE TEST BED PI REVIEWS

– *MME and Improving ISI Forecasts and Products*

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Recalibrating and Combining Ensemble Predictions

Michael K. Tippett, Tony Barnston, Lisa Goddard, Simon Mason
International Research Institute for Climate & Society (IRI), Columbia University
Malaquias Pena Mendez
SAIC-Environmental Modeling Center, NOAA/NWS/NCEP
Huug van den Dool
Climate Prediction Center, NOAA/NWS/NCEP

1. Introduction

The "model output statistics" (MOS) approach has long been used in forecasting to correct systematic errors of numerical models and to predict quantities not included in the model (Glahn and Lowry 1972). The MOS procedure is based on capturing the statistical relation between model outputs and observations and, in its simplest form, consists of a linear regression between these quantities. In theory, this procedure optimally calibrates the model forecast and provides reliable forecasts.

In practice, the regression parameters must be estimated from data. In seasonal forecasting, forecast histories are short, and skill is modest. Both factors lead to substantial sampling errors in the estimates. This work examines two problems where sampling error affects the reliability of regression-calibrated forecasts and provides solutions based on two "penalized" methods: ridge regression and lasso regression (Hoerl and Kennard 1988; Tibshirani 1996). The first problem comes from the observation that, even in a bivariate setting, ordinary least squares estimates lead to unreliable forecasts. The second problem arises in the context of multivariate MOS and is that common methods of predictor selection lead to negative skill and unreliable forecasts.

2. Are regression forecasts reliable?

The task of a forecaster is to make the best estimate of a future observation given available model output. In the case of probabilistic forecasts, the uncertainty of the estimate is also needed. Modeling the forecast f and its verifying observation o as random variables, the goal of the forecaster is to obtain the conditional distribution $p(o|f)$ which is defined to be the probability distribution of the verifying observation o given that the forecast f is known to have a particular value (DelSole 2005). The mean of the conditional distribution is the "best" estimate in the sense that it minimizes the expected squared error. Uncertainty information such as the forecast variance can

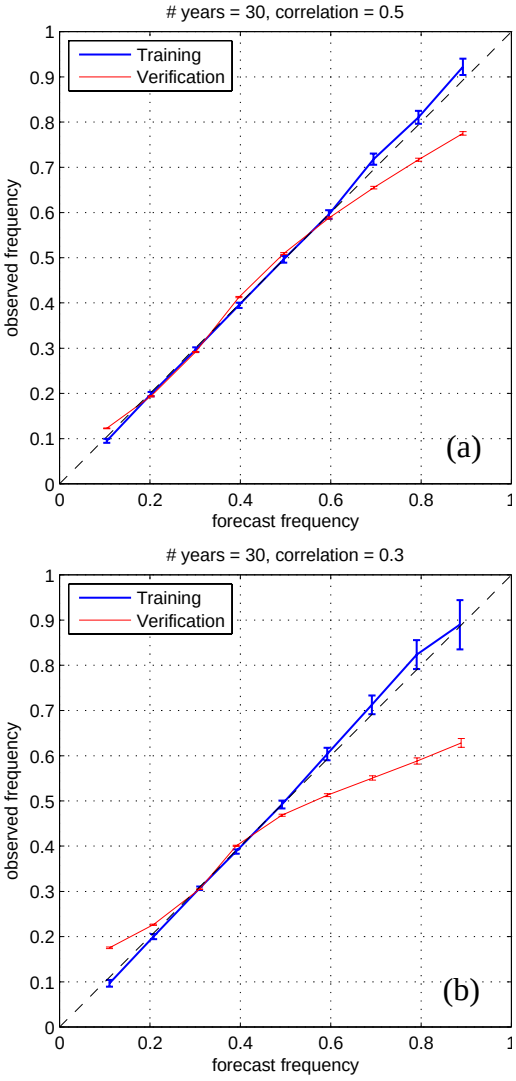


Fig. 1 Reliability diagrams for regression forecasts in (a) moderate skill (r=0.5) and (b) low skill (r=0.3) setting. The blue line is the in-sample reliability and the red line is the out-of-sample reliability.

Correspondence to: Michael K. Tippett, International Research Institute for Climate & Society (IRI), Columbia University, Palisades, NY; E-mail: tippett@iri.columbia.edu.

be computed from the conditional distribution and used to make probabilistic forecasts.

The challenge is to obtain the conditional distribution, which, in general, requires a complete description of the statistical relation between the forecast and the verifying observations. However, when forecast and observations have a joint Gaussian distribution, this only requires knowing the means and variances of *o* and *f*, and their correlation. In this case, the conditional distribution is itself Gaussian, and moreover, the conditional mean is simply given by linear regression; the conditional variance is the error variance of the regression. Such regression forecasts are known to be reliable when the regression parameters are known (Johnson and Bowler 2009). In practice, regression parameters must be estimated from data, and in the case of seasonal climate forecasts, from fairly short records. Therefore, an important issue is the impact of sampling error on the quality of forecasts, and in particular, on their reliability.

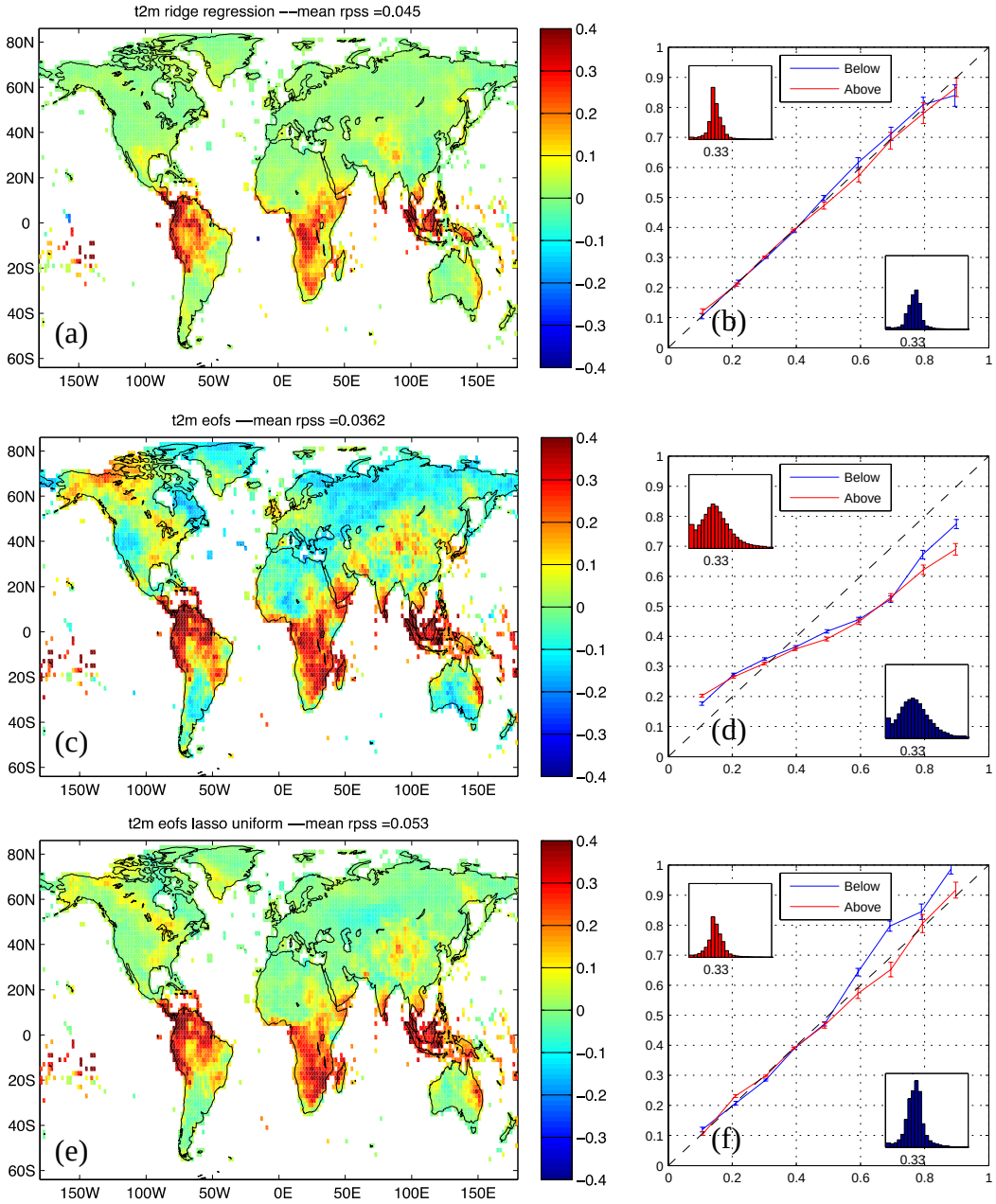


Fig. 2. RPSS and reliability diagrams for DJF t_{2m} forecasts based on gridpoint ridge regression (a,b), EOF predictors (c,d) and lasso applied to EOF predictors.

Monte Carlo experiments with bivariate synthetic Gaussian data show (Fig. 1) that regression forecasts are reliable with respect to the data used to estimate the regression parameters (in-sample). However, when regression forecasts are made for independent data (out-of-sample), they display systematic unreliability in the form of overconfidence. This problem is worse when the training sample is small and the underlying skill is low, both factors contributing to sampling error in the regression coefficient. Additional analysis of the synthetic data, as well as analytical expressions, indicate that the overconfidence is due to too strong signals rather than too small forecast error variance. Shrinkage methods like ridge and lasso reduce signal variance and improve reliability.

3. Selecting predictor patterns

A common multivariate regression forecast approach is to use spatial patterns as predictors. To illustrate our findings we use two-tier hindcasts of Dec-February (DJF) 2-meter temperature using the ECHAM4.5 atmospheric GCM with 24 ensemble members forced by Constructed Analogue (CA) forecast SST. The CA SST forecasts use data through the end of October. We use the period 1958-2001 and University of East Anglia observation data sets. Applying ridge regression on a gridpoint basis give reliable forecasts with skill in some regions as measured by the ranked probability skill score (RPSS; Figs. 2a,b). Using correlation-based EOFs of model output as predictors, results in increased RPSS in many regions (Fig. 2c), especially in the tropics, but is accompanied by negative RPSS in some regions, especially ones where the gridpoint MOS showed little or no skill. The average RPSS is 0.036, and the reliability diagram shows overconfidence (Fig. 2d). As in the bivariate case, this is due to excessive forecast signal, especially in areas where a climatological forecast of equal odds would be more appropriate. One way to proceed is to cast the problem as model-selection one, where one must choose between a pattern-based regression forecast or a climatological forecast. Methods like cross-validation and Akaike Information Criteria (AIC) can be used to select the model. This model selection approach offers some improvement, but does not entirely eliminate areas of negative skill and forecast overconfidence (not shown). Lasso regression is similar to ridge regression but more aggressively eliminates poor predictors. Lasso regression retains much of the skill improvements of EOF regression while not introducing negative skill, and improving reliability (Figs. 2e,f); average RPSS is 0.053.

4. Summary

Regression methods are often used to post-process model forecasts. When the regression parameters are known precisely, such regression forecasts lead to reliable probability forecasts. However, even in the idealized situation when the forecast and observation distributional forms are known (for instance, joint Gaussian), sampling error leads to unreliable forecast probabilities, with the lack of reliability being in the form of overconfidence. Sampling error is worse when the sample size is small or the underlying skill level is low. This systematic overconfidence is somewhat surprising since estimated regression coefficients are known to be unbiased estimates of the actual coefficients. The explanation for this behavior is that the signal variance of a regression forecast is the sum of the actual signal variance and a term that depends on the variance of the coefficient estimate. To the extent that the latter term is nonzero, the signal variance of the regression forecast is positively biased. This explanation suggests that shrinkage methods like ridge are useful. Shrinkage methods also are useful in the selection of predictors in multivariate pattern-based MOS approaches.

References

DelSole, T., 2005: Predictability and information theory. Part II: Imperfect forecasts. *J. Atmos. Sci.*, **62**, 3368–3381.

Glahn, H. R. and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, **11**, 1203–1211.

Hoerl, A. and R. Kennard, 1988: *Encyclopedia of Statistical Sciences*, **8**, chap. Ridge Regression, 129–136. Wiley, New York.

Johnson, C. and N. Bowler, 2009: On the reliability and calibration of ensemble forecasts. *Mon. Wea. Rev.*, **137**, 1717–1720.

Tibshirani, R., 1996: Regression shrinkage and selection via the lasso. *J. R. Statist. Soc. B*, **58**, 267–288.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Feasibility of Dynamical Seasonal Precipitation Prediction for the Pacific Islands

H. Annamalai[1], K.P. Sooraj[1], A. Kumar[2] and H. Wang[2]

[1]*IPRC/SOEST, University of Hawaii, Honolulu, HI*

[2]*Climate Prediction Center, NOAA/NWS/NCEP, Camp Springs, MD*

1. Introduction

With a substantial portion of the world's population influenced by climate variability, such as drought, flood, heat and cold waves, any capability to anticipate these fluctuations one or more seasons in advance would have measurable benefits for decision making in many sectors of society (Barnston *et al.* 2005). From the early 1980s to late 1990s, our ability to develop and improve coupled climate models have led to the ability to predict tropical climate variations with some success (Kang and Shukla 2006; Kirtman and Pirani 2009). Here, we examine the seasonal forecast performance of the National Centers for Environmental Prediction (NCEP) Coupled Forecast System (CFS) over the Tropics, and in particular, over the United States Affiliated Pacific Islands (USAPI).

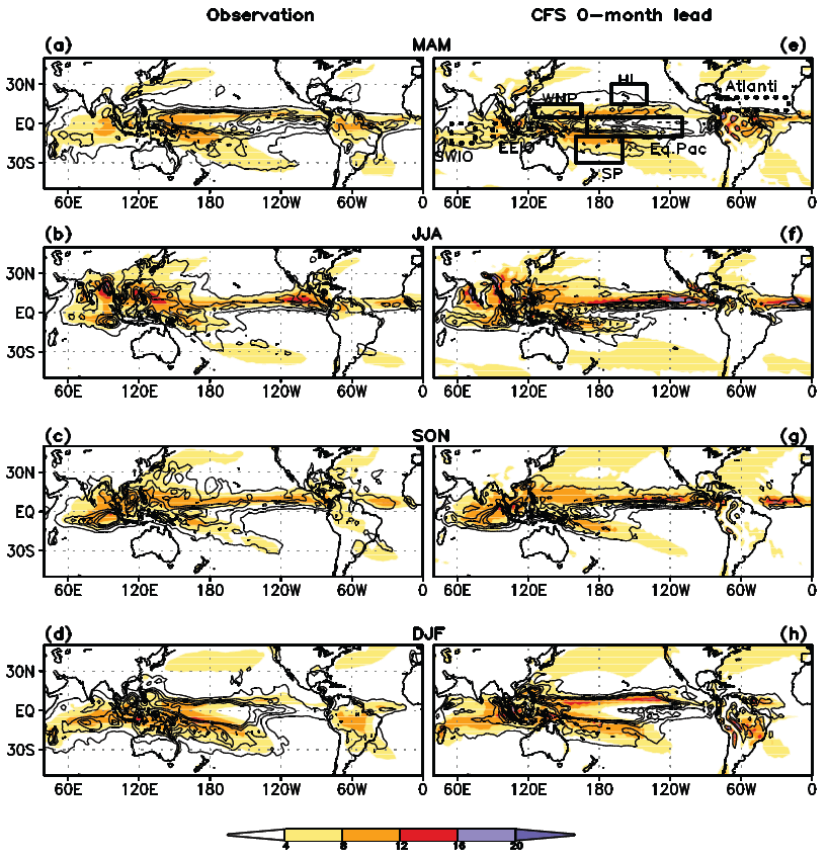


Fig. 1 Climatological precipitation (mm/day, shaded) and variance (mm^2/day, contours) for four standard seasons, (a-d) Observations and (e-h) 0-month CFS forecast. The boxed areas in (e) represent the regions where CFS ability in forecasting seasonal SST or precipitation anomalies are assessed. The area averaging used are: south west Indian Ocean (15°S-0°, 55°-75°E; SWIO); eastern equatorial Indian Ocean (10°S-0°, 90°-110°E; EEIO); northern Atlantic (20°-80°W, 10°-20°N) ; western north Pacific (5°-15°N, 125°-155°E; WNP); south Pacific (10°-30°S, 160°-200°E; SP); Hawaii (15°-30°N, 140°-170°W; HI), and equatorial Pacific (10°S-5°N, 170°-110°W; EPAC). Contours are drawn starting from 1 mm^2/day with an interval of 2 units.

Seasonal prediction over the tropics, and to a certain degree over the extratropics, is essentially linked to the accurate prediction of tropical SST (Shukla 1998; Kumar and Hoerling 1998; Goddard *et al.* 2001). Due to their impact on global climate anomalies (Ropelewski and Halpert 1987), predicting SST variations during the life cycle of El Niño–Southern Oscillation (ENSO) has been the focus of coupled model development. While efforts are underway to reduce model systematic errors, many Coupled General Circulation Models (CGCMs) have shown skill in capturing ENSO characteristics, paving ways for routine operational seasonal forecasts (Anderson *et al.* 2003; Saha *et al.* 2006). Recent studies have evaluated ENSO skills by analyzing hindcasts produced by CGCMs (Jin *et al.*

Correspondence to: H. Annamalai, International Pacific Research Center and School of Ocean and Earth Science and Technology, University of Hawaii, Honolulu, HI 96822; E-mail: hanna@hawaii.edu.

2008; Luo *et al.* 2008; Stockdale *et al.* 2011; Sooraj *et al.* 2011). Despite success, the predictive skills of models are degraded during the ENSO onset and decay periods (Barnston *et al.* 1999). Recent research interests have also focused on twotypes of El Niño namely, western Pacific and cold tongue events (*e.g.*, Kug *et al.* 2009).

While it is recognized that ENSO influences global climate anomalies, sensitivity experiments with AGCMs suggest that SST variations over the southwest Indian Ocean (SWIO) also modulate the amplitude of ENSO-induced circulation and precipitation anomalies over the tropical west Pacific (Watanabe and Jin 2003; Annamalai *et al.* 2005), and over the Pacific – North American (PNA) region (Annamalai *et al.* 2007). In summary, the expectation is that the combined effects of the tropical Pacific and Indian Oceans probably can further strengthen global climate anomalies.

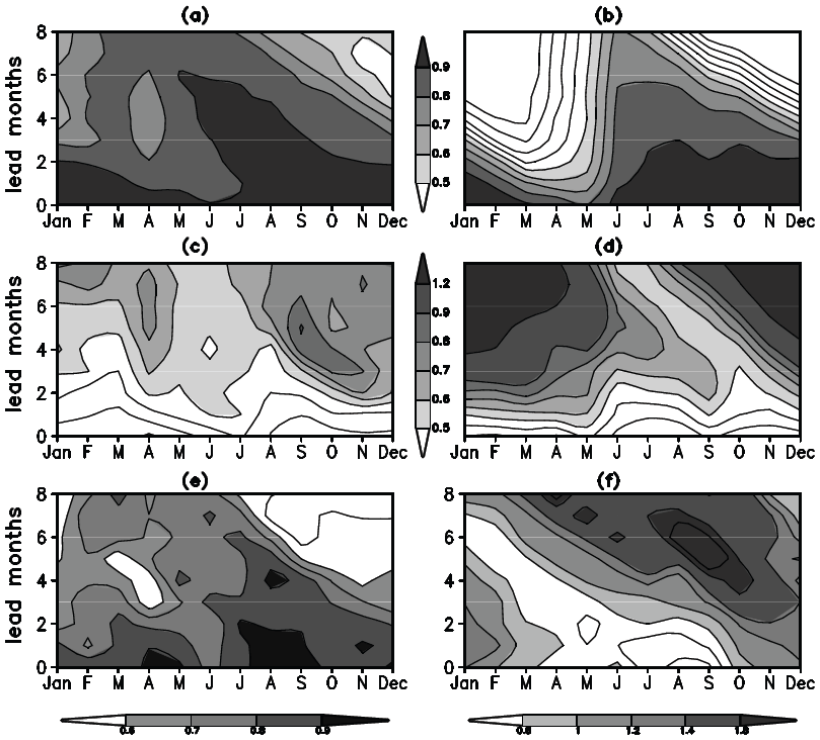


Fig. 2 (a) Anomaly correlation coefficient (ACC) of CFS ensemble mean forecasts of the monthly mean Nino3.4 SST over the period (1981-2005) as a function of initial condition month (x-axis) and lead months (y-axis). Nino3.4 is defined as the spatial mean SST over 5°S-5°N, 170°-120°W. (b) same as (a) but for persistence forecast. (c) same as (a) but for root mean square error (RMSE) for CFS ensemble forecast, and (d) same as (c) but for persistence forecast. (e) same as (a) but for CFS mean forecasts of the monthly mean precipitation anomalies over the equatorial Pacific (10°S-5°N, 170°E-110°W). (f) same as (c) but for monthly mean precipitation anomalies over the equatorial Pacific.

For the target regions over the USAPI (Fig. 1), the current operational seasonal precipitation prediction system is based on empirical methods in which SSTs provide the most reliable predictive information, and higher prediction skill is noticed during ENSO winters (He and Barnston 1996). During non-ENSO and weak to moderate ENSO events too, USAPI experience significant seasonal rainfall anomalies. In these circumstances, the precipitation forecast skill by empirical model is low, and the reasons may be manifold including: a) nonlinear relationship between ENSO SST and precipitation is not incorporated, b) details in the space-time evolution of SST during different flavors of ENSO are not properly accounted for, and c) SST anomalies other than ENSO may be responsible for rainfall variations. A prediction system based on a fully coupled dynamical model may overcome some of the above limitations.

2. Hindcasts and skill measures

Here, we analyze the output from CFS retrospective predictions (or alternatively referred to as hindcasts) that cover all 12 calendar months from 1981 to 2005. These hindcast runs, each of which is 9-month integration, are ensemble of 15 members starting from perturbed real-time oceanic and atmospheric initial conditions (ICs). Note that, for a 9-month integration, prediction at 6-month lead (L6) is the longest lead available for seasonal means while it is 8-month lead (L8) for prediction of monthly means. Variables examined in our analysis include SST, precipitation, and wind at 850 hPa. To infer the ocean Rossby waves in the tropical Indian Ocean (TIO), we also analyzed sea surface height (SSH). Hindcast anomalies are computed by removing the model climatology for each grid point, each initial month, and each lead time from the original ensemble hindcasts. Regarding verification of the real-time forecasts for the period 2006-09, we

examined 15-member ensemble members corresponding to the same ICs as in hindcasts. For the period 1981-2009, observed datasets used for verification include the CPC merged analysis of precipitation (CMAP) (Xie and Arkin 1996), winds from the NCEP/DOE reanalysis (Kanamitsu *et al.* 2002), and the NOAA optimally interpolated SST analysis (Reynolds *et al.* 2002). SSH is taken from the Global ocean data assimilation system. For verification observed data are interpolated to CFS' horizontal resolution (T62).

The entire range of available hindcasts (0 to 6 month lead) for all the four standard seasons is verified. To infer the uncertainty measures associated with forecasts, the skills are assessed using deterministic [anomaly correlation coefficient (ACC), the ensemble spread and signal-to-noise ratio (S/N)], categorical [Heidke skill score (HSS)], and probabilistic [rank probability skill score (RPSS)] methods. For the target regions (Fig. 1), the distribution of ensemble members, ensemble mean, and observed anomalies are plotted for every year. These plots aid in understanding year-to-year variations in the spread and its association with predicted value, and if there is asymmetry in model's predictive skill (positive *vs* negative values). More details are in Sooraj *et al.* (2011).

(a) Niño 3.4 region

Lead Months	DJF		MAM		JJA		SON	
	ACC	S/N	ACC	S/N	ACC	S/N	ACC	S/N
0	0.96	6.1	0.95	4.8	0.92	3.2	0.96	3.7
1	0.92	5.2	0.91	3.0	0.82	2.6	0.93	3.2
2	0.92	4.5	0.85	2.9	0.88	2.2	0.92	3.3
3	0.94	3.6	0.84	2.9	0.82	2.0	0.88	2.5
4	0.91	3.2	0.84	2.9	0.68	1.4	0.76	2.2
5	0.91	2.8	0.81	2.9	0.50	1.5	0.85	2.0
6	0.89	2.3	0.83	2.5	0.53	1.8	0.79	2.0

(b) SWIO region

Lead Months	DJF		MAM		JJA		SON	
	ACC	S/N	ACC	S/N	ACC	S/N	ACC	S/N
0	0.85	1.7	0.92	1.8	0.64	1.7	0.73	1.9
1	0.84	1.7	0.87	1.6	0.57	1.2	0.60	1.3
2	0.85	1.5	0.84	1.4	0.52	1.1	0.47	1.3
3	0.82	1.4	0.85	1.5	0.50	0.9	0.42	1.1
4	0.81	1.3	0.83	1.4	0.58	1.2	0.27	1.1
5	0.75	1.1	0.85	1.5	0.57	1.0	0.13	1.1
6	0.71	1.1	0.83	1.4	0.67	1.0	0.10	0.8

(c) EEIO region

Lead Months	DJF		MAM		JJA		SON	
	ACC	S/N	ACC	S/N	ACC	S/N	ACC	S/N
0	0.76	1.2	0.87	1.3	0.80	1.5	0.89	1.2
1	0.61	1.0	0.80	1.0	0.73	1.3	0.76	1.0
2	0.40	1.0	0.75	1.0	0.18	0.9	0.59	1.0
3	0.29	1.0	0.77	1.0	0.03	0.7	0.71	1.0
4	0.46	0.9	0.71	0.8	-0.04	0.4	0.51	0.9
5	0.48	0.9	0.67	0.9	-0.3	0.5	0.18	0.9
6	0.38	0.8	0.72	0.9	-0.1	0.5	0.04	0.8

Table 1 Values of anomaly correlation coefficient (ACC) and signal to noise ratio (SN) estimated for regional SST time series for four standard seasons, and for 0-6 month lead forecast.

3. Hindcast skills

Figure 1 shows climatological precipitation (shaded) and variance (contours) for standard seasons from observations (left) and at 0-month lead hindcast from CFS (right). Compared to observations, CFS captures the seasonal dependency in the position and intensity of the rainfall maximum and also regional variance maxima over the tropics with some systematic errors. Motivated by this, apart from examining SST skill, we also evaluate CFS' ability in forecasting: (i) tropical and regional precipitation anomalies; (ii) different flavors of El Niño and their associated regional response; and (iii) teleconnection between the tropical Pacific and Indian Ocean (TIO). In terms of regional indices, skill is examined for precipitation (area outlined in solid lines in Fig. 1), and SST influenced by thermocline variations over the TIO (area outlined in dotted lines in Fig. 1).

a. SST over the equatorial Pacific

Figure 2 shows ACC (Fig. 2a) and RMSE (Fig. 2c) for Niño3.4 SST anomalies estimated for the entire period. Results are shown for all lead times (0-8 months), and for all ICs (January through December). The corresponding measures estimated for persistence as the forecast are shown in Figs. 2b and 2d. The inverse association between ACC and RMSE holds good, and ensemble mean offers higher skills than persistence.

For instance, for leads up to 6-7 months and hindcasts initialized during late spring through early fall (May through October), the ensemble mean ACC is > 0.8 with relatively small RMSE (0.3-0.4) but for persistence forecast the ACC drops below 0.6 with higher RMSE (~0.7-0.8). This means that skill in forecasting the intensification and peak amplitude of ENSO is indeed high. However, for forecasts initialized during winter (November through January), skill in hindcasting the transition phase of ENSO is modest at best. Specifically, both methods share similar ACC and RMSE values in the first 3-5 months (Figs. 2a-d), but they drop off rapidly during spring with a minimum in July. This decay in skill, often referred to as the spring predictability barrier, is common to most models (*e.g.*, Jin and Kinter 2009), and the forecast performance of CFS over Niño3.4 region is comparable to the new ECMWF S3 system (Stockdale *et al.* 2011).

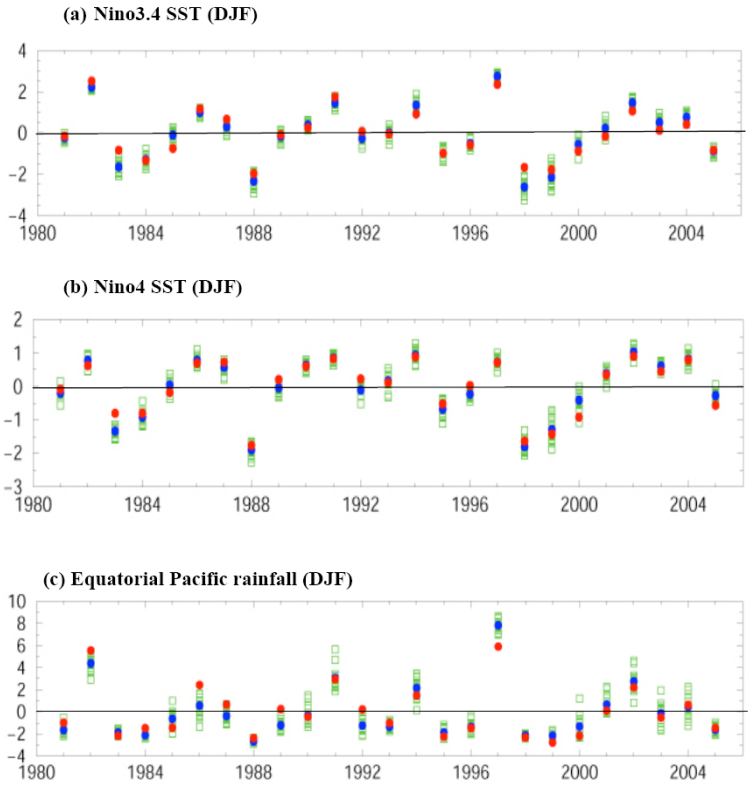


Fig. 3 (a-c) Temporal evolution of December through February average (DJF) SST anomalies (^{o}C) hindcast by CFS at lead 0-month. Ensemble mean (blue), all 15 individual members (green) and observations (red) are shown for three regions: (a) Nino3.4 (5^{o}S-5^{o}N, 190^{o}E-120^{o}W); (b) Niño4 (5^{o}S-5^{o}N, 160^{o}E-150^{o}W); (c) same as (a) but for equatorial Pacific rainfall anomalies (mm/day) averaged over (10^{o}S-5^{o}N, 170^{o}E-110^{o}W).

ENSO prediction is further assessed in many different ways. First, the model's relative skill in hindcasting individual El Niño and La Nina events (Fig. 3a) is assessed for the peak phase (December through February; DJF) at 0-month lead. This is necessitated because observations during the period 1982-2005 indicate unique feature that models need to forecast. They include: (i) two of the strongest El Niño of the 20th century (1982 and 1997) and (ii) persistent La Nina during 1998-2000 (thick red circles in Fig. 3a). It is encouraging that all the ensemble members (green circles) capture correctly the amplitude during 1982 and 1997, and the model's ability in hindcasting these two events is remarkable even for leads up to 6 months (Sooraj *et al.* 2011). For the prolonged La Nina episode, however, the predicted amplitude is higher in conjunction with a larger ensemble spread. We also note that the model is able to correctly forecast near-normal conditions in many years. Second, we analyzed CFS' skill in forecasting different flavors of El Niño (Fig. 3b). Observations indicate that SST maximum for cold tongue events is over the eastern Pacific or Niño3 region (5^{o}S-5^{o}N, 90^{o}W-150^{o}W), and that for warm pool events the maixmum lies over the west-central Pacific or Niño4 region (5^{o}S-5^{o}N, 160^{o}E-150^{o}W). Therefore, skills in hindcasting SST anomalies over Niño3 and Niño4 regions are examined. Owing to the role of wave-induced thermocline displacements influening SST through vertical advection even at 6-month lead time, the ACC remains high (0.9) for Niño3 region and other skill measures are as good as that over the Niño3.4 region, and hence, are not discussed further. Over the Niño4 region, local warming during boreal winter of 1990-91; 1994-95; 2002-03 and 2004-05 are forecasted with a minimum spread (Fig. 3b). An examination for the entire period (not shown) suggests that prediction of SST anomalies over both Nino3 and Nino4 regions are comparable to those over Nino3.4 for all short leads. In summary, CFS' skill is higher for forecasting stronger and persisting El Niño events that are primarily due to thermocline displacements. When seasonally stratified, predicting winter and autumn seasonal SST anomalies have higher skill. However, forecasting weaker El Niño events is limited to 0-2 month leads.

b. Precipitation along the equatorial Pacific

The over-all skill measures for precipitation forecast along the equatorial Pacific (170°E-110°W; 10°S-5°N) are shown in Figs. 2e-f. High values of ACC (> 0.7) with less RMSE are along the diagonal representing fall and winter seasons. Only for short lead times (0-2 months), predicting spring and summer rainfall is skilful, and similar to SST, forecasting precipitation anomalies during ENSO phase transition is difficult. Encouraged by the skill in forecasting aspects of El Niño-related SST anomalies in individual years, we turn our attention to precipitation skill along the equatorial Pacific at 0-month lead for DJF season (Fig. 3c). While the ensemble mean follows the observed, the spread is larger yielding a S/N of about 4.2, and this skill is considerably less than that for Niño3.4 SST (Table 1a). Other limitations include excess rainfall during the winter of 1997/98, and the failure in capturing above normal rainfall during the winter of 1986-87. In addition, during cold phases of 1983-84 and 1998-99 (Fig. 3a), precipitation forecast is weaker than observed (Fig. 3c) despite predicting stronger SST anomalies compared to observations. SST anomalies translating into rainfall anomalies depends on physical parameterizations employed, particularly convective schemes. Further, SST-rainfall relationship is also not local, and rainfall anomalies can be influenced by circulation anomalies forced by rainfall at other locations.

(a) West Pacific region (125°-165°E, 5°N-15°N)

Lead Months	DJF		MAM		JJA		SON	
	ACC	S/N	ACC	S/N	ACC	S/N	ACC	S/N
0	0.87	2.3	0.66	1.7	0.66	1.2	0.31	0.8
1	0.77	1.9	0.63	1.2	0.69	1.1	-0.07	0.8
2	0.73	1.5	0.57	1.4	0.61	0.9	-0.06	0.6
3	0.84	1.7	0.54	1.5	0.69	0.8	-0.44	0.5
4	0.75	1.5	0.47	1.3	0.65	0.5	0.28	0.5
5	0.72	1.3	0.50	1.5	0.56	0.5	-0.07	0.4
6	0.58	1.3	0.53	1.5	0.22	0.5	-0.09	0.4

(b) South Pacific region (160°E-160°W, 10°S-30°S)

Lead Months	DJF		MAM		JJA		SON	
	ACC	S/N	ACC	S/N	ACC	S/N	ACC	S/N
0	0.69	2.0	0.72	1.7	0.56	1.8	0.72	1.7
1	0.69	1.9	0.75	1.4	0.46	1.4	0.62	1.7
2	0.63	1.6	0.72	1.5	0.32	1.2	0.62	1.4
3	0.56	1.4	0.65	1.5	0.32	1.2	0.52	1.3
4	0.51	1.2	0.67	1.2	0.44	1.0	0.39	1.1
5	0.41	1.0	0.74	1.6	0.22	1.0	0.49	1.0
6	0.35	1.0	0.71	1.4	0.37	1.2	0.48	1.0

(c) Hawaiian region (170°W-140°W, 15°N-30°N)

Lead Months	DJF		MAM		JJA		SON	
	ACC	S/N	ACC	S/N	ACC	S/N	ACC	S/N
0	0.62	1.1	0.69	1.5	0.57	1.4	0.49	0.9
1	0.69	1.0	0.44	1.3	0.69	1.1	0.44	1.0
2	0.60	0.9	0.52	1.3	0.62	1.1	0.32	0.9
3	0.57	0.9	0.49	1.1	0.64	1.1	0.22	0.8
4	0.34	0.7	0.46	1.2	0.46	0.8	0.39	0.7
5	0.27	0.6	0.34	1.3	0.55	0.9	0.12	0.6
6	0.14	0.6	0.49	1.0	0.61	0.9	0.34	0.7

Table 2 As in Table 1 but for regional precipitation time series over the Pacific Islands.

Finally, to depict coherency among variables responsible for ocean-atmosphere interactions, lagged correlations between DJF Niño3.4 SST and SST (contours), rainfall (shaded), and 850 hPa wind (vector) anomalies at lead 0 month and spanning the 24-month period (entire life-cycle of ENSO) are shown in Fig. 4b. Similar results from observations are also shown in Fig. 4a. In CFS, the onset, development, mature and decay stages of El Niño, as well as the development of cold SST and negative precipitation anomalies over the Maritime Continent (120°E-150°E) together with low-level divergence corresponding to anomalous Walker Circulation are in good agreement with observations. To the west of this center of divergence, easterly wind anomalies cover the entire equatorial Indian Ocean, and their influence on local ocean dynamics is explained next.

c. Teleconnection to the TIO

Figure 4c shows lagged correlations between DJF Niño3.4 time series and SSH (shaded) and SST (contours) averaged over (8°S-12°S) the TIO from observations, and the corresponding results from the 0-lead CFS forecast are shown in Fig. 4c. Starting from May-June of Year [0], upwelling- favourable winds off Java-Sumatra coasts (Figs. 4a-b; 80°E-100°E) promote negative SSH values (*i.e.*, shallow thermocline anomalies, Fig. 4d) and subsequent local SST cooling attains a maximum in fall (Figs. 4a-b). The wind-stress curl

associated with the easterly wind anomalies forces downwelling oceanic Rossby waves, and the westward tilted structure in SSH and SST anomalies with respect to time (Figs. 4c-d) support that interpretation. These Rossby waves advect warm water and act to deepen the thermocline as they cross the ocean basin (Xie *et al.* 2002). The maximum perturbation to SSH (*i.e.*, deepened thermocline) and SST are noted over SWIO during boreal spring of Year [+1]. The warm SST anomalies persist for about 10-12 months primarily due to the presence of shallow-mean thermocline and passage of oceanic Rossby waves. In summary, CFS captures the teleconnection from the tropical Pacific to TIO, and also the essential mechanisms responsible for the anomalous conditions in the TIO.

Encouraged by the above results, we examined skill scores over both SWIO (55°E-75°E; 15°S-0°), and EEIO (90°E-110°E, 10°S-0°). Observations indicate SST anomalies peak during spring of Year [+1] over SWIO and during boreal fall over EEIO, respectively. For SWIO, CFS hindcasts initialized from June-July of Year [0] onwards depicts higher ACC and lower RMSE than persistence method but in contrast, over EEIO ensemble mean barely does better than persistence even at shorter leads (Sooraj *et al.* 2011).

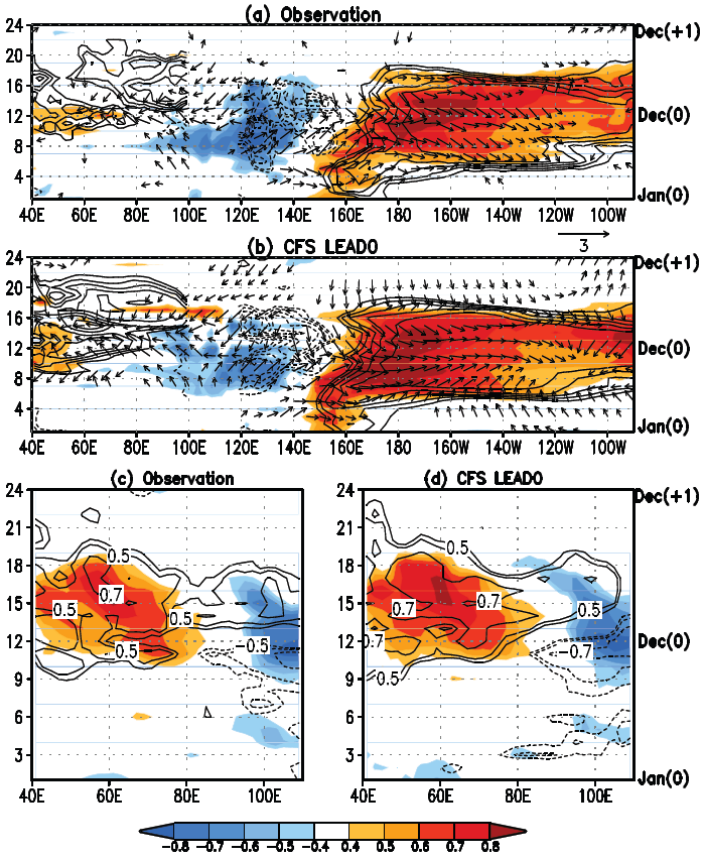


Fig. 4 (a) Lagged correlations of SST (contours), rainfall (shaded) and 850 hPa wind averaged in 3°S-3°N with winter (DJF) Nino3.4 SST index from observations. (b) same as (a) but from CFS ensemble mean 0-month lead forecast. Results are shown for a twoyear period representing the entire life-cycle of ENSO. (c) lagged correlations of SST (contours) and SSH (shaded) averaged in 8°-12°S with winter Nino3.4 SST index from observations. (d) same as (c) but from CFS ensemble mean 0-month lead forecast. Positive (negative) SST values are shown as solid (dashed) contours with an interval 0.1.

d. Precipitation over the USAPI

Figure 5 shows rainfall forecast over west Pacific islands (125°E-165°E; 5°N-15°N). The left (right) panels are results for 0-month (6-month) lead-time for standard seasons of summer (Figs. 5a-b), fall (Figs. 5c-d), winter (Figs. 5e-f) and spring (Figs. 5g-h). Observations (red circles) indicate that during strong El Niño years (*e.g.*, 1982-83; 1991-92; 1997-98), dryness (or below normal rainfall) persist from winter of Year [0] to summer of Year [+1]. Quite remarkably, CFS ensemble-mean forecast (blue circles) captures this drying tendency from 6-months lead time except for summer of Year [+1]. Barring the two strong El Niño events of 1982/83 and 1997/98, the spread among the ensemble members is large but the sign of the anomalies is well-captured. For 6-month lead forecast of JJA rainfall the predicted sign is wrong in many years. While correlations between rainfall anomalies over west Pacific and Nino3.4 SST anomalies is negative for most of the year, it is rather positive during July-August. Thus, the difficulty in capturing this seasonally-varying teleconnection may well be a factor in the low skill scores during summer. A possible interpretation is that large amplitude swings in precipitation occur during winter (Fig. 5e), and hence are more predictable.

Fig. 6 shows the results for the south Pacific islands (160°E-200°E; 10°S-30°S), and the skill statistics is summarized in Table 2b. In this region too, largest anomalies in precipitation are observed during ENSO winters (Figs. 6e-f) but the dryness starts in summer of Year [0] and persists until the following spring.

Interestingly, CFS correctly captures both the phase and amplitude of the dryness at longer lead times. However, for any given season ACC and S/N do not exceed 0.75 and 2.0, respectively. Compared to other seasons, forecasting rainfall anomalies during summer is less skilful (Table 2b) since ENSO teleconnection is weaker.

Another region of interest is the Hawaiian Archipelago (170°W-140°W; 15°N-30°N), and the precipitation forecasts are shown in Fig. 7 and the statistics are summarized in Table 2c. For the islands situated over the Northern Hemisphere, west Pacific and Hawaii, the dryness attains a maximum in winter of Year [0] and continues into the spring of Year [+1], and CFS has skill in forecasting them. However, the spread among the members is indeed large even during 1982-83 and 1997-98 El Niño years (Fig. 7) resulting in a low S/N (Table 2c). From Figs. 5-7 it is encouraging to note that observed anomalies (red dots), generally lie within the envelope of possible model solutions (green dots), and there are few instances of outliers.

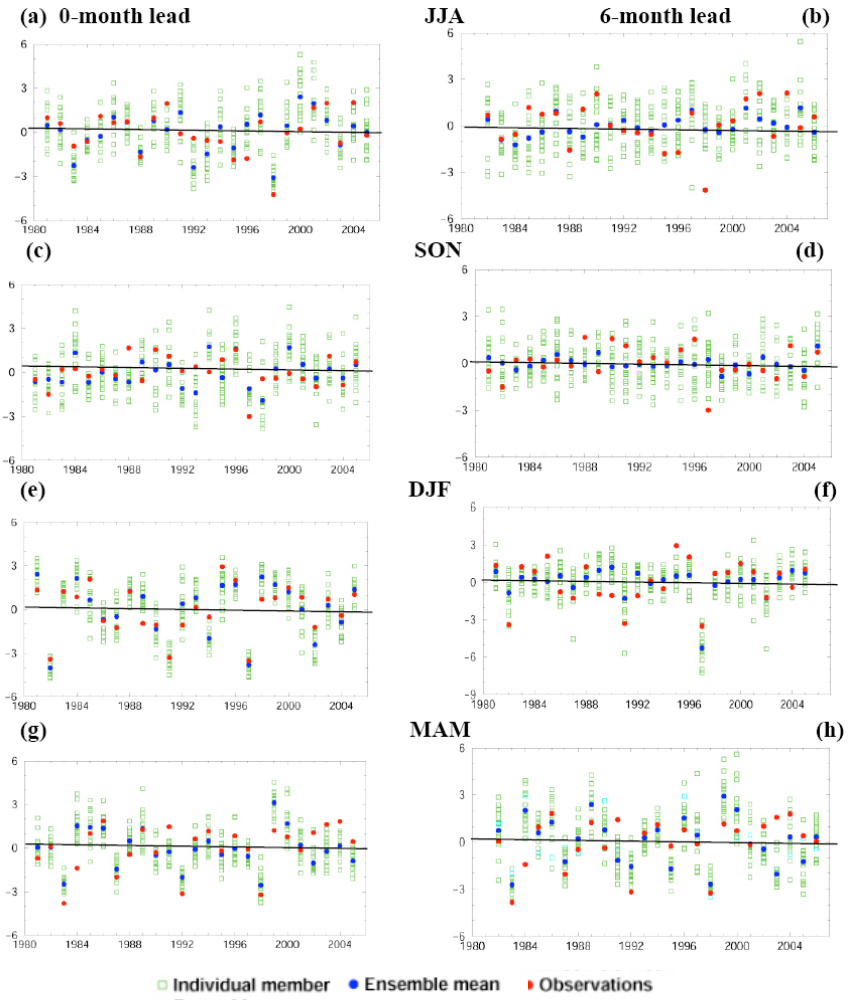


Fig. 5 Seasonal rainfall (mm/day) forecast at 0-month (left) and 6-month (right) lead time over tropical west north Pacific region for the period 1981-2005. Observations (red), ensemble-mean (blue) and all the 15 individual members (green) are shown.

4. Assessment of different aspects of forecasts

Here, ACC is compared and contrasted with HSS and RPSS to assess different aspects associated with forecasts that also rely on the spread information inherent in the ensembles (Kumar 2009). Our working hypothesis is that forecasts assessment based on ACC alone are perhaps not sufficient enough in the context of decision-making. As before, we assess the scores for SST indices (Fig.8). The left panels are scatter diagrams between ACC and HSS, while the right panels are scatter plots between ACC and RPSS, respectively.

CFS demonstrates highest confidence in predicting winter SST anomalies (open squares, Figs. 8a-b) over Niño3.4 region at all lead months with ACC upwards of 0.85. High skill in deterministic forecasts as measured by ACC is also captured for categorical and probabilistic forecasts with high HSS (> 50%) and RPSS (> 40%). This is consistent with the results discussed earlier (Fig. 2). The confidence in predicting fall season is also high but limited to 0-4 month leads. Predicting spring and summer anomalies has limited confidence at 0-1 month leads, and for other leads, even if ACC lies around 0.8, RPSS drops off to very low values sometimes even negative, indicating possible issues related to the spread among the ensemble members that influence the probabilistic forecasts. Therefore, forecasting the summer teleconnection features, *e.g.*, ENSO-monsoon association will be limited in CFS. The inference is that results of Fig. 2a and Table 1a

are encouraging, but additional diagnostics are necessary to attest the errors associated with the forecast for any decision-making.

For the SST indices over the TIO, the confidence in predicting the winter and spring variations over SWIO (Figs. 8c-d) are high at leads up to 5 months. Watanabe and Jin (2003) and Annamalai *et al.* (2005; 2007) noted that it is the winter and spring SST anomalies over SWIO that influence regional and global climate anomalies. In observations, summer and fall SST variations over SWIO are indeed small, and hence models have limited predictability. Confidence in predicting Indian Ocean Dipole Zonal Mode (IODZM) SST anomalies during fall is the least (Figs. 8e-f).

The scatter plots over the USAPI are shown in Fig. 9. Over the west Pacific region, forecasting rainfall variations during winter is trustworthy for at least 0-3 months lead, followed by prediction for spring season at lead 0-1 months. For summer rainfall variations, while ACC is greater than 0.6 and HSS around 15-25%, negative RPSS indicates that the forecast is not better than climatology. The model's forecast for fall rainfall anomalies is least skillful. A point to note here is that for the same value of ACC (~0.6) for 0-1 month lead forecast, RPSS is positive for spring but negative for summer indicating the seasonal dependency in the forecast errors. For the south Pacific region, predicting rainfall variations during winter and spring appear realistic for leads 0-4 months. Here also forecasting summer precipitation anomalies are not reliable. For the Hawaiian region, at shorter leads (0-1 months) and for all seasons except fall, convergence of all the three scores suggests that different rainfall forecast information are skilful and probably useful. In summary, for regional precipitation forecast over USAPI, ACC values greater than 0.7 and correspondingly high HSS and RPSS values occur particularly for winter and spring seasons.

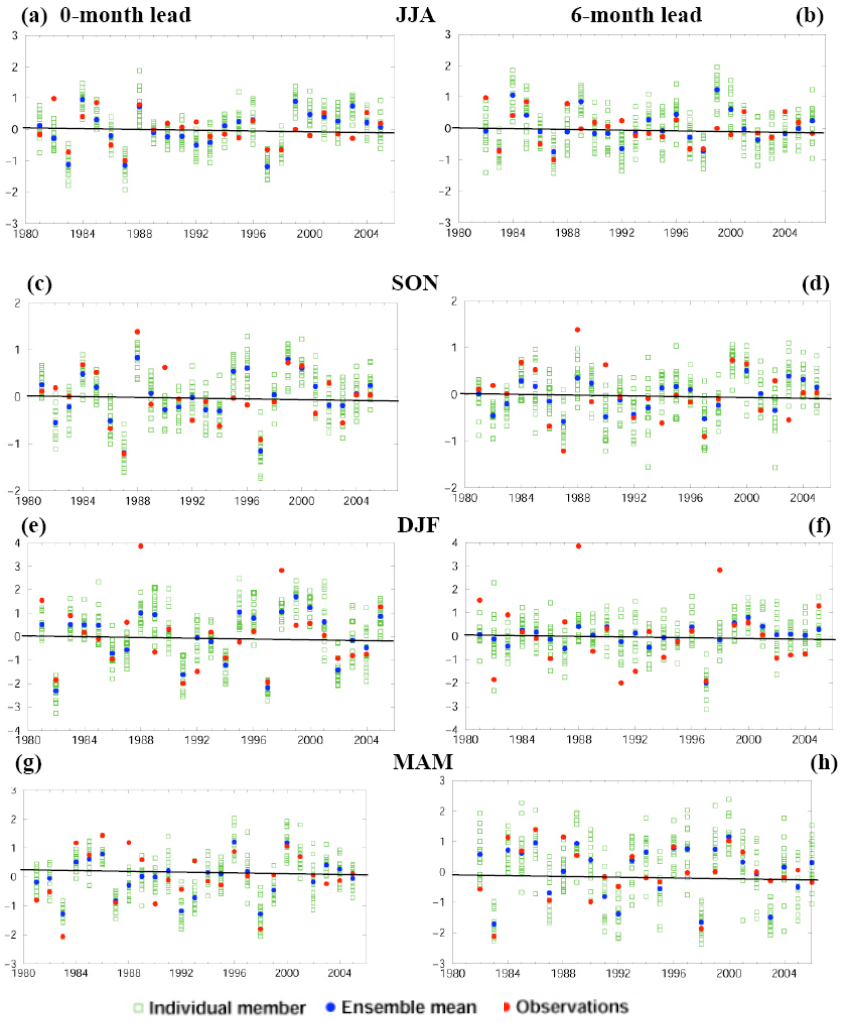


Fig. 6 Same as Fig. 5 but for the precipitation (mm/day) over south Pacific region.

5. Implications for dynamical seasonal prediction of precipitation

Both observational (Ropelewski and Halpert 1987; 1989) and modeling (*e.g.*, Shukla 1998; Su and Neelin 2002) studies provide a guide to the expected climatic impacts of ENSO over the tropics. The fact that an accurate prediction of the tropical Pacific SST is a necessary condition for successful prediction of rainfall along the equatorial Pacific is supported by the present analysis. Based on the lagged association between TIO and tropical Pacific SST anomalies, past seasonal prediction studies have used statistical methods to predict TIO SST anomalies (*e.g.*, Mason et *al.* 1999). However, recent studies also suggest that these ENSO-induced regional SST anomalies may possibly alter the strength of the circulation and rainfall anomalies elsewhere

(Annamalai *et al.* 2005; 2007). The ability of CFS in representing the coupled processes in the TIO, particularly over SWIO during ENSO years, deserves further attention.

Predictive skill of regional precipitation, however, is usually lower compared to SST and circulation (Kumar and Hoerling 1998) because of the "noisy", small-scale character, and complex physics of precipitation. Even averaged over a season substantial irregularities in the spatial pattern are likely, particularly over the Tropics where convective rainfall is most common (Gong *et al.* 2003). Even so, observational and modeling studies suggest that the large-scale circulation pattern responsible for the precipitation anomaly may be predictable several months in advance, particularly during ENSO events. An examination of spatial pattern of DJF SST and rainfall anomalies during all El Niño events indicate that CFS is capable of forecasting the "details", in particular the observed negative rainfall anomalies over all the three USAPI regions during 1982/93 and 1997/87, as well above normal rainfall over Hawaii during 1990-91, and 2004-05 events (figure not shown).

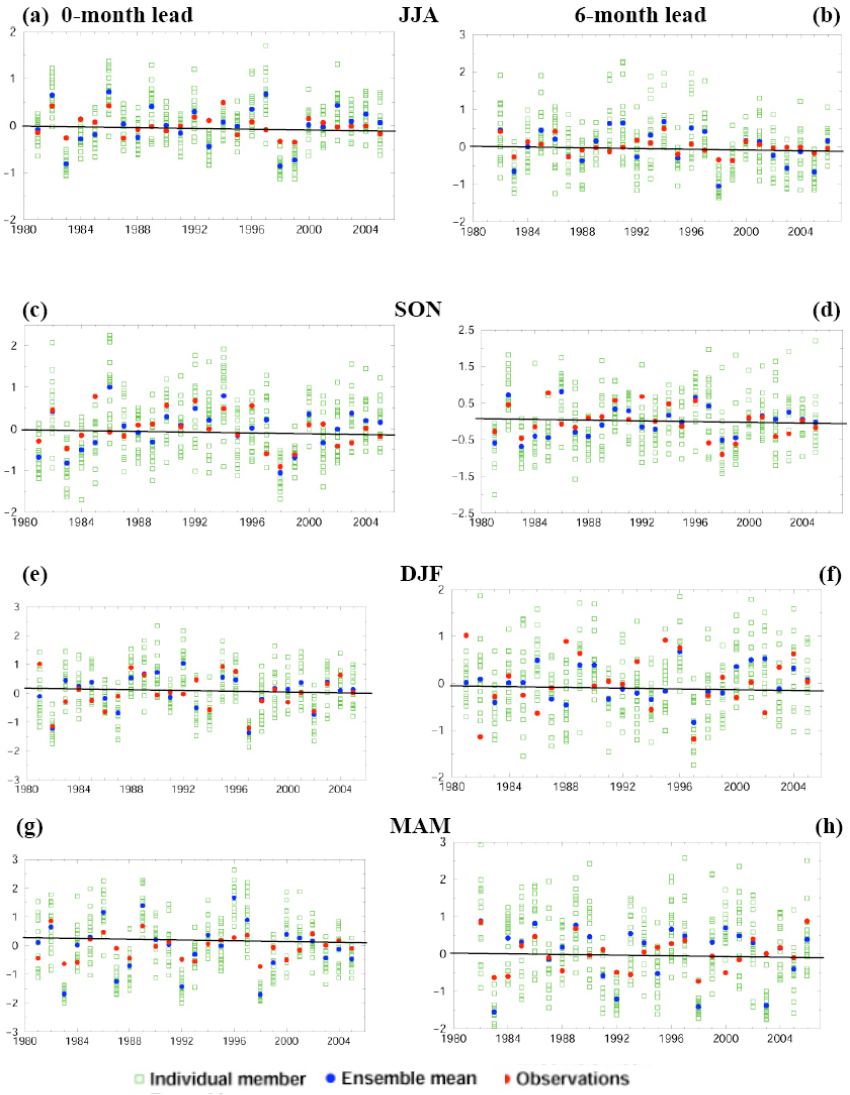


Fig. 7 Same as Fig. 5 but for the precipitation (mm/day) the Hawaii region.

Till date, the operational seasonal forecasting of precipitation over the USAPI stations relies on empirical method in which precipitation measured at individual stations itself is treated as a predictor (He and Barnston 1996). Their results based on the period 1955-94 suggest that at 1-month lead, ACC for predicting winter rainfall anomalies is ~ 0.4, 0.6, and 0.4 over south Pacific, west Pacific and Hawaiian regions, respectively (their Fig. 4). The dynamical forecast system based on CFS, on the other hand, demonstrates much higher skill at longer lead times over the USAPI (Table 2) when area-averaged fields are examined. Dynamical models represent the major components of the climate system (ocean, land and atmosphere) and can incorporate linear and nonlinear interaction processes among the components, and are expected to provide better seasonal forecasts than statistical models. A word of caution is that output from CFS represents averages over a relatively coarse grid, and therefore statistics of precipitation in regions of steep orography can differ substantially from that of station data. Nevertheless, owing to the success of CFS in forecasting the timing and amplitude of ENSO-related SST and precipitation anomalies along the equatorial Pacific and the associated teleconnection over the TIO we speculate that CFS has useful skill in forecasting regional rainfall anomalies over the USAPI. However, identifying individual physical processes responsible is beyond the scope of the present study. Lyon and Mason (2009) noted that the correct prediction of winter rainfall

anomalies during 1998 over southern Africa by some coupled models is probably not for the correct reasons. Our future study will examine the reasons for CFS performance over the USAPI.

6. Summary

The results presented here suggest the feasibility that a dynamical system based seasonal prediction of precipitation over the USAPI can be considered. It is necessary to continually assess the sources, and level of prediction skill as newer set of hindcasts based on improved models, and initial conditions obtained from more advanced data assimilation systems become available. For instance, one clear limitation in the hindcasts analyzed here is that for short-lead forecasts the initial conditions are from the old reanalysis system, a situation that would be rectified with the new CFS forecast model. In addition, in the updated CFS version the atmospheric model has a higher horizontal resolution (T126 compared to T62 in the current version), and may provide a better resolution necessary for precipitation prediction. In a future study, the physical processes that may be responsible for the performance of the model in predicting regional rainfall anomalies will also be examined. To attain station-level prediction of rainfall anomalies over the USAPI a downscaling system may need to be constructed.

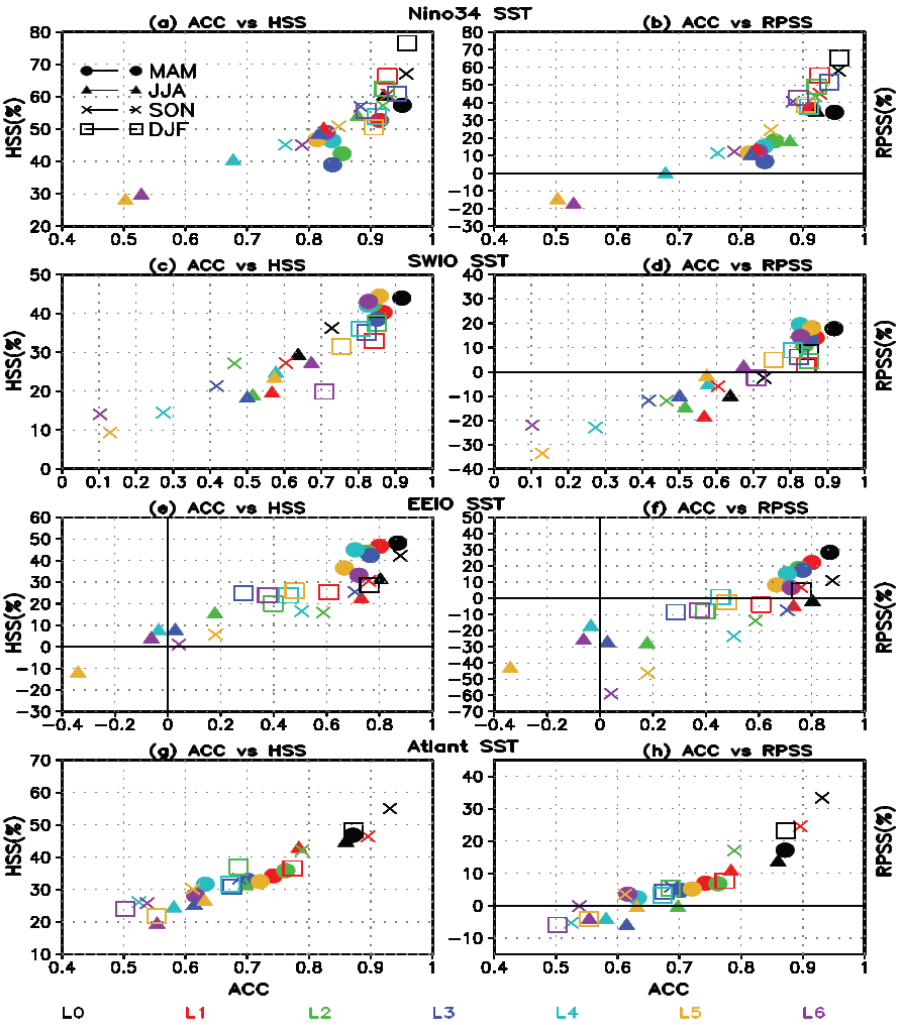


Fig. 8 Scatter diagram between ACC and HSS (left) and ACC and RPSS (right) for four standard seasons derived from CFS forecasts at 0-6 month leads; (a-b) for Nino3.4 SST, (c-d) SWIO SST, (e-f) EEIO SST indices and (g-h) Atlantic SST indices.

References

Anderson, D.L.T, and Coauthors 2003: Comparison of the ECMWF seasonal forecast system 1 and 2, including relative performance for the 1997/8 El Niño. Tech.Memo. **404**, ECMWF, Reading, UK, 93 pp.

Annamalai, H., P. Liu, and S.P. Xie, 2005: Southwest Indian Ocean SST variability: Its local effect and remote influence on Asian monsoons. *J. Climate,* **18,** 4150-4167.

—, H. Okajima, and M. Watanable., 2007: Possible impact of the Indian Ocean SST on the Northern Hemisphere circulation during El Niño. *J. Climate,* **20,** 3164-3189.

Barnston, A. G., M. H. Glantz, and Y. He., 1999: Predictive skill of statistical and dynamical climate models in SST forecasts during the 1997—98 El Niño episode and the 1998 La Niña onset. *Bull. Amer. Meteor. Soc.,* **80,** 217–244.

Barnston, A. G., A. Kumar, L. Goddard, and M. P. Hoerling, 2005: Improving seasonal prediction practices through attribution of climate variability. *Bull. Amer. Meteor. Soc.*, **86**, 59–72.

Goddard, L., S. J Mason, S. E. Zebiak, C. F. Ropelewski, R. Basher, and M. A. Cane, 2001: Current approaches to seasonal-to-interannual climate predictions. *Int. J. Climatol*, **21**, 1111-1152.

Gong, X., A.G. Barnston, and M.N. Ward, 2003: The effect of spatial aggregation on the skill of seasonal precipitation forecasts. *J. Climate,* **16,** 3059-3071.

He, Y., and A.G. Barnston, 1996: Long-lead forecasts of seasonal precipitation in the tropical Pacific islands using CCA. *J. Climate,* **9,** 2020-2035.

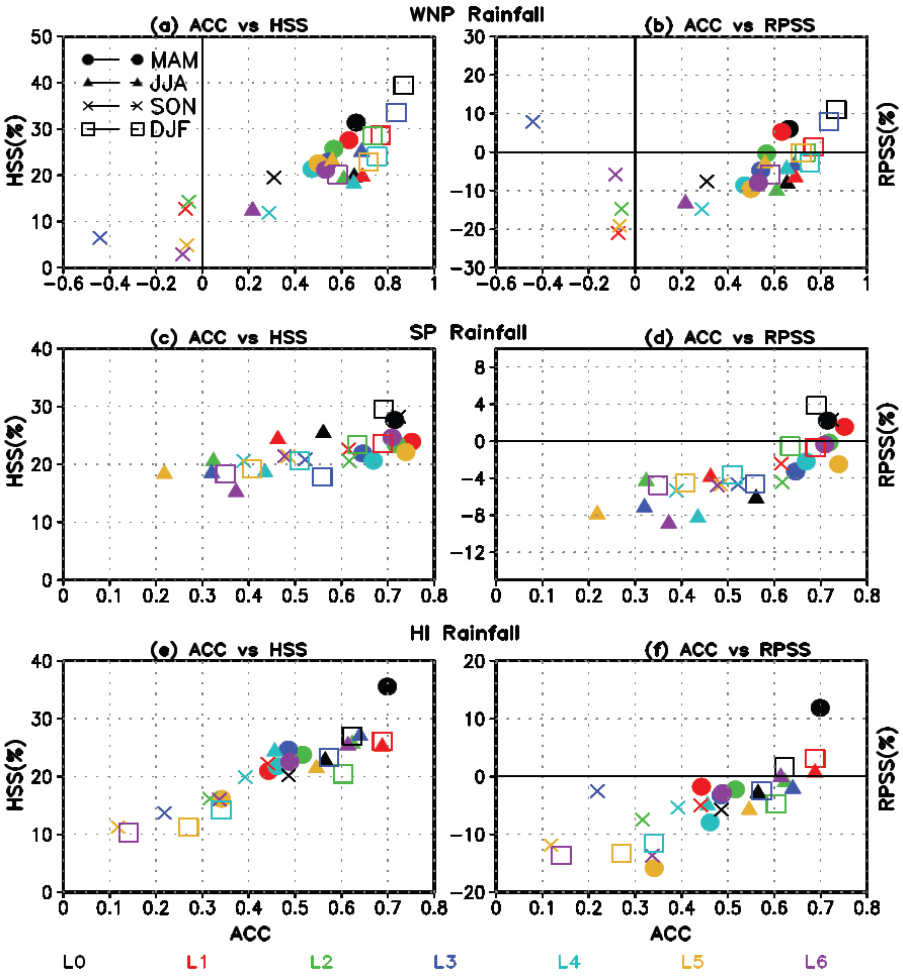


Fig. 9 Same as Fig. 8 but for regional precipitation indices; (a-b) west north Pacific, (c-d) south Pacific, and (e-f) Hawaii.

Jin, E. K, and J.L. Kinter, 2009: Characteristics of tropical Pacific SST predictability in coupled GCM forecasts using the NCEP CFS. *Climate Dyn*, **32,** doi, 10.1007/s00382-008-0418-2

Jin, E.K., and Coauthors 2008: Current status of ENSO prediction skill in coupled ocean-atmosphere models. *Climate Dyn,* **31**, 647–664.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J.J. Hnilo, M. Fiorino, and G.L. Potter, 2002: NCEP-DOE AMIP-II Reanalysis (R-2), *Bull. Amer. Met. Soc.*, **83,** 1631-1643.

Kang, I-S., and J. Shukla, 2006: Dynamical seasonal prediction and predictability of the monsoon. B. Wang (Ed), The Asian Monsoon. Springer, 586-612.

Kirtman, B., and A. Pirani, 2009: The state of the Art of Seasonal Prediction: Outcomes and Recommendations from the First World Climate Research Program Workshop on Seasonal Prediction, *Bull. Amer. Meteor. Soc.*, **90,** 455-458.

Kug, J.S., F.F. Jin, and S.I. An., 2009: Two types of El Niño events: cold tongue El Niño and warm pool El Niño. *J. Climate,* **22,** 1499-1515.

Kumar, A., and M. P. Hoerling, 1998: Annual cycle of Pacific–North American predictability associated with different phases of ENSO. *J. Climate,* **11,** 3295–3308.

—, 2009: Finite samples and uncertainty estimates for skill measures for seasonal prediction, *Mon. Wea. Rev.*, **137,** 2622-2631.

Luo, J.J., S. Masson, S.K. Behera, and T. Yamagata, 2008: Extended ENSO predictions using a fully coupled ocean-atmosphere model. *J. Climate*, **21,** 84-93.

Lyon, B., and S. Mason, 2009: The 1997/98 summer rainfall season in southern Africa. Part II: Model simulations and coupled model forecasts. *J. Climate,* **22**, 3802-3818.

Mason S J, L Goddard, N E Graham, E Yulaeva, L Sun and P A Arkin, 1999: The IRI seasonal climate prediction system and the 1997/98 El Niño event, *Bull. Amer. Meteor. Soc.,* **80,** 1853-1873.

Reynolds, R. W., N. A. Rayner, T. M. Smith, D. C. Stokes and W. Wang, 2002, An improved in situ and satellite SST analysis for climate. *J. Climate*, **15,** 1609-1625.

Ropelewski, C. F and M.S. Halpert, 1987: Global and regional scale precipitation associated with El Niño/Southern Oscillation, *Mon. Wea. Rev.*, **115,** 1606-1626.

—, and —, 1989: Precipitation pattern associated with the high index phase of the Southern Oscillation. *J. Climate*, **2,** 268-284.

Saha S., and Coauthors 2006: The NCEP climate forecast system. *J. Climate,* **19,** 3483-3517.

Shukla J., 1998: Predictability in the midst of chaos: A scientific basis for climate forecasting. *Science,* **282,** 728–731.

Sooraj, K.P., H. Annamalai, A. Kumar and H. Wang, 2011: A comprehensive assessment of CFS seasonal forecasts over the Tropics. *Weather and Forecasting*, (in press).

Stockdale, T.N., D.L. Anderson, M.A. Balmaseda, F.D. Reyes, L. Ferranti, K. Mogensen, T.N. Palmer, F. Molteni, and F. Vitart, 2011: ECMWF seasonal forecast system 3 and its prediction of sea surface temperature. Clim, Dyn., DOI 10.1007/s00382-010-0947-3.

Su., H and J. D. Neelin., 2002: Teleconnection mechanism for tropical Pacific descent anomalies during El Niño, *J. Atmos. Sci.*, **57,** 3767-3781.

Watanabe, M., and F. F. Jin 2003: A moist linear baroclinic model: Coupled dynamical convective response to El Niño, *J Climate*, **16,** 1121-1139.

Xie, P., and P. A. Arkin, 1996, Analysis of global monthly precipitation using gauge observations, satellite estimates and numerical model predictions. *J Climate*, **9,** 840-858

Xie S –P, Annamalai H, Schott F, McCreary JP Jr, 2002: Structure and mechanisms of South Indian Ocean climate variability. *J Climate*, **15,** 864-87.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Toward a Framework for Incorporating MJO and ENSO Information into CPC Probabilistic Extended Range Forecasts

Nat Johnson[1], Emily Riddle[2], Marshall Stoner[3],
Steven Feldstein[4], Dan Collins[3], Michelle L'Heureux[3]

[1]*IPRC, University of Hawaii at Manoa*
[2] *Wyle Information Systems, Climate Prediction Center, NCEP/NWS/NOAA*
[2] *Climate Prediction Center, NCEP/NWS/NOAA*
[4] *Pennsylvania State University*

1. Introduction

Numerous studies have discussed the role of the El Niño/Southern Oscillation (ENSO) in providing a source of skill in climate prediction on seasonal timescales across much of the globe, particularly over the Pacific/North America region. More recently, studies have shown that the dominant form of variability in the tropics on intraseasonal timescales, the Madden-Julian Oscillation (MJO), also impacts patterns of intraseasonal variability in the northern hemisphere extratropics, providing a possible source of predictive skill over this region on intraseasonal timescales. By modulating tropical convection, the MJO can initiate poleward propagating Rossby waves that impact extratropical weather patterns and influence the leading patterns of low-frequency northern hemisphere variability, the Arctic Oscillation (AO) and Pacific North America pattern (PNA) (*e.g.*, Higgins and Mo 1997, L'Heureux and Higgin 2008, Cassou 2008, Mori and Watanabe 2008, Lin *et al.* 2009, Johnson and Feldstein 2010). Through these mechanisms, the MJO may provide some degree of enhanced predictability for precipitation and temperature in the northern hemisphere extratropics, especially during the winter months at extended range timescales (~10 – 30 days) (*e.g.*, Cassou 2008, Vitart and Molteni 2010, Lin and Brunet 2010, Jones *et al* 2011).

One challenge for forecasters at NOAA Climate Prediction Center (CPC) is finding a way to incorporate these known relationships among the MJO, ENSO, and the large-scale atmospheric circulation into existing extended-range forecast products for the one-to-four week time period. In an effort to provide a framework for incorporating these influences, we present preliminary results that demonstrate a strong influence of the MJO and ENSO on the frequency of winter geopotential height patterns over the PNA region. The purpose of this study is 1) to explore when and for how long tropical MJO activity impacts commonly occurring intraseasonal climate patterns over North America and the surrounding oceans; and 2) to examine how these impacts change during different phases of ENSO. In addition, we discuss the potential for a Bayesian framework to combine this information with dynamical model forecast performance to generate enhanced extended-range forecasts in the one-to-four week time period.

2. Data and methods

We characterize the continuum of wintertime 500 hPa geopotential height patterns in the PNA region with the use of k-means cluster analysis (*e.g.*, Michelangeli *et al.* 1995, Johnson and Feldstein 2010). The method of k-means clustering essentially partitions a potentially large, high-dimensional dataset to a smaller number (K) of representative clusters. Each seven-day geopotential height field (described below) is assigned to a best-matching cluster, and each cluster centroid represents the mean of all height fields assigned to the cluster. Because the method essentially maximizes the similarity between the seven-day height fields and their corresponding cluster centroids, we have reasonable assurance that the resulting centroids correspond with physical, representative patterns.

The cluster analysis is performed on 500 hPa height anomalies from the NCEP/NCAR reanalysis at 2.5º x 2.5º resolution. Only days during the winter months (DJFM) are considered over years ranging from January

Correspondence to: Nat Johnson, International Pacific Research Center (IPRC), University of Hawaii at Manoa, Honolulu, HI 96822; E-mail: natj@hawaii.edu.

1979 to March 2011. This time period is chosen to match the dates when outgoing longwave radiation (OLR) data are also available to compute the MJO index (discussed below). The domain for the cluster analysis ranges from 20ºN to 87.5ºN and from 157.5ºE to 2.5ºW, covering North America and the surrounding ocean basins. This domain was chosen because it encompasses regions with the strongest MJO response, and because our focus is on prediction over North America. Anomalies in the 500-hPa height data are calculated with respect to the daily 1981-2010 climatology, and smoothed by a flat seven-day filter. The cluster analysis is performed on the smoothed daily height anomalies. This smoothing ensures that the cluster analysis focuses on large scale features in the height field and matches the averaging timescale of extended range week-2, week-3 and week-4 forecasts.

To examine how the MJO affects the cluster occurrence frequencies, we must identify periods when the MJO is active, and summarize the spatial location and propagation of MJO during these periods. To do this, we use the Wheeler-Hendon MJO index (Wheeler and Hendon 2004) as provided by the Australian Bureau of Meteorology. The index is derived from the leading two empirical orthogonal functions (EOFs) of three combined fields: tropical OLR, equatorial zonal wind at 850 hPa, and equatorial zonal wind at 200 hPa. Based on the values of these leading EOFS, the Wheeler-Hendon (WH) index traces through eight phases as the MJO signature propagates eastward. Between phase 2 and phase 6, for example, a convectively active region propagates from the western Indian Ocean across the maritime continent and into the western Pacific. The OLR and zonal wind signatures of each phase of the MJO are referenced in a number of previous papers (*e.g.*, Wheeler and Hendon 2004, Cassou *et al.* 2008, Johnson and Feldstein 2010) and can be viewed on the CPC website, http://www.cpc.ncep.noaa.gov/products/precip/CWlink/MJO/Composites/Tropical/.

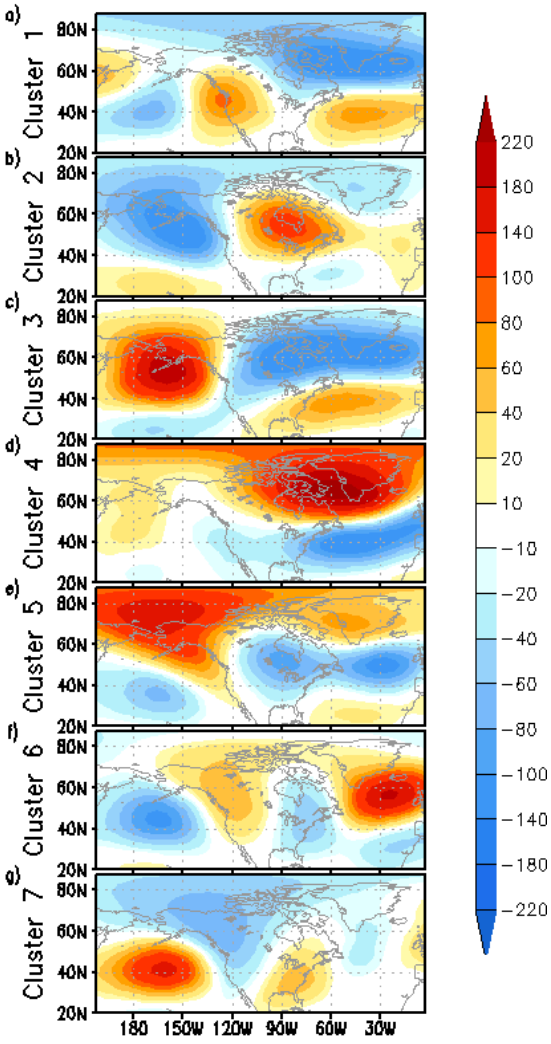


Fig. 1 K-means cluster centroid patterns of 500 hPa geopotential height anomalies (*m*) over the extended North America region.

Following L'Heureux and Higgins (2008) we identify active MJO events based on a pentad-averaged version of the WH index. An MJO episode is identified when the following requirements are met for at least six consecutive pentads: 1) The index amplitude is greater than one and 2) The index phase progresses in a counter-clockwise direction without reversing directions or stalling in a particular phase for more than four pentads. El Niño, La Niña and neutral days are classified by month based on the central month in the Oceanic Niño Index (ONI) definition provided on the CPC website.

3. Results

3.1 Geopotential height cluster patterns

In Figure 1 we show the seven 500 hPa geopotential height cluster patterns that represent the continuum of dominant wintertime teleconnection patterns in the Pacific/North America region. The choice of seven for *K* is based on an optimal volume index ratio, which essentially signifies the number of cluster patterns that efficiently cover the hyperspace without excessive cluster overlap, as described more thoroughly in Riddle *et al.* (2011). Overall, the results we present, however, are not sensitive to the precise choice of *K*. The seven cluster patterns bear a strong resemblance to well known teleconnection patterns such as the positive (clusters

2 and 6) and negative (clusters 3 and 7) phases of the PNA as well as the positive (clusters 1 and 3) and negative (clusters 4 and 6) phases of the North Atlantic Oscillation (NAO)/AO.

The remainder of this paper focuses on clusters 4, 6 and 7 because, unlike the other four clusters, these three clusters are significantly influenced by MJO activity, as discussed in the next section. Clusters 4, 6, and 7 are associated with substantial upper tropospheric zonal wind, precipitation, and surface temperature anomalies over regions of the continental United States (Figure 2). In particular, the negative phase AO-like cluster 4 is associated with a southward shift of the North American jet exit region, and a merging of this jet with the entrance region of the North African jet (Fig. 2a). In association with this southward shift, cluster 4 features widespread cold anomalies across the eastern and mid-western United States (Fig. 2g) and a southward shift in the storm track that is reflected in the precipitation anomalies (Fig. 2d).

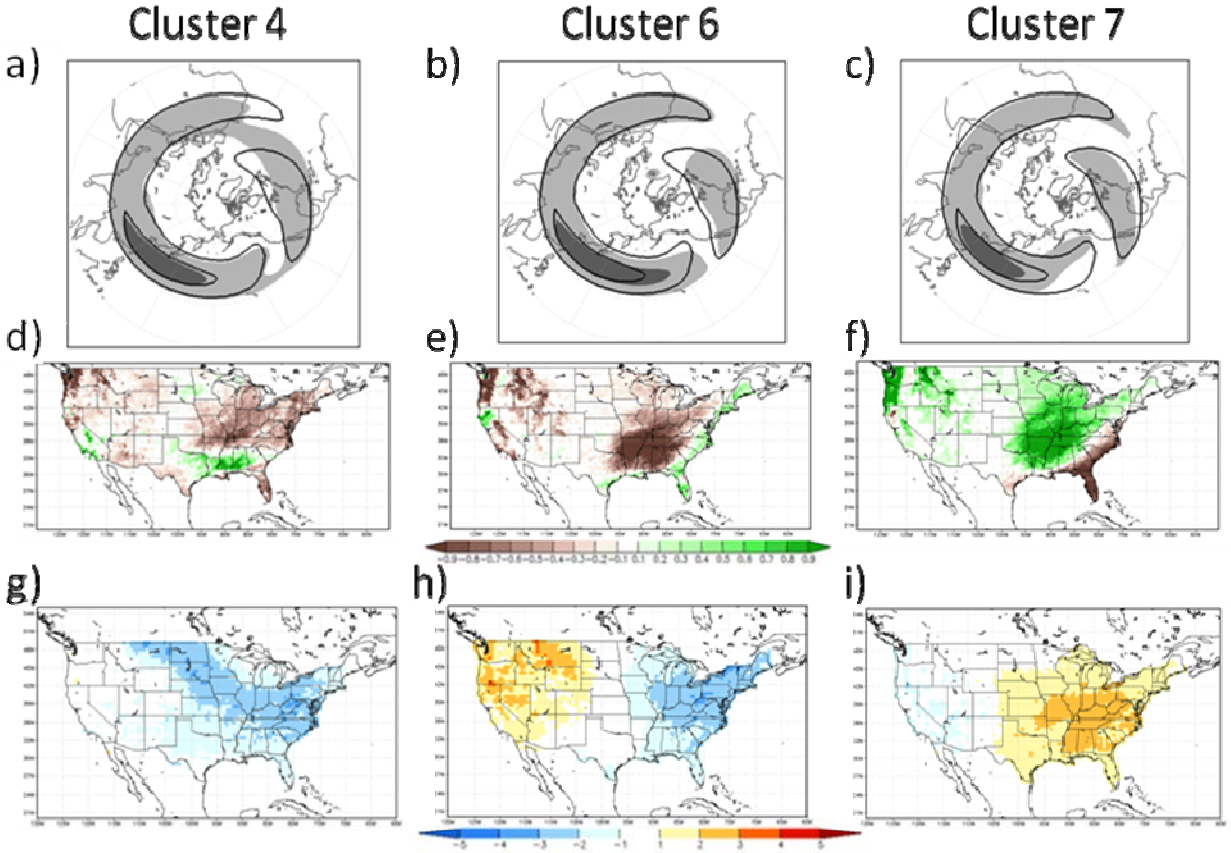


Fig 2 (a) – (c) Shading shows 25 m/s (light grey) and 50 m/s (dark grey) contours for a composite of 200 hPa zonal winds for all days in Clusters 4, 6, and 7 compared with the same two contours for the winter (DJFM) climatology (solid lines). (d) – (i) Composite of (d,e,f) precipitation and (g,h,i) temperature anomalies over the United States for all winter days in Clusters 4, 6, and 7. Zonal wind data are from the NCEP reanalysis, surface temperature composites are derived from the gridded daily cooperative dataset of Janowiak *et al.* (1999), and the precipitation composites are derived from the Climate Prediction Center Merged Analysis of Precipitation (CMAP) (Xie and Arkin 1997).

Over the Pacific, the positive phase PNA-like cluster 6 is associated with a strengthening and southward shift of the jet exit region of the East Asian jet (Fig. 2b). Although both cluster 6 and 4 are associated with negative AO index values, the two clusters have very different signatures in terms of perturbations to jets over the North Atlantic. Cluster 4 is associated with an extension and southward shift of the jet exit region in the Atlantic whereas cluster 6 is associated mainly with a retraction of the North American jet over the Atlantic. Rainfall anomalies over the continental United States for cluster 6 are positive along the eastern coastal US and negative over an extensive region in the interior southeastern states and the Midwest (Fig. 2e). Except for

a wet region in northern California, dry anomalies are also observed over most of the west coast and western mountain states. Temperature anomaly composites are negative over the eastern US and positive over the western and northern mountain states in association with the PNA ridge-trough pattern (Fig. 2h).

The negative phase PNA-like cluster 7 composites (Fig. 2c, f, i) largely oppose those of cluster 6. The biggest asymmetry is seen in the temperature composites: whereas cluster 6 is associated with widespread warm anomalies over the western states (Fig. 2h), cluster 7 features only weak anomalies in the west (Fig. 2i).

3.2 Influence of ENSO and the MJO on cluster pattern frequency of occurrence

To examine how the frequency of cluster occurrence is modulated in the days and weeks following an MJO event and during different phases of ENSO, we follow a very similar approach to that of Cassou (2008). Like Cassou, we examine how the frequency of a cluster occurring under certain external conditions *E* (*e.g.*, 7 days after the MJO is active in phase 1) is elevated or suppressed with respect to the cluster's climatological frequency of occurrence over all 3962 days. The percent change in frequency, *C*, is a function of the external conditions, *E*, and the cluster number *i*:

$$C(i, E) = 100 * \left[\frac{\left(\frac{N_{i,E}}{N_E} - \frac{N_i}{N_T} \right)}{\frac{N_i}{N_T}} \right] \quad (1)$$

where N_T is 3962, the total number of days in the study, N_i is the number of times in the study that cluster *i* occurs, N_E is the number of days when the external conditions, *E*, are met, and $N_{i,E}$ is the number of times that cluster *i* occurs under the conditions *E*. The percentage $C(i, E)$ is equal to 100 if cluster *i* occurs twice as frequently under the conditions *E* as it does in the full record, and is equal to -100 if the cluster never occurs under the conditions *E*. *C* is calculated for a range of external conditions, *E*, including an active MJO during each of the 8 phases and at lead times ranging from zero to 40 days, and for La Niña and El Niño MJO days only. In all of these cases, the full reference climatology is always used for comparison.

Figure 3 shows *C* as a function of lag with respect to MJO phase for clusters 4, 6, and 7. For each cluster pattern in Fig. 3, we consider all MJO episodes and MJO episodes during El Niño and La Niña episodes only. Statistical significance of *C* is assessed with a Monte Carlo test whereby 10,000 synthetic, first-order Markov chain cluster pattern time series are generated with transition probabilities that follow the observed transition probabilities, as discussed more thoroughly in Riddle et al. (2011). Moreover, global significance, which tests against the null hypothesis that the MJO has no overall effect on cluster frequency of occurrence, is assessed by controlling the "false discovery rate" (Benjamini and Hochberg 1995, Wilks 2006). Again, further details can be found in Riddle *et al.* (2011).

Consistent with previous studies, Figure 3 reveals that the MJO exerts a significant influence on the dominant teleconnection patterns of the Pacific/North America region over lags of a few weeks. The occurrence frequency of cluster 4, which resembles a negative AO, is elevated significantly with respect to climatology following active MJO episodes in phases 6 and 7 (Fig. 3a), which are indicative of enhanced convection and upper level divergence over the eastern Pacific Ocean, and suppressed convection and upper level convergence over the western Pacific and maritime continent. The largest positive anomalies in these frequencies occur approximately 20-25 days after MJO phase 6, and approximately 8-13 days after phase 7. Under these conditions Cluster 4 occurs between 2 and 2.5 times as likely as it is in the overall climatology. Though weaker, significant anomalies are observed as far out as 40 days after an occurrence of MJO phase 5. The most significant suppression of Cluster 4 frequencies occurs approximately 24-28 days after phase 2 of the MJO and 18-22 days after phase 3 of the MJO. The phasing of these responses is expected and consistent with an MJO signal propagating from phase 1 through phase 8. The results shown in Figure 3a are consistent with findings from L'Heureux and Higgins (2008), Cassou (2008), Roundy *et al.* (2010), and others who have noted an increase in negative AO events approximately 10-20 days after the occurrence of an MJO phase 6-7 event. Moreover, Fig. 3a shows that this enhancement is more pronounced during El Niño episodes relative

to La Niña episodes, which is consistent with findings that ENSO can modulate the response to the MJO (Schrage *et al.* 1999, Roundy *et al.* 2010, Moon *et al.* 2011).

Figure 3b shows that the frequencies of cluster 6 are elevated significantly with respect to climatology following active MJO episodes in phases 5 and 6, which are indicative of convection and upper level divergence over the central Pacific, and suppressed convection and upper level convergence over the maritime continent and eastern Indian Ocean. The strongest positive anomalies occur 23-26 days after MJO phase 5, and 12-14 days after MJO phase 6. These results are consistent with previous studies that link an increase in

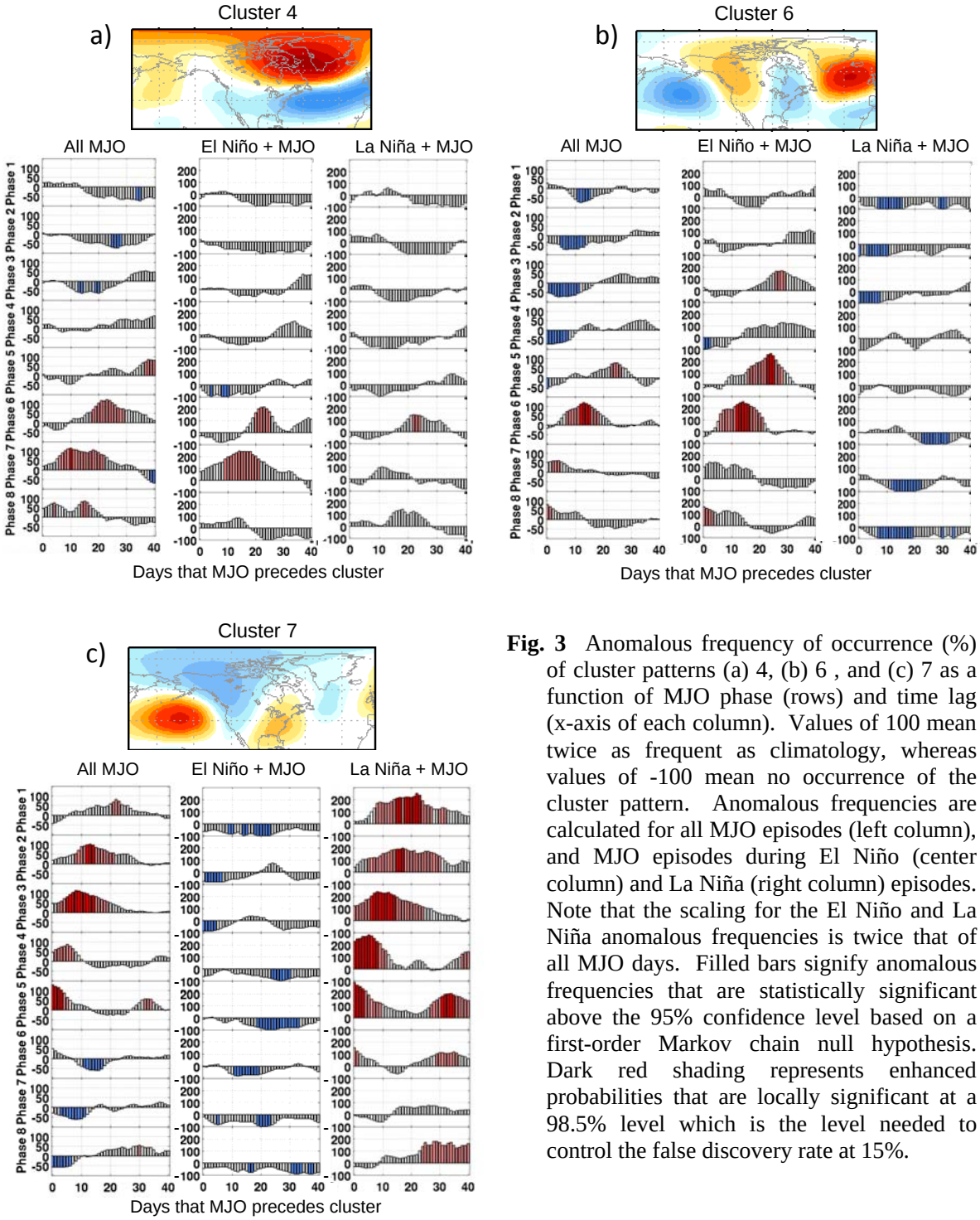


Fig. 3 Anomalous frequency of occurrence (%) of cluster patterns (a) 4, (b) 6 , and (c) 7 as a function of MJO phase (rows) and time lag (x-axis of each column). Values of 100 mean twice as frequent as climatology, whereas values of -100 mean no occurrence of the cluster pattern. Anomalous frequencies are calculated for all MJO episodes (left column), and MJO episodes during El Niño (center column) and La Niña (right column) episodes. Note that the scaling for the El Niño and La Niña anomalous frequencies is twice that of all MJO days. Filled bars signify anomalous frequencies that are statistically significant above the 95% confidence level based on a first-order Markov chain null hypothesis. Dark red shading represents enhanced probabilities that are locally significant at a 98.5% level which is the level needed to control the false discovery rate at 15%.

convection over the central Pacific with a strenthening of the local Hadley circulation, an extension of the East Asian jet over the Pacific, and the development of a cyclonic circulation anomaly over the northern Pacific. Higgins and Mo (1997) and Hoskins and Karoly (1981) and others show that positive anomalies in the PNA occur approximately 10 days following MJO phase 6, in line with the results presented here. Consistent with this finding, the anomalous frequencies of the negative PNA-like cluster 7 (Fig. 3c) are generally opposite to those of cluster 6. The strongest changes in frequency occur simultaneously with MJO phase 5, 3-8 days after phase 4, 8-12 days after phase 3, and 12-16 days after phase 2. These are consistent with studies that suggest that negative PNA anomalies occur approximately 10 days after phase 3, which is indicative of suppressed convection and upper level convergence over the central Pacific. Figures 3b and c reveal that the response of clusters 6 and 7 to the MJO is substantially altered by ENSO. In fact, the enhanced probabilities of cluster 6 (cluster 7) following phase 6 (phase 2) of the MJO are completely absent in La Niña (El Niño) years. This may be expected given that convection anomalies over the Pacific are in phase between El Niño (La Niña) and phase 6 (phase 2) of the MJO.

4. Future work

The preceding analysis demonstrates that the MJO significantly impacts atmospheric circulation patterns with strong surface temperature and precipitation signatures in the continental United States for lead times generally between one and four weeks. An important question remains: how can forecasters at NOAA CPC use this information to enhance extended range forecasts? A suitable framework should incorporate both the expected MJO impacts and the information provided by dynamical model forecast guidance. We propose to merge these two sources of information in a framework based on Bayes' theorem. In the present context, we informally may express Bayes' theorem by

$$\mathrm{P}(\theta_j | f_j) \propto P(\theta_j)\, P(f_j|\theta_j), \qquad (2)$$

where $\mathrm{P}(\theta_j | f_j)$ is the probability of the occurrence of cluster pattern j, given the model forecast of cluster pattern $_j$ in the time period of interest. Bayes' theorem states that the posterior probability, $P(\theta_j|f_j)$, which reflects the forecaster's confidence in the occurrence of cluster j, is proportional to the prior probability $P(\theta_j)$ multiplied by the conditional probability $P(f_j|\theta_j)$. In our proposed application, the expected, combined influence of the MJO and ENSO provides the forecaster with prior information (*i.e.*, before viewing the patterns forecast by the dynamical models) on the occurrence of each cluster pattern that can be incorporated into $P(\theta_j)$. In addition, archived model hindcasts may provide information on model forecast performance that can be used to construct $P(f_j|\theta_j)$. Therefore, Bayes' theorem provides a basic framework with which both the expected MJO/ENSO influence can be combined with dynamical model forecast performance to generate probabilities of cluster pattern occurrence in the forecast period of interest. These posterior probabilities then may be used to weight model forecast fields in a multi-model ensemble forecast. Current efforts are focused on converting this informal framework into a formal Bayesian forecast system for extended range forecasts in the one- to four-week time period.

References

Benjanini, Y., and Y. Hochberg, 1995: Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B (Methodological),* **57**, 289-300.

Cassou, C., 2008: Intraseasonal interaction between the Madden-Julian Oscillation and the North Atlantic Oscillation. *Nature,* **455**, 523-527.

Higgins, R. W, and K. C Mo, 1997: Persistent North Pacific circulation anomalies and the tropical intraseasonal oscillation. *J. Climate,* **10**, 223–244.

Hoskins, B.J., and D. J. Karoly, 1981: The steady linear response of a spherical atmosphere to thermal and orographic forcing. *J. Atmos. Sci.*, **38**, 1179-1196.

Janowiak, J. E, G. D Bell, and M. Chelliah. 1999: *A gridded data base of daily temperature maxima and minima for the conterminous United States: 1948-1993.* U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service (Washington, D.C.), 49 pp.

Johnson, N. C, and S. B. Feldstein. 2010. The continuum of North Pacific sea level pressure patterns: Intraseasonal, interannual, and interdecadal variability. *J. Climate,* **23**, 851-867.

Jones, C., J. Gottschalck, L. M. V. Carvalho, and W. Higgins, 2011: Influence of the Madden-Julian Oscillation on forecasts of extreme precipitation in the contiguous United States. *Mon. Wea. Rev.,* **139**, 332-350.

L'Heureux, M. L., and R. W. Higgins, 2008: Boreal winter links between the Madden-Julian Oscillation and the Arctic oscillation. *J. Climate,* **21**, 3040-3050.

Lin, H., G Brunet, and J. Derome, 2009: An observed connection between the North Atlantic Oscillation and the Madden-Julian Oscillation. *J. Climate,* **22**, 364-380.

Lin, H., and G. Brunet, 2010: The influence of the Madden–Julian Oscillation on Canadian wintertime surface air temperature. *Mon. Wea. Rev.*, **137**, 2250-2262.

Michelangeli, P.- A, R. Vautard, and B. Legras, 1995: Weather regimes: Recurrence and quasi stationarity. *J. Atmos. Sci.,* **52**, 1237–1256.

Moon, J-Y., B. Wang, and K-J. Ha, 2011. ENSO regulation of MJO teleconnection. *Climate Dyn.,* **37**, 1133-1149.

Mori, M., and M. Watanabe, 2008: The growth and triggering mechanisms of the PNA: A MJO-PNA coherence. *J. Meteor. Soc. Japan*, **86**, 213-236.

Riddle, E., M. Stoner, D. Collins, S. Feldstein, N. C. Johnson, and M. L'Heureux, 2011: The impact of ENSO and the MJO on wintertime intraseasonal climate patterns over the Pacific/North America region. To be submitted to *J. Climate*.

Roundy, P. E., K. MacRitchie, J. Asuma, and T. Melino, 2010: Modulation of the global atmospheric circulation by combined activity in the Madden-Julian Oscillation and the El Niño-Southern Oscillation during boreal winter. *J. Climate,* **23**, 4045-4059.

Schrage, J. M, D. G. Vincent, and A. H. Fink, 1999: Modulation of intraseasonal (25–70 day) processes by the superimposed ENSO cycle across the Pacific basin. *Meteor. Atmos. Phys.,* **70**, 15–27.

Vitart, F., and F. Molteni, 2010: Simulation of the Madden–Julian Oscillation and its teleconnections in the ECMWF forecast system. *Quart. J. Roy. Meteor. Soc.,* **136**, 842-855.

Wilks, D. S., 2006: On 'field significance' and the false discovery rate. *J. Appl. Meteor. Climatol.,* **45**, 1181-1189.

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.

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.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Spatial-Intensity Variations in Extreme Precipitation in the Contiguous United States and the Madden-Julian Oscillation

Charles Jones[1] and Leila M. V. Carvalho[1,2]

[1] *Earth Research Institute, University of California, Santa Barbara, California*
[2] *Department of Geography, University of California, Santa Barbara, California*

1. Introduction

The Madden–Julian Oscillation (MJO) is the most prominent mode of tropical intraseasonal variability in the climate system (Madden and Julian 1994; Lau and Waliser 2005; Zhang 2005; Jones and Carvalho 2006; Jones 2009; Jones and Carvalho 2011a). Important linkages have been found between the MJO and precipitation variability including occurrences of extreme events (Jones *et al.* 2004; Donald *et al.* 2006). Significant signals have been shown over the contiguous United States (CONUS) (Mo and Higgins 1998a, 1998b; Mo 1999; Higgins *et al.* 2000a; Jones 2000; Bond and Vecchi 2003; Becker *et al.* 2011; Ralph *et al.* 2011; Zhou *et al.* 2011).

The present study focuses on the MJO and occurrences of precipitation over the CONUS during boreal winter. This problem is investigated by considering two joint properties: intensity and spatial extent of extreme precipitation. The following questions are investigated: 1) What is the probability of extreme precipitation over the CONUS when the MJO is active? 2) Does the spatial-intensity probability of extreme precipitation associated with the MJO significantly vary with ENSO? 3) Do probabilities of spatial-intensity characteristics of extreme precipitation over the CONUS vary with MJO phases? 4) Are large amplitudes of the MJO associated with high probabilities of extreme precipitation over the CONUS?

2. Data

Daily gridded precipitation from the NOAA Climate Prediction Center (CPC) unified gauge (CPC-uni) (Higgins *et al.* 2000b; Chen et al. 2008) is used to investigate the variability of extreme events. Data with 0.5º in latitude by longitude are used for the period 1 January-31 December, 1979-2010. Figure 1 shows the mean daily precipitation over the CONUS during boreal winter seasons defined from 1 November to 31 March (1979-2010). In the western CONUS, mean precipitation exceeds 2 mm day^{-1} and shows large gradients associated with topographic features over the Coastal Ranges, Sierra Nevada and Rocky Mountains. In contrast, the mean winter precipitation in the eastern CONUS shows smooth horizontal gradients and maximizes over the southeastern States. To make the presentation manageable, the CONUS is divided into six sectors: southwest (SW), central-south (CS), southeast (SE), northwest (NW), central north (CN) and northeast (NE). The spatial-intensity characteristics

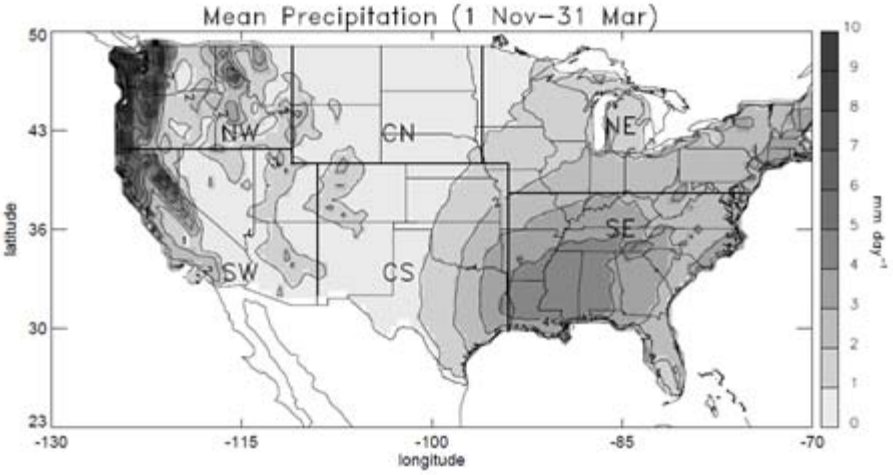


Fig. 1 Mean precipitation during 1 November - 31 March, 1979-2010 (1 mm day^{-1} interval). Thick solid lines indicate six sectors dividing the continuous United States (CONUS).

Correspondence to: Charles Jones, Earth Research Institute, University of California, Santa Barbara, CA 93106.
Email: cjones@eri.ucsb.edu

of extreme precipitation, their probabilities of occurrence and the importance of the MJO are aggregated in each sector over the CONUS.

To identify MJO events, daily averages of zonal wind components at 850-hPa (U850) and 200-hPa (U200) from the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalysis (Kalnay et al. 1996) are used (1 January-31 December 1979-2010). MJO events are identified according to the method discussed in Jones (2009) and Jones and Carvalho (2011b).

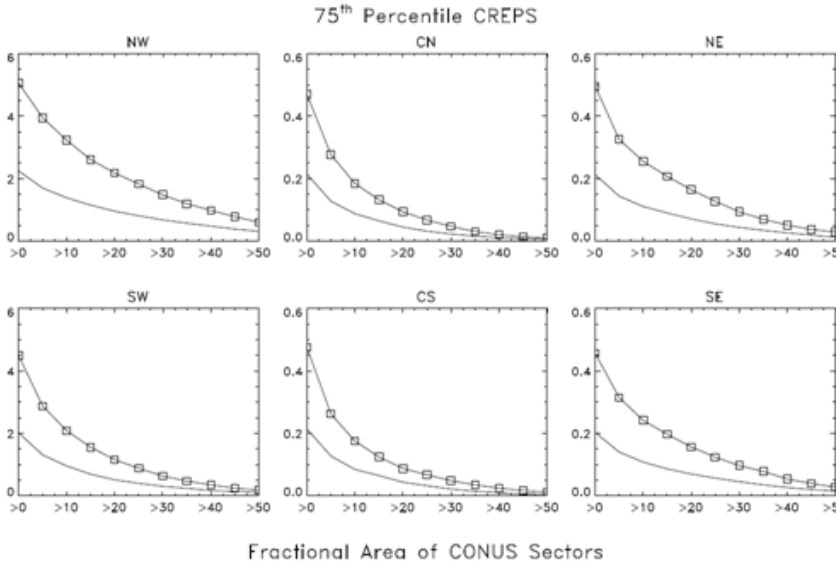


Fig. 2 Joint Probabilities of 75th percentile CREP during active and inactive MJO days. Panels are for each sector in the CONUS. Solid lines with squares show joint probabilities that the fractional area of the sector due to 75th percentile CREPs exceeds a given threshold (horizontal axis; percentages) and the NJO being active (in any phase). Solid lines show joint probabilities that the fractional area of the sector due to 75th percentile CREP exceeds a given threshold and the MJO being inactive.

The spatial-intensity variability of extreme precipitation in the CONUS is analyzed by first identifying extreme events in daily gridded precipitation. Two thresholds of daily precipitation intensity are used: exceeding the 75th and 90th percentiles. In additional, the spatial extent of extreme precipitation is analyzed by considering the frequency distribution of areas. Regions of spatially connected gridpoints in which precipitation exceeds the 75th or 90th percentiles are identified. In summary, two types of *contiguous regions of extreme precipitation* (hereafter CREP) are analyzed. Type I: intensity of precipitation and size exceeds the 75th percentiles of frequency distributions. Type II: intensity and size exceed the 90th percentiles of the frequency distributions.

3. The MJO and probabilities of CREPs

This section presents a quantitative analysis of probabilities of CREP occurrences and relationships with the MJO. To make the presentation more manageable, the results are aggregated for each CONUS sector. We define fractional area as the area of the sector covered by CREPs divided by the total area of the sector. The calculation for each sector considers the area of the CREP contained within the sector. Likewise, intensity is defined as the total precipitation associated with CREPs falling in each sector.

Figure 2 shows joint probabilities that the fractional area in each sector associated with 75th percentile CREPs exceeds specific thresholds (horizontal axis) and the MJO is active (in any phase). Similarly, joint probabilities for fractional area exceedance and inactive MJO days are plotted. Over the NW CONUS, for instance, the probability that the MJO is active and the sector is covered by more than 10% with 75th percentile CREPs is about 0.33. In contrast, the joint probability for the NW sector being covered by more than 10% CREPs during inactive MJO days is ~0.14. As the thresholds of fractional area increase, the joint probabilities decrease. The joint probabilities are significantly higher when the MJO is active than in inactive days. Joint probabilities decrease slower over the NW, NE and SE sectors relative to the CN and CS sectors, probably because precipitation over the central CONUS is not too intense during winter and weather systems quickly move eastward over those regions.

Joint probabilities of precipitation intensity and active (inactive) MJO days associated with 75th extremes are shown in Fig. 3. The joint probabilities are at least twice higher when the MJO is active than during

inactive days. The results above demonstrate quantitatively that the MJO has a substantial influence on both the spatial and intensity characteristics of extreme precipitation over the CONUS during winter.

4. Conclusions

Jones and Carvalho (2011b) investigated the spatial-intensity variability of extreme precipitation over the CONUS during boreal winter and relationships with the MJO. Daily gridded precipitation is used to define two types of contiguous regions of extreme precipitation (CREPs): intensity and spatial extent exceeding the 75th and 90th percentiles of frequency distributions. Extreme precipitation occurs twice more frequent when the MJO is active than inactive. Joint probabilities of fractional area of CONUS sectors when the MJO is active are 2.0-2.5 higher than probabilities during inactive days for both 75th and 90th percentiles CREPs (similarly for intensity of CREPs). Probabilities of fractional area of 75th percentile CREPs when the MJO is active in neutral ENSO are higher than during warm or cold ENSO. Joint probabilities of fractional area during MJO and warm ENSO are higher than MJO and cold ENSO and statistically significant over southern sectors. Results are similar for joint probabilities of intensity exceedance and MJO activity in warm and cold ENSO phases. Proportions of 75th and 90th percentile CREPs for each sector and phase of the MJO are predominantly large when MJO convective signals are over the central Indian Ocean or western Pacific. Probabilities of fractional area of 90th percentile CREPS conditioned on MJO phases, however, do not show clear predominance. This indicates that the MJO is not the sole player in the occurrences of CREPs. Lastly, this study concludes that probabilities of fractional area and intensity of 75th and 90th percentile CREPs in the CONUS do not depend on the amplitude of the MJO.

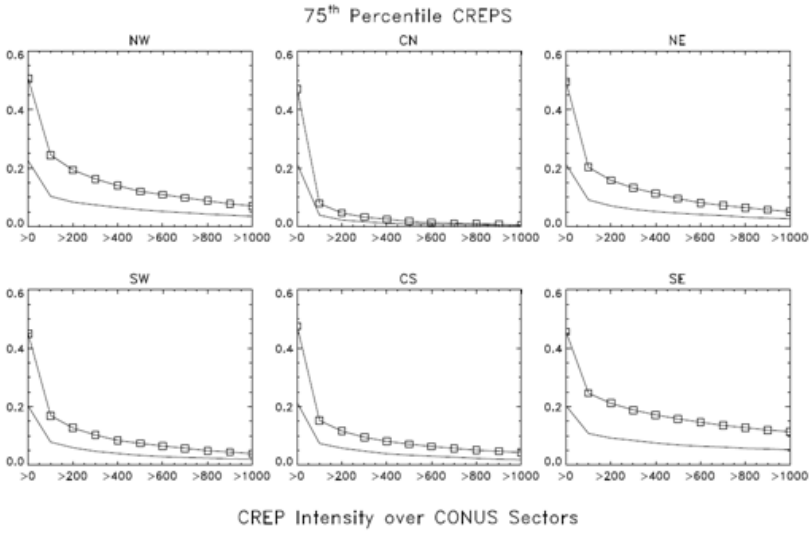


Fig. 3 Joint Probabilities of 75th percentile CREP during active and inactive MJO days. Panels are for each sector in the CONUS. Solid lines with squares show joint probabilities that the precipitation intensity in the sector due to 75th percentile CREPs exceeds a given threshold (horizontal axis; mm day^{-1}) and the MJO being active (in any phase). Solid lines show joint probabilities that the precipitation intensity in the sector exceeds a given threshold and the MJO being inactive.

Acknowledgements. This research was supported by NOAA's Climate Program Office, Climate Test Bed program (NA08OAR4310698) and Climate and Large-scale Dynamics Program from the National Science Foundation (AGS-1053294). NCEP/NCAR Reanalysis and OLR data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA (www.esrl.noaa.gov).

References

Becker, E. J., E. H. Berbery, and R. W. Higgins, 2011: Modulation of Cold-Season U.S. Daily Precipitation by the Madden–Julian Oscillation. *J. Climate*, **24**, 5157-5166.

Bond, N. A., and G. A. Vecchi, 2003: The influence of the Madden-Julian oscillation on precipitation in Oregon and Washington. *Weather and Forecasting*, **18**, 600-613.

Chen, M. Y., W. Shi, P. 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-Atmos*, **113**, D04110, 04111-04113.

Donald, A., and Coauthors, 2006: Near-global impact of the Madden-Julian Oscillation on rainfall. *Geophysical Research Letters*, **33**, L09704, doi:10.1029/2005GL025155.

Higgins, R. W., J. K. E. Schemm, W. Shi, and A. Leetmaa, 2000a: Extreme precipitation events in the western United States related to tropical forcing. *J. Climate*, **13**, 793-820.

Higgins, R. W., W. Shi, E. Yarosh, and R. Joyce, 2000b: Improved United States precipitation quality control system and analysis. *NCEP/Climate Prediction Center ATLAS No. 7,* National Oceanic and Atmospheric Administration, 40 pp.

Jones, C., 2000: Occurrence of extreme precipitation events in California and relationships with the Madden-Julian oscillation. *J. Climate*, **13**, 3576-3587.

——, 2009: A homogeneous stochastic model of the Madden-Julian Oscillation. *J. Climate*, **22**, 3270-3288.

Jones, C., and L. M. V. Carvalho, 2006: Changes in the activity of the Madden-Julian oscillation during 1958-2004. *J. Climate*, **19**, 6353-6370.

——, 2011a: Stochastic simulations of the Madden-Julian Oscillation activity. *Climate Dynamics*, **36**, 229-246, doi:210.1007/s00382-00009-00660-00382.

——, 2011b: Spatiotemporal variations in extreme precipitation in the contiguous United States and the Madden-Julian Oscillation. *J. Climate (in revision)*.

Jones, C., D. E. Waliser, K. M. Lau, and W. Stern, 2004: Global occurrences of extreme precipitation and the Madden-Julian oscillation: Observations and predictability. *J. Climate*, **17**, 4575-4589.

Kalnay, E., and Coauthors, 1996: The NCEP-NCAR 40 Year Reanalysis Project. *Bulletin of the American Meteorological Society*, **77**, 437-471.

Lau, W. K. M., and D. E. Waliser, 2005: *Intraseasonal Variability in the Atmosphere-Ocean Climate System.* Springer, Chichester, UK, 436 pp.

Madden, R. A., and P. R. Julian, 1994: Observations of the 40-50-Day Tropical Oscillation-A Review. *Mon.Wea. Rev.*, **122**, 814-837.

Mo, K. C., 1999: Alternating wet and dry episodes over California and intraseasonal oscillations. *Monthly Weather Review*, **127**, 2759-2776.

Mo, K. C., and R. W. Higgins, 1998a: Tropical influences on California precipitation. *J. Climate*, **11,** 412-430.

——, 1998b: Tropical convection and precipitation regimes in the western United States. *J. Climate*, **11,** 2404-2423.

Ralph, F. M., P. J. Neiman, G. N. Kiladis, K. Weickmann, and D. W. Reynolds, 2011: A Multiscale Observational Case Study of a Pacific Atmospheric River Exhibiting Tropical–Extratropical Connections and a Mesoscale Frontal Wave. *Monthly Weather Review*, **139,** 1169-1189.

Zhang, C. D., 2005: Madden-Julian oscillation. *Reviews of Geophysics*, **43,** 1-36.

Zhou, S., M. L'Heureux, S. Weaver, and A. Kumar, 2011: A composite study of the MJO influence on the surface air temperature and precipitation over the Continental United States. *Climate Dynamics***,** 1-13.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

A GOES Thermal-Based Drought Early Warning Index for NIDIS

Christopher R. Hain[1], Martha C. Anderson[2], Xiwu Zhan[3], Mark Svoboda[4], Brian Wardlow[4], Kingtse Mo[5], John R. Meckalski[6], William P. Kustas[2], and Jesslyn Brown[7]

[1]*Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD*
[2]*USDA-ARS Hydrology and Remote Sensing Lab, Beltsville, MD*
[3]*National Drought Mitigation Center, University of NE-Lincoln, Lincoln, NE*
[4]*NESDIS, NOAA, Camp Springs, MD*
[5]*NCEP Climate Prediction Center, Camp Springs, MD*
[6]*University of Alabama in Huntsville, Huntsville, AL*
[7]*USGS Earth Resources Observations and Science (EROS) Center, Sioux Falls, SD*

1. Introduction

The interpretation of drought signals has proven to be difficult because of a general lack of ground-based "truth" metrics available at continental scales; therefore, forecasters must rely on a convergence of evidence strategy using multiple drought index datasets. Standard indicators currently used in drought monitoring focus on different components of the water budget: precipitation, soil moisture, groundwater, runoff and streamflow. The goal of our NOAA CPO-funded project is to develop a thermal-based drought index based on estimates of the actual to potential evapotranspiration ratio provided by the Atmosphere Land-Exchange Inverse (ALEXI) model (Anderson, *et al.*, 2007a,b; 2011), focusing on the water-use component of the hydrologic cycle.

Current drought indices include precipitation-based analyses (*e.g.*, Standardized Precipitation Index (SPI; McKee *et al.* 1995); the Palmer indices (Palmer 1965)), and satellite-based vegetation/TIR indices (*e.g.*, Vegetation Health Index (VHI; Kogan 1997); VegDRI (Brown *et al.* 2008)), along with soil moisture and ET datasets generated with land-surface models (LSMs) in the National Land Data Assimilation System (NLDAS; Mitchell *et al.* 2003). Each of these index classes has issues: datasets like NLDAS and SPI require precipitation and/or soil texture fields that are difficult to observe/specify accurately over large spatial domains; while empirical TIR-based drought indices currently in use (like the VHI) do not account for important forcings on land-surface temperature (LST) (*e.g*, available energy, atmospheric demand), and can therefore generate spurious drought detections under certain circumstances – particularly at high latitudes (Karnieli *et al.* 2006; Karnieli *et al*. 2010).

In contrast, diagnostic LSMs based on TIR remote sensing of LST, like ALEXI, require no information regarding antecedent precipitation or soil moisture storage capacity - *the current surface moisture status is deduced directly from the remotely sensed radiometric temperature signal*. This results in a seamless implementation over the continent, unaffected by discontinuities in soils and precipitation dataset collected by individual countries. In contrast with the VHI, ALEXI is based on energy balance, so radiation, atmospheric and soil moisture controls are all considered in the interpretation of the LST signal. The TIR remote sensing data used in the Evaporative Stress Index (ESI) also provide information about non-precipitation related moisture inputs to the land-surface system arising from processes such as irrigation, shallow groundwater sources, and lateral flows – processes that must be known a priori and modeled explicitly in prognostic LSMs (such as in NLDAS) but may significantly mitigate drought impacts during local rainfall deficits. And whereas vegetation index (VI) is a relatively slow response variable to moisture deficits, showing decline only after the damage has been done, thermal remote sensing has the potential to provide valuable drought early warning preceding detectable degradation in VIs.

Correspondence to: Christopher R. Hain, Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD; E-mail: Chris.Hain@noaa.gov.

2. Model methodology

The ALEXI surface energy balance model (Anderson *et al.* 1997, 2007a; Fig. 1) was specifically designed to minimize the need for ancillary meteorological data while maintaining a physically realistic representation of land-atmosphere exchange over a wide range in vegetation cover conditions. It is one of few land-surface models designed explicitly to exploit the high temporal resolution afforded by geostationary satellites.

a. Interpretation of the thermal land-surface signature

Surface energy balance models estimate ET by partitioning the energy available at the land surface ($RN - G$, where RN is net radiation and G is the soil heat conduction flux, in Wm^{-2}) into turbulent fluxes of sensible and latent heating (H and λE, respectively, Wm^{-2}):

$$RN - G = H + \lambda E \qquad (1)$$

where λ is the latent heat of vaporization ($J\ kg^{-1}$) and E is ET ($kg\ s^{-1}\ m^{-2}$ or $mm\ s^{-1}$). The land-surface representation in ALEXI model is based on the series version of the two-source energy balance (TSEB) model of Norman *et al.* (1995), which partitions the composite surface radiometric temperature, T_{RAD}, into characteristic soil and canopy temperatures, T_S and T_C, based on the local vegetation cover fraction apparent at the thermal sensor view angle, $f(\theta)$:

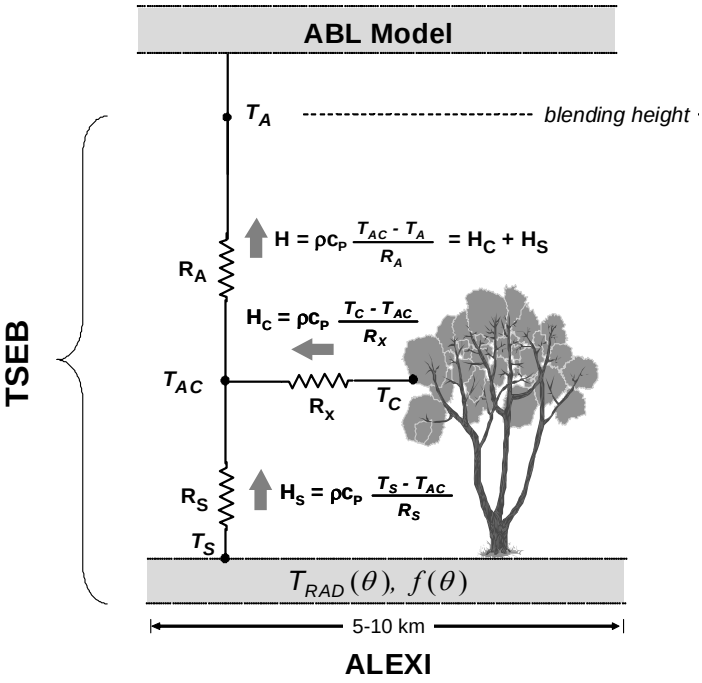


Fig. 1 Schematic diagram representing the ALEXI modeling framework, highlighting fluxes of sensible heat (H) from the soil and canopy (subscripts 's' and 'c') along gradients in temperature (T), and regulated by transport resistances R_A (aerodynamic), R_x (bulk leaf boundary layer) and R_S (soil surface boundary layer).

$$T_{RAD} \approx \{ f(\theta) T_c + [1 - f(\theta)]\ T_s \} \qquad (2)$$

(Fig. 1), where $f(\theta)$ can be related to leaf area index (LAI) using Beer's law. Eq. 2 is a linear approximation to an aggregation of surface radiance values. With information about T_{RAD}, LAI, and radiative forcing, the TSEB evaluates the soil (subscript 's') and the canopy ('c') energy budgets separately, computing system and component fluxes of net radiation ($RN = RN_C + RN_S$), sensible and latent heat ($H = H_C + H_S$ and $\lambda E = \lambda E_C + \lambda E_S$), and soil heat ($G$). Importantly, because angular effects are incorporated into the decomposition of T_{RAD}, the TSEB can accommodate TIR data acquired at off-nadir viewing angles by geostationary satellites. The TSEB has a built-in mechanism for detecting thermal signatures of stress in the soil and canopy. An initial iteration assumes the canopy transpiration (λE_C) is occurring at a potential (non-moisture limited) rate, while the soil evaporation rate (λE_S) is computed as a residual to the system energy budget. If the vegetation is stressed and transpiring at significantly less than the potential rate, λE_C will be overestimated and the residual λE_S will become negative. Condensation onto the soil is unlikely midday on clear days, and therefore $\lambda E_S < 0$ is considered a signature of system stress. Under such circumstances, the λE_C is iteratively down-regulated until $\lambda E_S \sim 0$ (expected under dry conditions).

b. ALEXI Evaporative Stress Index

The Evaporative Stress Index (Anderson *et al.* 2007 a, b; 2011) represents standardized anomalies in the ratio of actual-to-potential ET, f_{PET} = ET/PET, where ET and PET are instantaneous clear-sky estimates at shortly before local noon, retrieved using the ALEXI algorithm. Normalization by PET serves to minimize variability in ET due to seasonal variations in available energy and vegetation cover, further refining focus on the soil moisture signal. Limiting the assessment to clear-sky conditions separates signals of soil moisture

variability from that of cloud climatology. To highlight differences in moisture conditions between years, standardized anomalies in f_{PET} are expressed as a pseudo z-score, normalized to a mean of zero and a standard deviation of one with respect to baseline fields describing "normal" (mean) conditions over the period of record.

c. Model archive generation

Input datasets required by ALEXI are listed in Table 1. The sources of input data are a mix of operational inputs that are available daily (GOES Sounder / insolation and North American Regional Reanalysis (NARR) model output), and MODIS land products (such as LAI and albedo), which are routinely generated but available with a several week time lag. The ALEXI domain covers the CONUS at a spatial resolution of 10 km and a temporal resolution of 1 day.

Data	Purpose	Source	Spatial Resolution	Temporal Resolution
LST	ΔT_{rad}, RN	GOES	10 km	1 hr
LAI	T_{rad} partitioning	MODIS	0.01°	8-day
Insolation	RN	GOES	20 km	1 hr
Longwave radiation	RN	GOES	20 km	1 hr
Albedo	RN	MODIS	0.05°	16-day
Wind Speed	Aerodynamic resistances	NARR	32 km	3 hr
Atmos lapse [dθ/dz]	ABL growth model	NARR	32 km	3 hr
Landcover type	Canopy characteristics	UMD	0.01	fixed

Table 1 Primary inputs used by the current CONUS ALEXI ESI system.

3. Results

An intercomparison study of ALEXI ESI and a suite of drought indices (Table 2) was conducted from 2000 to 2011 over the CONUS, focusing on the primary growing season for most of the United States (April – October). As is the case with ESI, standardized anomalies were computed for each drought index. Temporal and spatial correlations between index anomalies were examined to assess the similarity between drought indices in their ability to rank drought severity and to visualize spatial patterns in index congruity.

Index	Acronym	Type
U.S. Drought Monitor	USDM	Multi-index synthesis
Evaporative Stress Index (X-month composite)	ESI-X	Remote sensing of f_{PET}
Vegetation Health Index	VHI	Remote sensing of LST, VI
Standardized precipitation index (X-month composite)	SPI-X	Precipitation
Palmer Z Index	*Z*	Precipitation + storage
Palmer drought severity index	PDSI	Precipitation + storage
Palmer modified drought index	PMDI	Precipitation + storage
Palmer hydrologic drought index	PHDI	Precipitation + storage
Noah-LDAS evapotranspiration	Noah-ET	Land surface model
Noah-LDAS soil moisture	Noah-SM	Land surface model

Table 2 Drought indices included in the intercomparison study.

Drought features in the USDM classifications are generally reflected in one or more of the other indices but to varying degrees depending on drought type and time scale. Figure 2 shows the seasonal anomalies for a selection of drought indices listed in Table 2 for the study period of 2000 to 2011. In general, ESI reproduces patterns evident in the precipitation-based indices, indicating the value of the LST signal as a surface moisture proxy. For example, the thermal band inputs to ALEXI capture the major drought events in 2002 and 2007, even in the eastern United States, an area in which dense vegetation is dominant during the warm season. This is particularly important because standard soil moisture retrievals based on microwave remote sensing tend to lose sensitivity due to strong attenuation of the surface soil moisture signal by the overlying vegetative canopy. However, in this case, the thermal signal is able to detect vegetation stress related to root zone soil moisture deficits and elevated canopy temperatures.

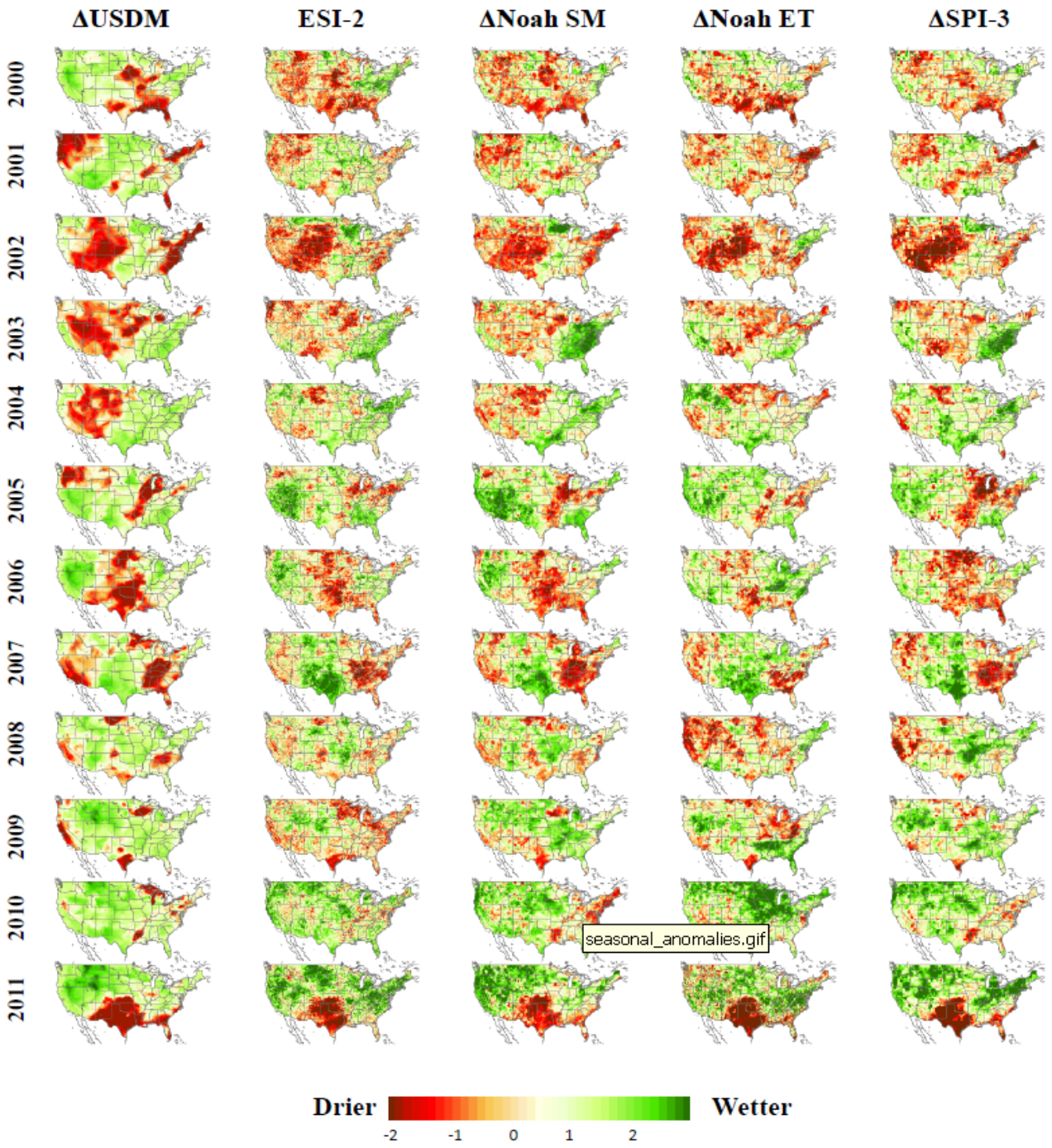


Fig. 2 Seasonal (Apr-Oct) anomalies in US Drought Monitor classes, ESI, Noah soil moisture, Noah evapotranspiration and SPI-3.

Of the products included in the intercomparison, Noah soil moisture anomalies were found to be the most similar to the USDM with respect to temporal correlation (averaged over CONUS). The ALEXI 2-month ESI composite (ESI-2) shows higher average temporal correlations with the USDM than do the precipitation indices of shorter or comparable time scales (Z and SPI-1 to SPI-3). ESI-2 also outperforms Noah ET anomalies and VHI in terms of correlation with USDM. The strongest correlations between ESI and USDM are observed over the Great Plains and in the southeastern United States. These are areas identified by Karnieli *et al.* (2010) where LST and NDVI tend to be anticorrelated, indicating moisture-limiting (as opposed to energy limiting) vegetation growth conditions. In these areas, ET will be most sensitive to

changing root-zone soil moisture condition, and subsequently providing indications of drought. There are also regions where ESI shows reduced correlations with the USDM, e.g. over the Mississippi River basin, where shallow water tables and intensive irrigation tend to decouple ET rates from precipitation to some extent. ESI also shows lower correlations with the USDM over the Everglades in southern Florida, an area which is largely inundated with water over much of the year, and ET variations at the seasonal scale may be more related to climatic variability than to moisture availability. Finally, lower correlations are also found in the northern states where, particularly in early spring, ET is driven more by radiation and climate and is less tightly coupled with moisture conditions.

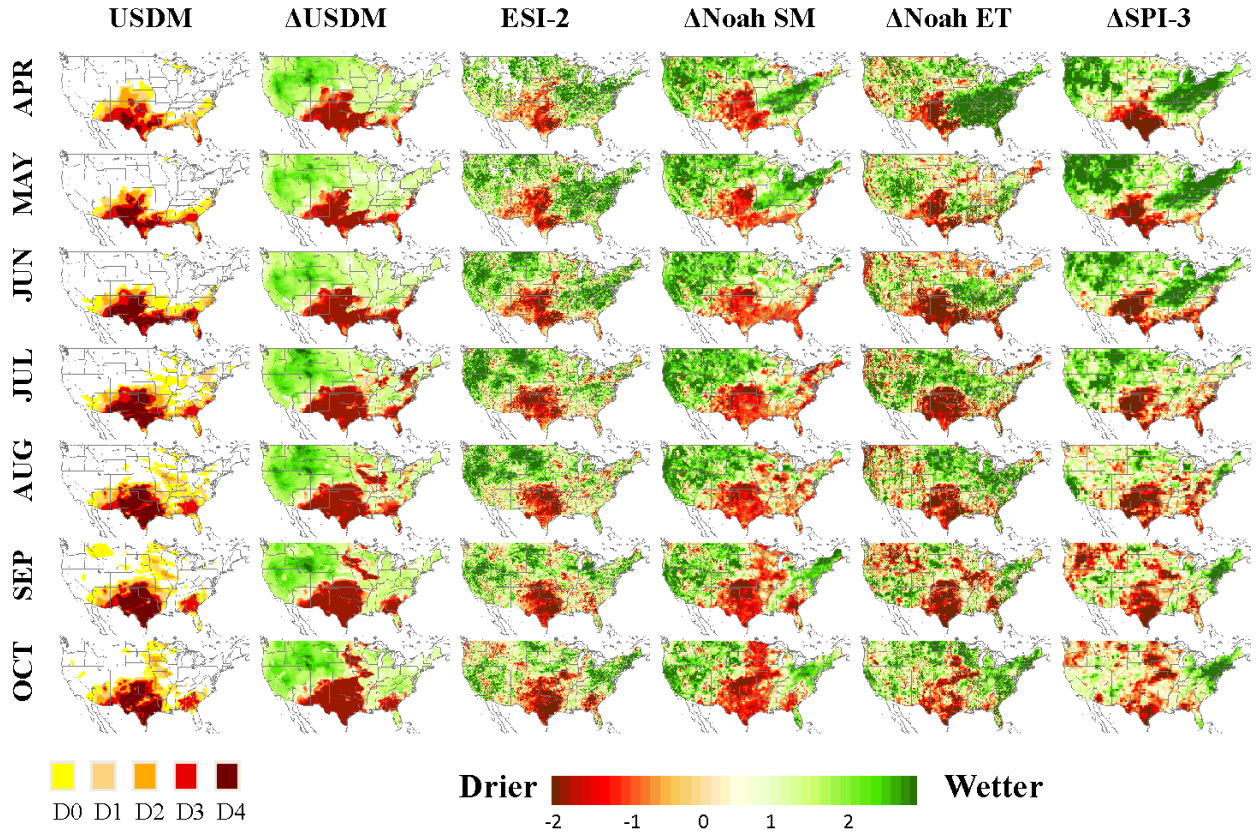


Fig. 3 Monthly USDM drought classification and anomalies in USDM, ESI-2 Noah soil moisture, Noah evapotranspiration, and SPI-3 for 2011.

Spatial anomaly correlations were also computed to assess how well the indices agree on a spatial rather than temporal scale. On a seasonal (April-October) time-scale, all indices show the weakest correlations in 2003 during the long-term hydrologic drought event in the western CONUS, which is captured only by indices with time constants exceeding one year. The highest correlations, among all indices, is found during 2007 where there was a strong contrast in moisture conditions, with extensive drought conditions across the southeast US and anomalously wet conditions in the south central US. At the monthly time-scale, correlations between ESI-2 and the USDM are the weakest during April and May, likely due to poor temporal sampling in the ESI related to increased snow and cloud cover. However, spatial correlation with ESI-2 increases throughout the warm season as ET becomes more closely coupled to moisture conditions.

In contrast, correlations between the short-term precipitation indices (*Z*, SPI-2, SPI-3) and the USDM tend to degrade in August and September. An example of the monthly comparison of spatial anomalies between the USDM and ESI-2 is shown in Figure 3 for the year 2011. All five drought indices show excellent agreement with the extent and severity of the drought conditions across much of the south central US. However, an area of disagreement between ESI-2 and the USDM is evident in southern GA, where the USDM shows more severe drought conditions than is shown in ESI-2 during the period from May to August. ESI-2 does show small areas of dry conditions in May and June, yet conditions return to near normal during

July and August, before a rapid expansion of dry conditions in September through October. The improvement shown in July may be related to several precipitation events which likely lead to a replenishment of root-zone soil moisture and a decrease in vegetation stress. This is to some extent shown in both Noah SM and ET anomalies which showed slight improvement in July and August, although both Noah indices still had stronger negative anomalies than ESI-2 during both months. However, as dry conditions returned in August, root-zone soil moisture deficits likely increased and lead to a rapid appearance of significant dry anomalies in ESI-2 as shown in September and October.

A few caveats must be considered in this case, first, the USDM in not independent of many of the indices listed in Table 2, as they are commonly used in the construction of USDM drought classifications. ALEXI ESI was not used in the USDM classification process during the period of record in this analysis, and therefore is wholly independent. Second, the USDM drought classes incorporate information relevant to different kinds of drought over varying timescales, and we cannot expect a single indicator to agree perfectly with the USDM. For example, socioeconomic drought features in the USDM may indicate increased human demand for water rather than natural hydrological deficits. A more detailed analysis of all the drought index intercomparison results can be found in Anderson *et al.* (2011).

4. Future Work

The final year of our NOAA-CPO project will mainly focus on two core objectives: (1) developing an open interface with end-users at the Climate Prediction Center (CPC) and the National Drought Mitigation Center (NDMC) to provide feedback on the use of ESI maps and (2) automating the ALEXI ESI system to provide weekly ESI maps to the National Integrated Drought Information System (NIDIS) portal and end-users at CPC and NDMC. Figure 4 shows an example of the ALEXI ESI website developed to provide ESI maps to end-users at CPC and NDMC. The website will be open to the entire drought community during the spring of 2012 and through the NIDIS portal at (www.drought.gov).

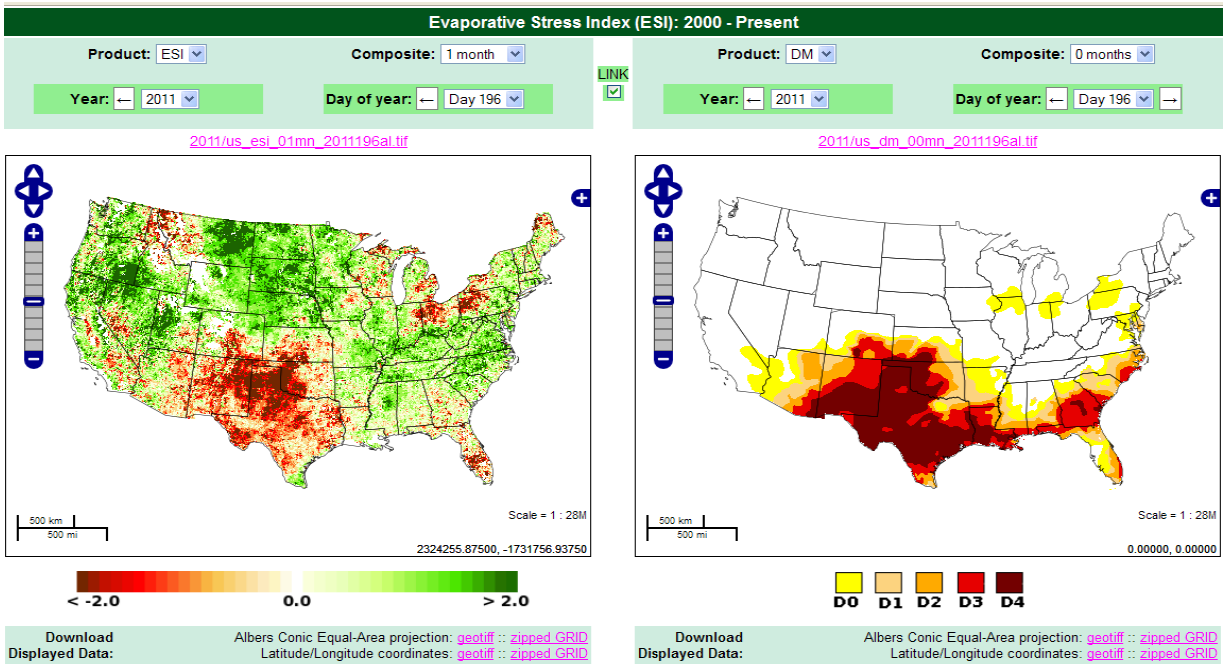


Fig. 4 Screenshot of the ALEXI Evaporative Stress Index (ESI) website developed at the USDA Hydrology and Remote Sensing Lab.

Furthermore, the use of ALEXI as a proxy for soil moisture conditions will be expanded to produce an operational data assimilation system for the optimal assimilation of thermal and microwave soil moisture into the Noah LSM component of the NLDAS towards the goal of improved LSM-based drought monitoring. As mentioned earlier, ALEXI has been shown to perform well over densely vegetation regions such as the southeast US, an area in which microwave retrievals can suffer from significant vegetation-related errors.

Therefore, thermal and microwave retrievals methods have been shown to be quite complementary: thermal methods provide soil moisture information over a wide range of vegetation conditions, while microwave methods provide high temporal information (can retrieve through cloud cover) over areas of low vegetation cover (Hain *et al.* 2001). Finally, although this application of ESI focused solely on the CONUS, ALEXI domains are currently being developed both on a global scale (spatial resolution of 0.25°) and on a regional scale (*e.g.*, Europe, Africa, and Australia; spatial resolution of 3 to 10 km), facilitating production of ESI maps over these domains in the near future.

References

Anderson, M. C., J. M. Norman, G. R. Diak, W. P. Kustas, and J. R. Mecikalski, 1997: A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing. *Remote Sens. Environ.*, **60**, 195-216.

Anderson, M. C., J. M. Norman, J. R. Mecikalski, J. P. Otkin, and W. P. Kustas, 2007a: A climatological study of evapotranspiration and moisture stress across the continental U.S. based on thermal remote sensing: I. Model formulation. *J. Geophys. Res.*, **112**, D10117, doi:10110.11029/12006JD007506.

Anderson, M. C., J. M. Norman, J. R. Mecikalski, J. P. Otkin, and W. P. Kustas, 2007b: A climatological study of evapotranspiration and moisture stress across the continental U.S. based on thermal remote sensing: II. Surface moisture climatology. *J. Geophys. Res.*, **112**, D11112, doi:11110.11029/12006JD007507.

Anderson, M. C., C. R. Hain, B. D. Wardlow, A. Pimstein, J. R. Mecikalski and W. P. Kustas, 2011: Evaluation of drought indices based on thermal remote sensing of evapotranspiration over the continental United States. *J. of Climate*, **24**, 2025-2044.

Brown, J. M., B. D. Wardlow, T. Tadesse, M. J. Hayes and B. C. Reed, 2008: The vegetation drought response index (VegDRI): A new integrated approach for monitoring drought stress in vegetation. *GIScience Remote Sens*., **45**, 16-46.

Hain, C. R., W. T. Crow, J. R. Mecikalski, M. C. Anderson, and T. Holmes (2011), An intercomparison of available soil moisture estimates from thermal infrared and passive microwave remote sensing and land surface modeling. *J. Geophys. Res.*, **116**, D15107, doi:10.1029/2011JD015633.

Karnieli, A., N. Agam, R. T. Pinker, M. C. Anderson, M. L. Imhoff, G. G. Gutman, N. Panov, and A. Goldberg, 2010: Use of NDVI and LST for assessing vegetation health: merits and limitations. *J. Climate,* **23**, 618-633.

Karnieli, A., M. Bayasgalan, Y. Bayarjargal, N. Agam, S. Khudulmur, and C. J. Tucker, 2006: Comments on the use of the Vegetation Health Index over Mongolia. *Int. J. Remote Sensing*, **27**, 2017-2024.

Kogan, F. N., 1997: Global drought watch from space. *Bull. Amer. Meteorol. Soc.*, **78**, 621-636.

McKee, T. B., N. J. Doesken, and J. Kleist, 1995: Drought monitoring with multiple time scales. Paper presented at AMS Ninth conf. on Applied Climatology, Dallas, TX.

Mitchell, K. E. and Coauthors, 2003: The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling sytem. *J. Geophys Res.*, **102**, doi:10.1029/2003JD002823.

Norman, J. M., W. P. Kustas, and K. S. Humes, 1995: A two-source approach for estimating soil and vegetation energy fluxes from observations of directional radiometric surface temperature. *Agric. For. Meteorol.*, **77**, 263-293.

Palmer, W.C., 1965: Meteorological drought. *Research Paper No. 45*, U.S. Weather Bureau, NOAA Library and Information Services Division, Washington, D.C. 20852, 58p.

Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

Seasonal Prediction of Ecosystems, Fire, Carbon Using NCEP/CFS and a Dynamic Vegetation Model

Ning Zeng[1], Eugenia Kalnay[1] and Arun Kumar[2]

[1]*Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD*
[2]*Climate Prediction Center, NOAA/NWS/NCEP, Camp Springs, MD*

1. Introduction

In recent years, many advances have been made in the science and practice of seasonal climate predictions. For example, seasonal climate predictions have attained operational status and have come to rely increasingly more on dynamical prediction models. Such advances notwithstanding, application of seasonal climate outlooks to applications of societal importance has been slow to materialize. The aim of this project is to develop one such application, *i.e., a capability to forecast terrestrial ecosystem productivity and carbon sources and sinks on seasonal-interannual time-scale.* The modeling system is global, but the focus of validation and application will be for North America.

The development of an outlook capability for the ecosystem will rely on several components that have evolved following independent pathways and have reached a state of maturity in their respective domains of interest. The key effort here brings together these modeling and prediction component systems.

The modeling components of the predictive capability include:

a) A dynamic Vegetation-Global-Atmosphere-Soil (VEGAS; Zeng *et al.* 2005, 2008) model with full terrestrial carbon cycle
b) Operational climate forecasts at the Climate Prediction Center and dynamical seasonal forecasts based on the Climate Forecast System (CFS) (both at NCEP)

Specific targets include:

Developing a procedure to specify vegetation and soil initial conditions derived from some form of data assimilation system.

a) Developing procedures to forecast ecosystem and carbon variables using ensemble climate prediction information from CFS
b) Validation of prediction system based on hindcast skill by comparing model predictions against a suite of observed variables such as satellite vegetation index, CO_2 flux measurements, and assimilated carbon fluxes
c) Comparison of the CFS based skill with other baseline estimates of skill for predicting eco-carbon variables, e.g., prediction based on operational CPC forecasts
d) Testing the prediction system in a real-time operational setting, getting feedbacks from a wider community, improving the system.

Deliverable of this project will be a seasonal forecasting system for terrestrial ecosystem productivity and carbon fluxes that later will be transitioned to operations using the Climate Test-Bed (CTB) infrastructure.

2. Hindcast experiments

We have conducted a 25-year hindcast experiment to explore the possibility of seasonal-interannual prediction of terrestrial ecosystem and the global carbon cycle. This has been achieved using a prototype forecasting system in which the dynamic vegetation and terrestrial carbon cycle model VEGAS was forced with the 15-member ensemble climate prediction and lead time up to 9 month from the NCEP/CFS climate forecast system. The results show that the predictability is dominated by the ENSO signal for its major

Correspondence to: Ning Zeng, Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD 20742; E-mail: zeng@umd.edu.

infuence on the tropical and subtropical regions, including the Amazon, Indonesia, western US and central Asia. The hindcasted ecosystem variables and carbon flux show significantly slower decrease in skill compared to the climate forcing, partly due to the memories in land and vegetation processes that filter out the higher frequency noise and sustain the signal. (Examples shown in Figs.1 and 2)

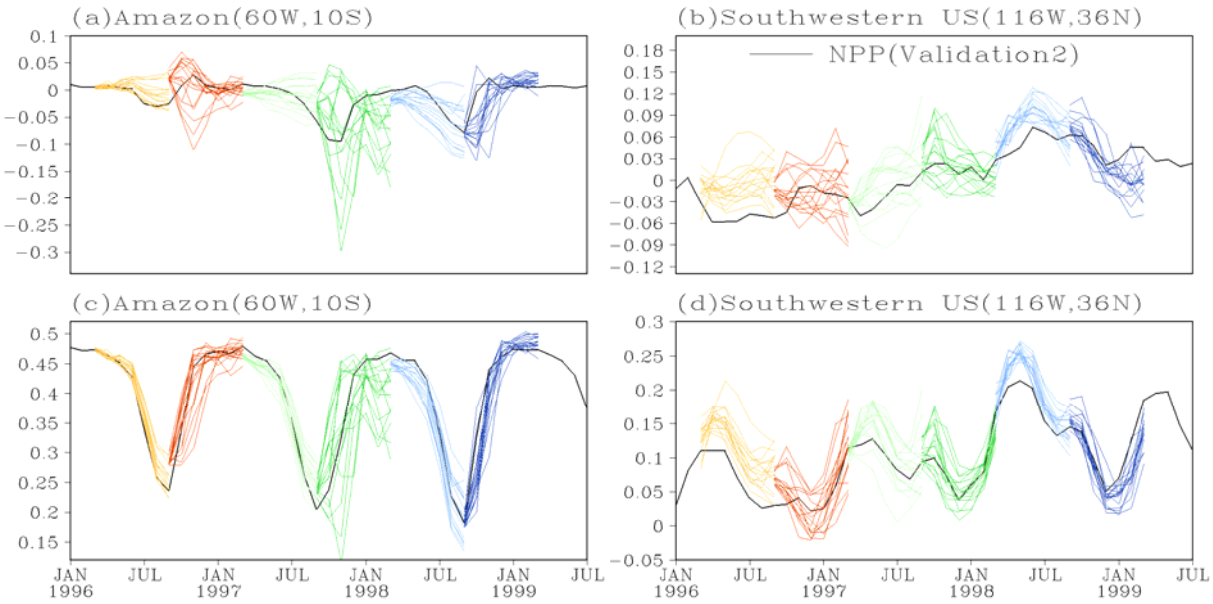


Fig. 1 Example of the CFS/VEGAS handcast: a time section of the predicted Net Primary Productivity (NPP) anomalies kg C m^{-2} y^{-1} for two grid points, one over the Amazon, the other one southwestern US, compared to the validation (black line). Each line represents one individual member of a 15-member ensemble forecast. For clarity, the forecasts were 'thinned' to show only every 6 months and for a 6-month long forecast while the actual forecasts were monthly and 9 month long. The top two panels are for anomalies while the lower panels include seasonal cycle

3. Pseudo-operational forecast

We have completed the initial setup of a one-way pseudo-operational forecast system. This is 'pseudo' in the sense that it was not actually issued, and the run was not always done in real time. However, it uses only the operational CFS input so it is what one would have got if it was done operationally. The system consists of the following key steps, which involved:

a) A shell script was developed to automatically download CFS operational forecast once a day. These forecasts are archived only for 1 week, and the daily download is for safety.

b) The data are processed by spatial interpolation. Climate variables such as precipitation and temperature anomalies are computed, then added to a climatology.

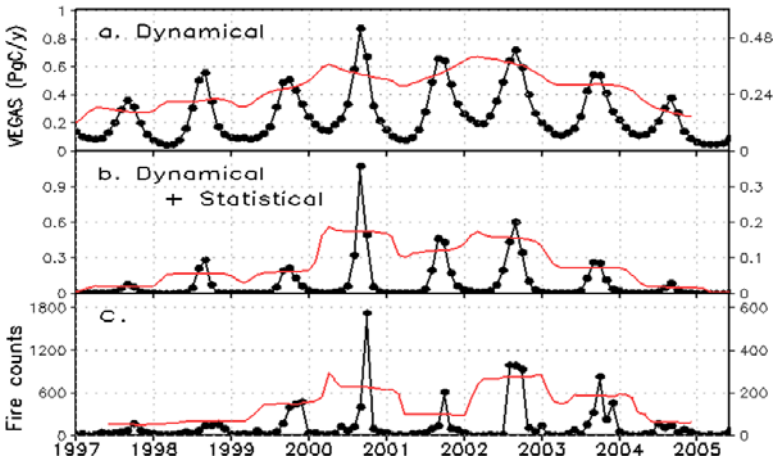


Fig. 2 Probabilistic prediction of direct carbon fluxes due to fire for the CONUS region (monthly in black and 12 month running mean in red) by (a) dynamical model using climate information alone, and (b) dynamical model but also including nonlinear effect of human management and fire suppression. Panel (c) shows satellite observations of fire counts. Large amount of carbon was released during major drought years of 2000 and 2002.

c) An ensemble of 9 climate predictions drives the vegetation/carbon model.

The forecast is conducted each month, with initial condition comes from the ensemble mean of the 1 month lead forecast from the previous month's forecast.

This system is being actively tested.

References

Zeng, N., H. Qian, C. Roedenbeck, and M. Heimann (2005), Impact of 1998-2002 midlatitude drought and warming on terrestrial ecosystem and the global carbon cycle, Geophys. Res. Lett., 32, L22709, doi:10.1029/2005GL024607.

Zeng, N., J. Yoon, A. Vintzileos, G. J. Collatz, E. Kalnay, A. Mariotti, A. Kumar, A. Busalacchi, and S. Lord (2008), Dynamical prediction of terrestrial ecosystems and the global carbon cycle: A 25-year hindcast experiment, Global Biogeochem. Cycles, 22, GB4015, doi:10.1029/2008GB003183.

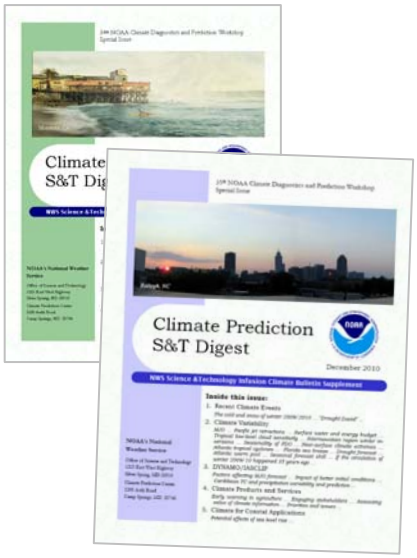
ORAL SESSION

POSTER SESSION

NWS Science and Technology Infusion Climate Bulletin

Featured Special Collections

Climate Prediction Science and Technology Digest



1. 35th Annual Climate Diagnostics and Prediction Workshop Digest

http://www.nws.noaa.gov/ost/climate/STIP/35CDPW/Digest-35CDPW.pdf

2. 34th Annual Climate Diagnostics and Prediction Workshop Digest

http://www.nws.noaa.gov/ost/climate/STIP/34CDPW/34cdpw-CollectionVol.pdf

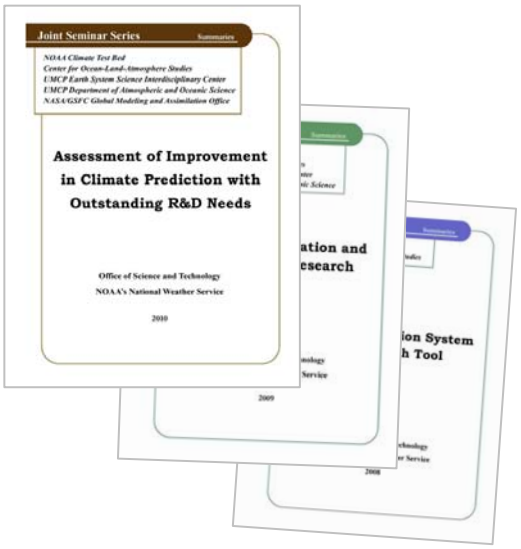
3. 33rd Annual Climate Diagnostics and Prediction Workshop Extended Summaries

http://www.nws.noaa.gov/ost/climate/STIP/33cdpw.htm

NOAA Climate Test Bed Joint Seminar Series Extended Summaries Collection Volume

1. Unified Modeling, Seamless Prediction and Integrated Services (2010-2011)

http://www.nws.noaa.gov/ost/climate/STIP/fy11jsctb.htm



2. Assessment of Improvement in Climate Prediction with Outstanding R&D Needs (2009-2010)

http://www.weather.gov/ost/climate/STIP/FY10CTBSeminars/CollectionVol_FY10CTBJS.pdf

3. Research to Operation and Operation to Research (2008-2009)

http://www.weather.gov/ost/climate/STIP/FY09CTBSeminars/CollectionVol_FY09CTBJS.pdf

4. CFS as a Prediction System and Research Tool (2007-2008)

http://www.weather.gov/ost/climate/STIP/CTB-COLA/ctb-cola_seminar_summaries.pdf