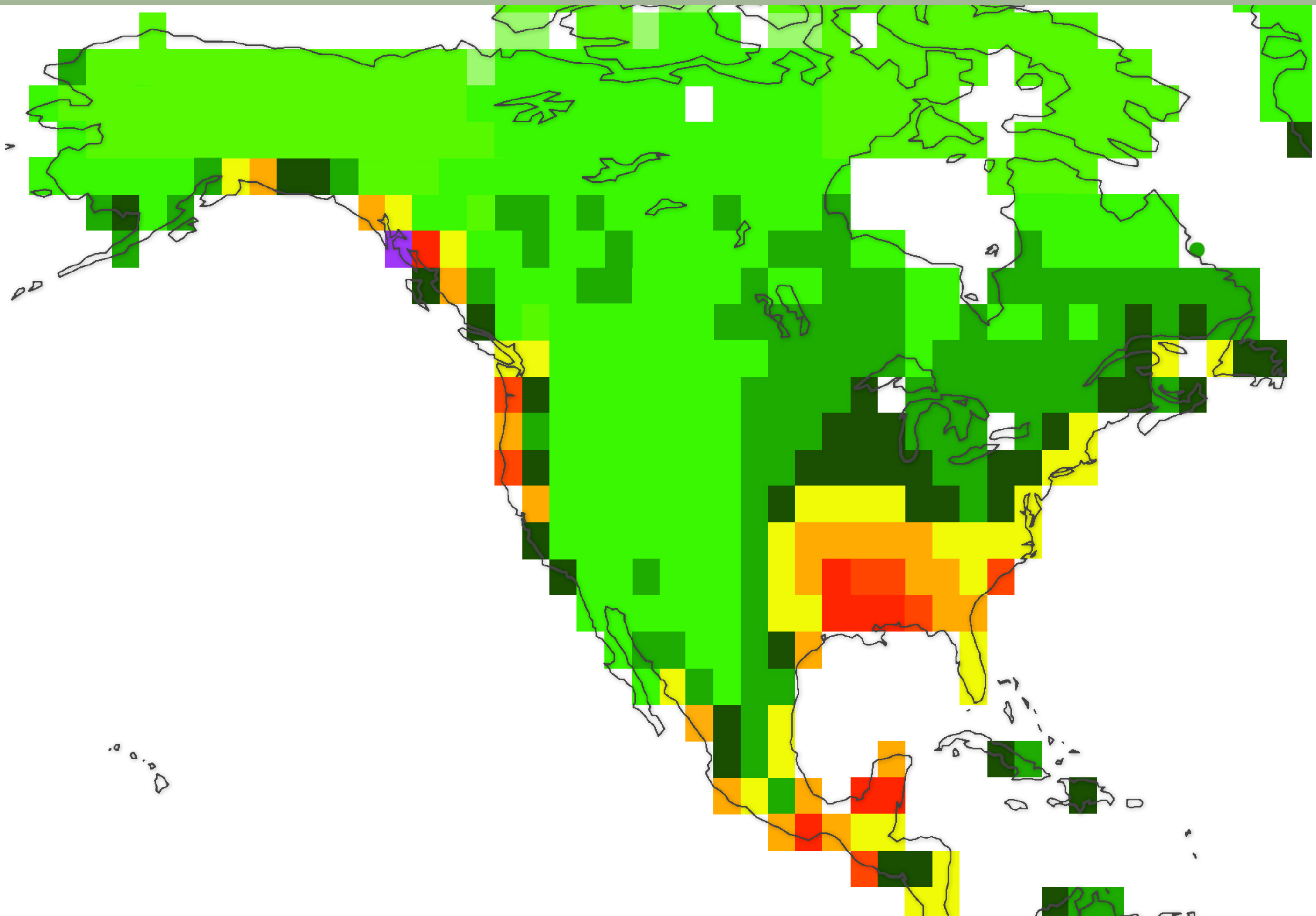
MAPP
Modeling, Analysis, Predictions, and Projections

NOAA TECHNICAL REPORT

REGIONAL CLIMATE PROCESSES AND PROJECTIONS FOR NORTH AMERICA:

CMIP3/CMIP5 DIFFERENCES, ATTRIBUTION AND OUTSTANDING ISSUES



NOAA Technical Report OAR CPO-2
doi:10.7289/V5DB7ZRC

REGIONAL CLIMATE PROCESSES AND PROJECTIONS FOR NORTH AMERICA: CMIP3/CMIP5 DIFFERENCES, ATTRIBUTION AND OUTSTANDING ISSUES

CLIMATE PROGRAM OFFICE
Silver Spring, MD
December 2014

REPORT TEAM AUTHORS

REGIONAL CLIMATE PROCESSES AND PROJECTIONS FOR NORTH AMERICA: CMIP3/CMIP5 DIFFERENCES, ATTRIBUTION AND OUTSTANDING ISSUES

Justin Sheffield [1], Andrew Barrett [2], Dan Barrie [3], Suzana J. Camargo [4], Edmund K. M. Chang [5], Brian Colle [5], D. Nelun Fernando [6,7,8] Rong Fu [6], Kerrie L. Geil [9], Qi Hu [10], Xianan Jiang [11], Nathaniel Johnson [12], Kristopher B. Karnauskas [13], Seon Tae Kim [14], Jim Kinter [15], Sanjiv Kumar [15], Baird Langenbrunner [16], Kelly Lombardo [5], Lindsey N. Long [17,18] Eric Maloney [19], Annarita Mariotti [3], Joyce E. Meyerson [16], Kingtse C. Mo [18], J. David Neelin [16], Sumant Nigam [21], Zaitao Pan [20], Tong Ren [26], Alfredo Ruiz-Barradas [21], Richard Seager [22], Yolande L. Serra [23], Anji Seth [25], De-Zheng Sun [24, 25] Jeanne M. Thibeault [26], Julienne C. Stroeve [2], Chunzai Wang [27], Shang-Ping Xie [28], Ze Yang [6], Lei Yin [6], Jin-Yi Yu [29], Tao Zhang [24,25] Ming Zhao [12]

AFFILIATIONS

[1] Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ

[2] National Snow and Ice Data Center, Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO

[3] National Oceanic and Atmospheric Administration, Office of Oceanic and Atmospheric Research, Climate Program Office Silver Spring, MD

[4] Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY

[5] School of Marine and Atmospheric Sciences, Stony Brook University – SUNY

[6] Jackson School of Geosciences, University of Texas at Austin, TX

[7] University Corporation for Atmospheric Research, Boulder, Colorado

[8] Surface Water Resource Division, Texas Water Development Board, Austin, TX

[9] Kerrie L. Geil, Department of Atmospheric Sciences, University of Arizona, Tucson, AZ

[10] School of Natural Resources and Department of Earth and Atmospheric Sciences, University of Nebraska-Lincoln, Lincoln, NE

[11] Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, CA

[12] Nathaniel Johnson, NOAA Geophysical Fluid Dynamics Laboratory, Princeton, NJ

[13] Woods Hole Oceanographic Institution, Woods Hole, MA

[14] CSIRO, Marine and Atmospheric Research, Aspendale, Victoria, Australia

[15] Center for Ocean-Land-Atmosphere Studies, Fairfax, VA

[16] Department of Atmospheric and Oceanic Sciences, University of California Los Angeles

[17] Wyle Science, Technology and Engineering, College Park, MD

[18] Climate Prediction Center/NCEP/NWS/NOAA, College Park, MD

[19] Department of Atmospheric Science, Colorado State University, Fort Collins, CO

[20] Saint Louis University, St. Louis, MO

[21] Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD

[22] Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY

[23] Department of Atmospheric Sciences, University of Arizona, Tucson, AZ

[24] Cooperative Institute for Environmental Studies/University of Colorado

[25] NOAA Earth System Research Laboratory, Boulder, CO

[26] Department of Geography, University of Connecticut, Storrs, CT

[27] Physical Oceanography Division, NOAA Atlantic Oceanographic and Meteorological Laboratory, Miami, FL

[28] Atmospheric Science & Physical Oceanography, Scripps Institution of Oceanography, University of California, San Diego, CA

[29] Department of Earth System Science, University of California, Irvine, CA

The authors wish to thank the following individuals for their careful review of the report:
Ruby Leung (Pacific Northwest National Laboratory), Michael Wehner (Lawrence Berkley National Laboratory), Tom Knutson (NOAA Geophysical Fluid Dynamics Laboratory), and one additional anonymous reviewer.

TABLE OF CONTENTS

ABSTRACT

The Coupled Model Intercomparison Project, phase 5 (CMIP5) provides an unprecedented set of climate model data from coordinated experiments that can be used to address a variety of questions related to climate change and climate variability. The CMIP5 builds on the previous model intercomparison phase (CMIP3) in several ways, including a larger number of modeling centers and models, the use of generally moderately higher resolution models, and the inclusion of more complex and complete representation of Earth system processes. A key question is whether the CMIP5 results have improved since CMIP3 in terms of the representation of observed climate features and processes, and whether the future projected changes are more robust, and why. This report addresses these questions for a suite of climate variables and regional processes for North America and provides recommendations for future analyses and experiments to resolve some of the ongoing issues.

Overall, the multi-model ensemble (MME) mean performance has not improved substantially in CMIP5 relative to CMIP3 for climatological variables (precipitation, sea surface temperature - SST), except for a slight improvement for near surface air temperature over land. CMIP5 models tend to underestimate the frequency of heavy and extreme daily precipitation events, despite a slight improvement over CMIP3, especially in the southeastern US. Projected increases in moderate to extreme precipitation events are similar to CMIP3. Generally, the CMIP5 models show better skill for basic attributes of El Niño and the Southern Oscillation (ENSO) with performance related to the mean SST state. It is unclear whether the representation of teleconnections with North American climate has improved. It is likely that the structure of Pacific Decadal Variability (PDV) as indexed by the PDO is slightly better simulated and teleconnections for precipitation also are improved slightly but remain poor overall. Atlantic Multidecadal Variability (AMV) as represented by the AMO is better represented in CMIP5 models in terms of decadal variability and persistence than CMIP3 models, but its SST footprint and teleconnections with North American climate are still poorly represented. Regionally, projections of changes in precipitation from CMIP5 for the sub-tropics tend to be more robust overall than CMIP3, in particular for summer drying in the Caribbean and southwest Mexico. The boundary between winter wetting and drying in the Southwest US is projected to move southward in CMIP5 relative to CMIP3 results, although the changes are highly dependent on the region and season. The CMIP5 models project a more significant decrease in extra-tropical storm track activity than CMIP3, which may be related to a larger projected decrease in the temperature gradient between lower and higher latitudes.

1. INTRODUCTION

This report summarizes the results of an analysis of the Coupled Model Intercomparison Project, phase 5 (CMIP5; Taylor et al., 2012; http://cmip-pcmdi.llnl.gov/cmip5/) data sets for North America with special emphasis on the qualitative and quantitative changes in results from the previous model intercomparison phase (CMIP3) for regional climate processes, attribution of changes where possible, as well as ongoing issues in model evaluation. This report has been developed as an aid in communications with the community interested in the results from CMIP5, to address specific questions on climate impacts, adaptation and vulnerability that are of high interest, and in particular the National Climate Assessment community (NCA; http://globalchange.gov/what-we-do/assessment/nca-overview.html). This report is not an exhaustive list of questions/issues – rather it is intended as an initial effort that can determine the efficacy of such an assessment and advise future such efforts.

The CMIP5 provides an unprecedented set of climate model output data that can be used to address a variety of questions related to climate change and climate variability, including: the assessment of future climate projections, evaluations of models for historical climate, the attribution of observed climate change and understanding of climate processes and feedbacks. The CMIP5 builds on the previous CMIP3 project in several ways that are expected to result in better skill in representing current climate and improved understanding of the uncertainties in future projections: 1) a larger number of modeling centers and models have participated in the experiments; 2) the models are generally run at higher spatial resolution, which helps better resolve topographic and coastline features that are important for representing regional processes; 3) some models are more comprehensive in the processes they represent; 4) many modeling centers have provided multiple realizations of simulations that permit quantification of uncertainties.

Differences between CMIP5 and CMIP3 include a larger number of participating modeling centers, generally higher model resolution, more comprehensive representation of Earth system processes, and more ensemble members for each model.

A key question is whether the CMIP5 results have improved compared with CMIP3, and why. While it is possible to compare published results from CMIP5 models with earlier publications based on CMIP3 models, a direct comparison of CMIP5 and CMIP3 data is more conducive to establishing differences between the performance of the models. This report takes both approaches to provide a more comprehensive evaluation of the representation of climate processes in the historical simulations and the robustness of changes for the future projections.

This report was compiled under the auspices of the National Oceanic and Atmospheric Administration (NOAA) Modeling, Analysis, Predictions and Projections (MAPP) program, which is managed by the Climate Program Office (CPO) in the Office of Oceanic and Atmospheric Research (OAR). A special CMIP5 Task Force (http://cpo.noaa.gov/ClimatePrograms/ModelingAnalysisPredictionsandProjections/MAPPTaskForces/CMIP5TaskForce.aspx) composed of MAPP principal investigators was formed, originally intended to evaluate the fidelity of CMIP5 simulations of the past century of climate and to determine the scientifically-defensible conclusions that can be drawn from CMIP5 projections of future climate, focused on the climate of North America. The results of the Task Force's analysis have been documented in a Special Collection of the *Journal of Climate* including 22 papers and three summary papers (Sheffield et al., 2013a,b; Maloney et al., 2013) that provide an overview of the scientific results. The results reported in this and other relevant literature are provided in this report together with new analysis of the CMIP3 and CMIP5 databases to provide a consensus on the differences and ongoing issues. Where there is a lack of information or consensus, recommendations for future research are also made, as appropriate. The results are intended to be complementary to Sun et al. (2014), who discuss regional-scale differences between CMIP3 and CMIP5 for surface temperature and precipitation means and extremes over the US.

This report is organized into seven sections, including the introduction. Section 2 describes the differences in the emission scenarios between CMIP3 and CMIP5 and discusses the implications for comparisons between the two sets of simulations. This section also describes the model simulations and observations used. We begin the analysis in section 3 by discussing changes in basic surface climate variables (precipitation, land air and sea surface temperature) and then extremes of precipitation and temperature at the daily time scale (section 4). Section 5 addresses inter-annual to decadal variability, in terms of the representation of the El Niño Southern Oscillation (ENSO), the Pacific Decadal Oscillation (PDO) and Atlantic Multidecadal Variability (AMV) and their teleconnections with North American near-surface climate. Some regional climate processes are evaluated in section 6, including changes in precipitation in the Southwest, a region expected to get drier overall in the future with implications for water resources, and the warming hole phenomenon.

2.1. EMISSION SCENARIOS IN CMIP3 AND CMIP5

A key difference between the CMIP3 and CMIP5 is the set of redesigned emissions scenarios that were used to force the models. The CMIP3 models were forced by the emissions scenarios from the Special Report on Emissions Scenarios (SRES; Nakićenović et al., 2000) while the CMIP5 models were forced by Representative Concentration Pathways (RCPs; van Vuuren et al., 2011). The CMIP3 scenarios represent futures with different mixes of population growth and policies on energy sources. In contrast, the CMIP5 RCP scenarios do not relate to particular policy actions but reflect mitigation scenarios that lead to one of the RCPs. Figure 1 shows the total radiative forcing for a selection of the SRES scenarios and RCPs for CO_2 from fossil fuel sources, and indicates those scenarios that are comparable. Furthermore, there are differences in the specification of other anthropogenic greenhouse gases (GHG) and aerosols, with more species prescribed in CMIP5, such as black and organic carbon, alongside the inclusion of land use change in the historic and future projections for CMIP5 (Taylor et al., 2012).

> Differences in projected changes are in large part due to different assumptions about non-greenhouse gas forcing between the CMIP3 and CMIP5 scenarios, in particular the specification and treatment of aerosols.

Recent studies indicate that the equilibrium climate sensitivity (ECS) to a doubling of CO_2 concentration ($2xCO_2$) and the transient response of global/regional temperature and precipitation to increasing CO_2 remain similar in CMIP5 and that differences in projected changes are in large part due to different assumptions about non-GHG forcing between the CMIP3 and CMIP5 scenarios, in particular the specification and treatment of aerosols. For example, initial investigation of ECS for temperature by Andrews et al. (2012) indicates similar behavior between CMIP3 and CMIP5. Analysis of global-mean temperature, tropical Atlantic temperature, their difference, and inferred hurricane activity by Villarini and Vecchi (2012, 2013), show similar responses for CMIP3 and CMIP5 models in terms of the response to CO_2 but differences among scenarios. For one model, GFDL-CM3, they showed that the difference could be attributed to the treatment of aerosols. Further work by Knutson et al. (2013) shows that the difference in Atlantic SST, wind shear, and potential tropical cyclone intensity projections between CMIP3 and CMIP5 are not attributable to different response to GHGs, but to different non-GHG forcing. This also appears to be true for precipitation (Fig. 2; G. Vecchi, 2013, unpublished manuscript), implying that differences from model changes are small, and that aerosol treatment and GHG forcing differences are responsible for the differences seen in the transient scenario projections. As discussed later in the context of regional climate processes, further analysis is needed to see if these results extend to other models and whether non-GHG forcings are driving differences in the projections.

2.2. MODEL AND OBSERVATIONAL DATA

We use data from multiple climate models from the CMIP3 and CMIP5 databases for the historic simulations (“20C3M” for CMIP3; “historical” for CMIP5) and a range of future scenarios. The “historical” simulations are run in coupled atmosphere-ocean mode forced by historical estimates of changes in atmospheric composition from natural and anthropogenic sources, volcanoes, GHG and aerosols, as well as changes

Figure 1. Total radiative forcing from fossil fuel CO_2 emissions for selected SRES and RCPs (Figure created at live.magicc.org; Meinshausen et al., 2011).

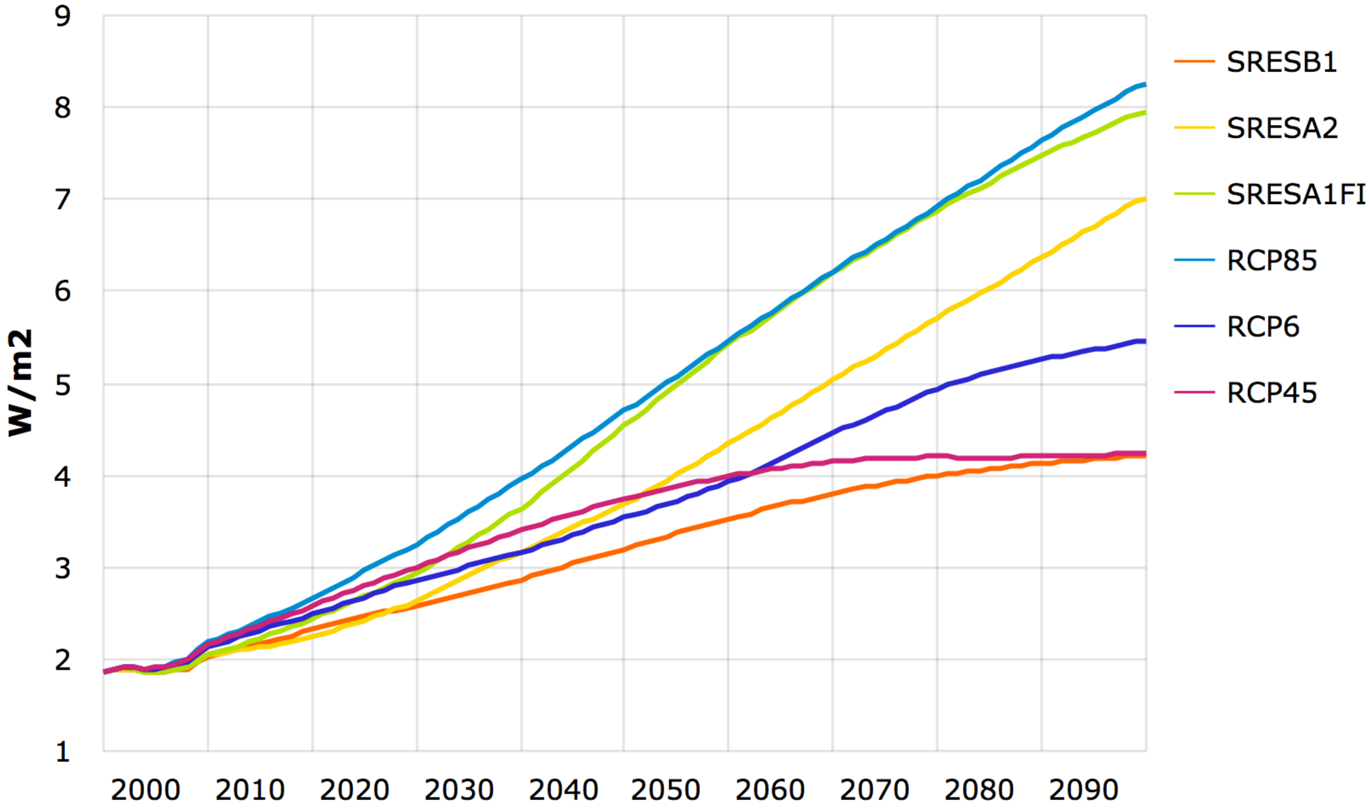
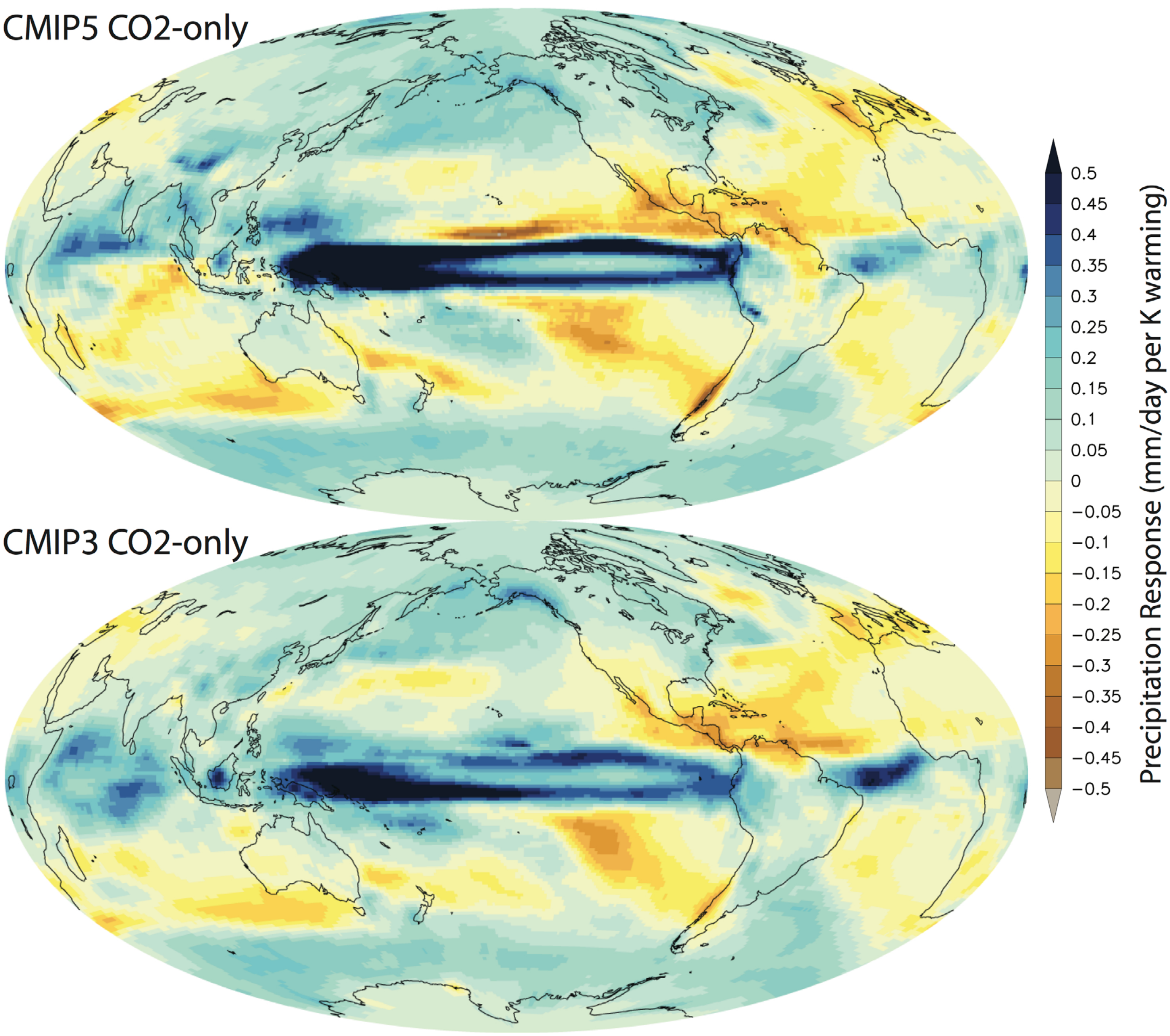


Figure 2. Precipitation response (change per K of warming) to doubling of CO_2 for CMIP5 (top) and CMIP3 (bottom) models.



in solar output and land cover. Note that only anthropogenic GHGs and aerosols are common prescribed forcings for all models; each model differs in the set of other forcings that it uses and how it implements those forcings, such as land use change. For Earth system models (ESMs), the carbon cycle and natural aerosols are modeled and can therefore introduce feedbacks.

The historical simulations are generally carried out for the period 1850-2005 but we evaluate the models for the most recent 30 years, depending on the availability of observations. In the case of remote sensing observations, this restricts the analysis to the satellite period from about 1979 onwards. In some cases, multiple observational datasets are available and are used to estimate the uncertainties in the observations. Details of the datasets and data processing are given in the relevant sub-sections and figure captions. Most of the models have multiple ensemble members and in general we use the first ensemble member (all members have equal utility for the analysis and the first member is only chosen as a matter of convenience, possibly at the expense of over- or under-estimating model errors especially at smaller scales). In some cases, the results for multiple ensembles are averaged where appropriate or used to assess the variability across ensemble members. Results are generally shown for the multi-model ensemble (MME) mean and for the individual models using performance metrics that quantify the errors relative to the observations or the change in a future period relative to the historic period.

2.3. SOME DEFINITIONS AND CAVEATS

The CMIP models are evaluated on an individual and ensemble basis, also drawing from existing evaluations in the literature. Model performance is dependent on the variable and scale of the evaluation and as the report covers a wide range of climate processes, it is not possible to have a universal definition of performance. In general, the performance of the MME is based on whether the ensemble envelops the observation, and the performance of individual models is based on measures of distance from the observation, relative to other models. In each case there are uncertainties associated with the choices made in evaluating the models and assessing whether attribution of historic changes and projected changes are robust. Here, robustness generally refers to the level of indifference of the results to differences in these choices, for example, to different models or to alternative observations. There are several factors that can influence the evaluations, including: model selection and ensemble size, model internal variability, climate drift and observational uncertainties. Although we generally do not take these factors into account because they are deemed to have a relatively small influence on the overall results (e.g. internal variability versus inter-model differences), they are discussed next and also within sub-sections because they may be important for particular models and metrics. These caveats also have implications on whether we can determine if an individual model or the MME has improved from CMIP3 to CMIP5. For individual models, in particular, the concept of improvement is muddled because newer models will generally incorporate better physics and can therefore be considered more comprehensive, but this does not necessarily lead to improved performance against observations. The conclusions of this report should therefore be interpreted with this in mind.

The set of models analyzed here and within the cited references varies in terms of the number and membership. The analysis is generally based on a core set of about 15 models that was used by Sheffield et al. (2013a,b) and Maloney et al. (2014) (and references within) and selected to subjectively span a range of modeling centers and model types (AOGCMs, ESMs). The actual number used in each analysis varied based on data availability and the effort required to process high temporal resolution data. No account was taken to determine model independence or the sensitivity of results to different subsets of models. Studies of model selection are generally geared towards robust estimates of future projections (e.g. McSweeney et al., 2012; Thober and Samaniego, 2014) and are based on model performance, physical plausibility of projections and spanning the range of projections, rather than quantifying the uncertainty in sub-sampling. However, we expect that the error associated with using an ensemble of at least 15 models is small compared to the inter-model uncertainty.

Evaluations of individual models against observations are subject to errors induced by internal climate variability that ensures that coupled models will not replicate observed variations in climate on annual to decadal time scales, unless there is a strong external forcing (e.g. volcanic eruption). For this reason, evaluations against observations are subject to differences in the current states of the observed and modeled climates. Evaluating the models using metrics calculated over multiple decades and large regions, and for multiple model realizations, will reduce the impact somewhat. However, there is potential to unduly penalize models, for example, for extreme events for which the number of events is small, or for decadal climate variability for which the modeled phase will not necessarily line up with the observations. We note throughout the report where there is potential for these types of discrepancies. For future projections, internal variability can obfuscate the detection of forced changes for several decades, with changes in temperature being detectable earlier than for precipitation (e.g. Hawkins and Sutton, 2009, 2010). This is dependent on the size of the ensemble (Deser et al., 2012). For this reason, we evaluate projected changes at the end of the 21st century when the forced signal should be detectable.

The magnitude of climate drift in the models is relatively small for the upper ocean, atmospheric and land processes considered in this report and has improved considerably in CMIP5 relative to CMIP3.

Climate drift in coupled models is another aspect that can affect the evaluation of historic climate and future projections. Drift is generally a result of model component coupling and deficiencies in model physics and numerical schemes (Sen Gupta et al.,2012), and can be significant especially at smaller spatial scales and for the deep ocean for which equilibrium has not been reached. The magnitude of the drift is relatively small for the upper ocean, atmospheric and land processes considered in this report and has improved considerably in CMIP5 relative to CMIP3 (Sen Gupta et al., 2013). Furthermore, the magnitude and direction of drift is model dependent and so evaluations based on the MME are less affected than for individual models. Individual model drift in precipitation (which is highly dependent on SST drift) for CMIP5 models is generally less than 10% of the historical model trend, but for some models this can exceed 30%.

A limited analysis of errors for North American winter/summer precipitation (P), surface air temperature (T) and sea surface temperature (SST) was carried out in Sheffield et al. (2013a). In analyzing errors, model outputs were compared with the best available observational or reanalysis data sets. Figure 3 shows a direct comparison of CMIP5 results with CMIP3 results for basic climate variables over the period 1971-1999 (Sheffield et al. 2013a). The figure shows root-mean-square error (RMSE) values for CMIP5 and CMIP3 models for seasonal mean P and T over North America and SST over the surrounding oceans. The majority of the CMIP5 models analyzed have an equivalent CMIP3 model, that is either the same model (HadCM3), a newer version, or an earlier related version, and so a direct comparison of changes since CMIP5 is feasible but subject to uncertainties in how the two versions are related (see Section 2.3).

Overall, the multi-model ensemble mean performance has improved slightly in CMIP5 for nearly all climatological variables.

Overall, the MME mean performance has improved slightly in CMIP5 for nearly all climatological variables. For example, there is a reduction in the MME mean RMSE for summer P (0.90 mm day^{-1} for CMIP3, 0.86 mm day^{-1} for CMIP5), and for winter SST (1.72°C to 1.55°C). The largest percentage reduction in RMSE for the MME mean is for summer T (11.8% reduction in RMSE). The spread in model performance (as quantified by the standard deviation) has remained about the same for P, increased for T and decreased for SST. The increase in spread for T is due to both increases and decreases in individual model performance relative to the CMIP3 models. Several models have improved considerably and across nearly all variables and seasons, such as the CCSM4, INMCM4, IPSL-CM5A-LR, and MIROC5, although it is unclear how these models are related to their previous versions. Reductions in performance for individual models are less prevalent across variables, but are large for CSIRO-Mk3.6.0, HadCM3, and MRI-CGCM3 for SST in both seasons. The CanESM2 has worse performance than its CMIP3 equivalent (CGCM3.1) for all variables, although it is likely that these are very different models and so a direct comparison may be unwarranted. The HadCM3 model, which is used for both the CMIP3 and CMIP5 simulations, appears to have degraded in performance for SST. We assume that the model has not changed and that changes in variability derived from different initial conditions and/or external forcings have caused the changes in performance.

A more detailed regional examination of the errors is shown in the next sections in terms of maps of the MME mean fields for climatological winter and summer in Figs. 4-6 (Ruiz-Barradas et al. 2013). As for the continental evaluations above, the MME mean fields have not changed much from CMIP3 to CMIP5 models, except for a slight improvement in T.

3.1. TEMPERATURE

The MME mean of CMIP5 models shows a slight improvement over the corresponding mean of CMIP3 models of winter and summer T over North America (Fig. 4). Winter T is characterized by below zero (Celsius) temperatures in the inland northern latitudes with a minimum over the Rocky Mountains and a maximum over Mexico and Central America; the effect of the oceans in the northern latitudes temperatures is seen through the above-zero T along the coastal strips

Figure 3. Comparison of CMIP5 and CMIP3 model performance for metrics of relevance to North America for seasonal (DJF and JJA) precipitation (P), surface air temperature (T) and SST. Results are shown as RMSE values calculated for 1971-1999 relative to the GPCP, CRU and HadISST observational datasets. Precipitation and temperature RMSE values are calculated over North America (130°-60°W, 0°-60°N) and SST RMSE values are calculated over neighboring oceans (170°-35°W, 10°S-40°N). The core set of CMIP5 models and their equivalent CMIP3 models where available (otherwise indicated by N/A) are shown. The MME mean values are also shown. This is Figure 20 in Sheffield et al. (2013a).

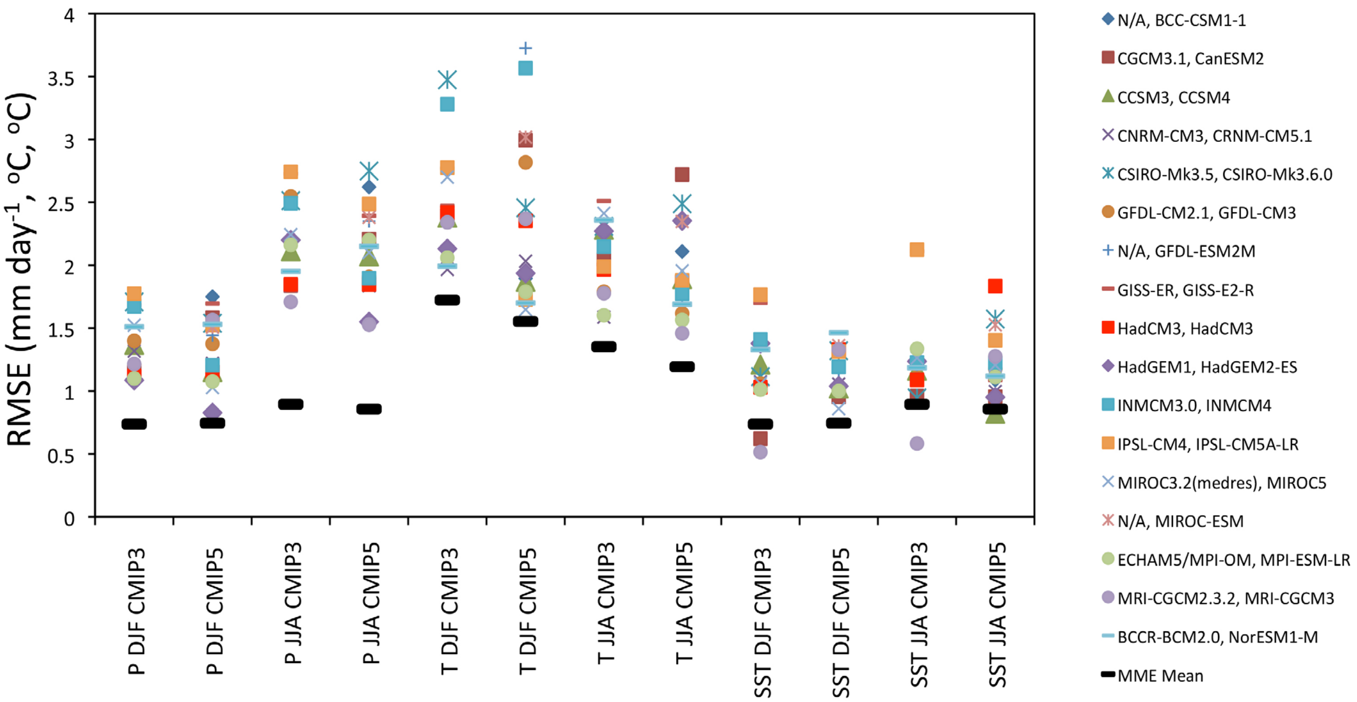
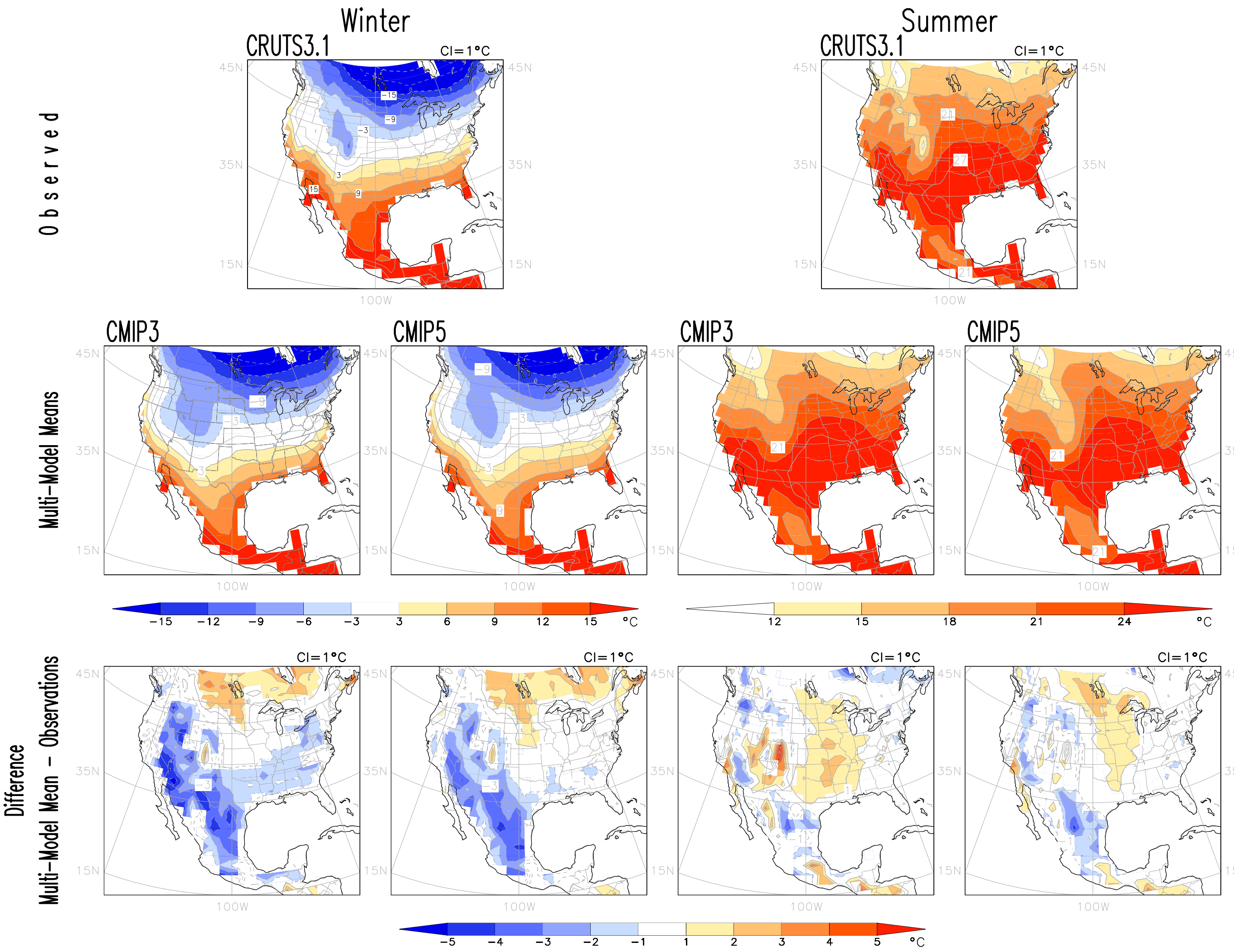


Figure 4. Mean winter and summer climatological surface air temperature for the period 1971-1999. Observations (upper row from CRUTS3.1), CMIP3 and CMIP5 multi-model means (middle row) and difference of multi-model mean minus observations (lower row). Red/blue shading denotes positive/negative temperatures (upper and middle rows) and their differences (lower row); contour interval is 1mm day^{-1} for both mean values and differences. Statistically significant differences calculated at the 5% level using a two-tailed Student's t-test are shaded. Fields have been interpolated to a common 1.5°×1.5° grid. Figure adapted from Ruiz-Barradas et al (2013).



of the Pacific Northwest and Northeastern US. The structure of the CMIP3 and CMIP5 MME means agree with each other and with the general structure of the observed winter T; however the extension of the minimum over the Rocky Mountains and the southward extension of cool temperatures over the highlands of Mexico are larger than observed in both of the MME means. Differences from observations of the MME means are smaller in the CMIP5 than in the CMIP3, especially over the East South Central and South Atlantic states of the US and to a lesser degree over the Rocky Mountain states of the US and the highlands of Mexico. Mean summer T is characterized by relatively low values over the Pacific Northwest (15-18°C) and the Rocky Mountains where they reach a minimum (~9°C), and a tongue of warm T penetrating the US from the Gulf of Mexico through the South Central states and extending to the northern Great Plains. High T (larger than 27°C) is evident over the South Central states of the US as well as over southern California and Arizona and the Sonora Desert over Mexico; minimum T in Mexico is found over the highlands (18-21°C). The CMIP3 and CMIP5 MME means of summer T have structures similar to the observations with some differences between them, mainly on the extension of the minimum over the Rocky Mountains and the Mexican highlands; the extension of these minima is closer to observations in the CMIP5 mean than in the CMIP3 mean. The CMIP5 MME mean error over the Rockies and the Mexican highlands is smaller than for the CMIP3 mean; errors also show that the tongue of warm air over the southern Great Plains and the high T over northwestern Mexico (including the Sonora desert) are better captured by the CMIP5 mean than by the CMIP3 mean.

In terms of variability, Knutson et al. (2013b) showed that decadal-scale variability of surface air temperature is generally overestimated by CMIP3 and CMIP5 models (based on long control simulations) in high latitudes globally, and generally overestimated across most of North America, with underestimation in northwest Mexico. The patterns of variability are reasonably well reproduced (spatial correlation coefficients of 0.5 – 0.7) by individual models from both CMIP3 and CMIP5. Substantial differences between CMIP3 and CMIP5 were not found.

When temperature changes are normalized by the global temperature change, the geographic patterns of mean changes between CMIP3 and CMIP5 are very consistent.

Temperature is projected to increase in CMIP5 models in all seasons under RCP8.5 with the greatest warming in wintertime at high latitudes (up to 15°C increase), but with higher inter-model variability (Maloney et al., 2014). Direct comparison of CMIP3 and CMIP5 projections is difficult because of the different scenario forcings (see section 2.1). However, as the mean and range of the climate sensitivity and transient climate responses are similar (Andrews et al., 2012), the differences are likely to be due more to the scenario forcing rather than changes in the models. Direct comparison by Knutti and Sedlácek (2012) showed that the there is little difference in the projections, although warming globally and over North America is slightly higher in CMIP5. When changes are normalized by the global temperature change, the geographic patterns of mean changes between CMIP3 and CMIP5 are very consistent (pattern correlation = 0.98 globally), although the absolute changes are statistically significantly different in high latitudes and parts of Mexico (Collins et al., 2013).

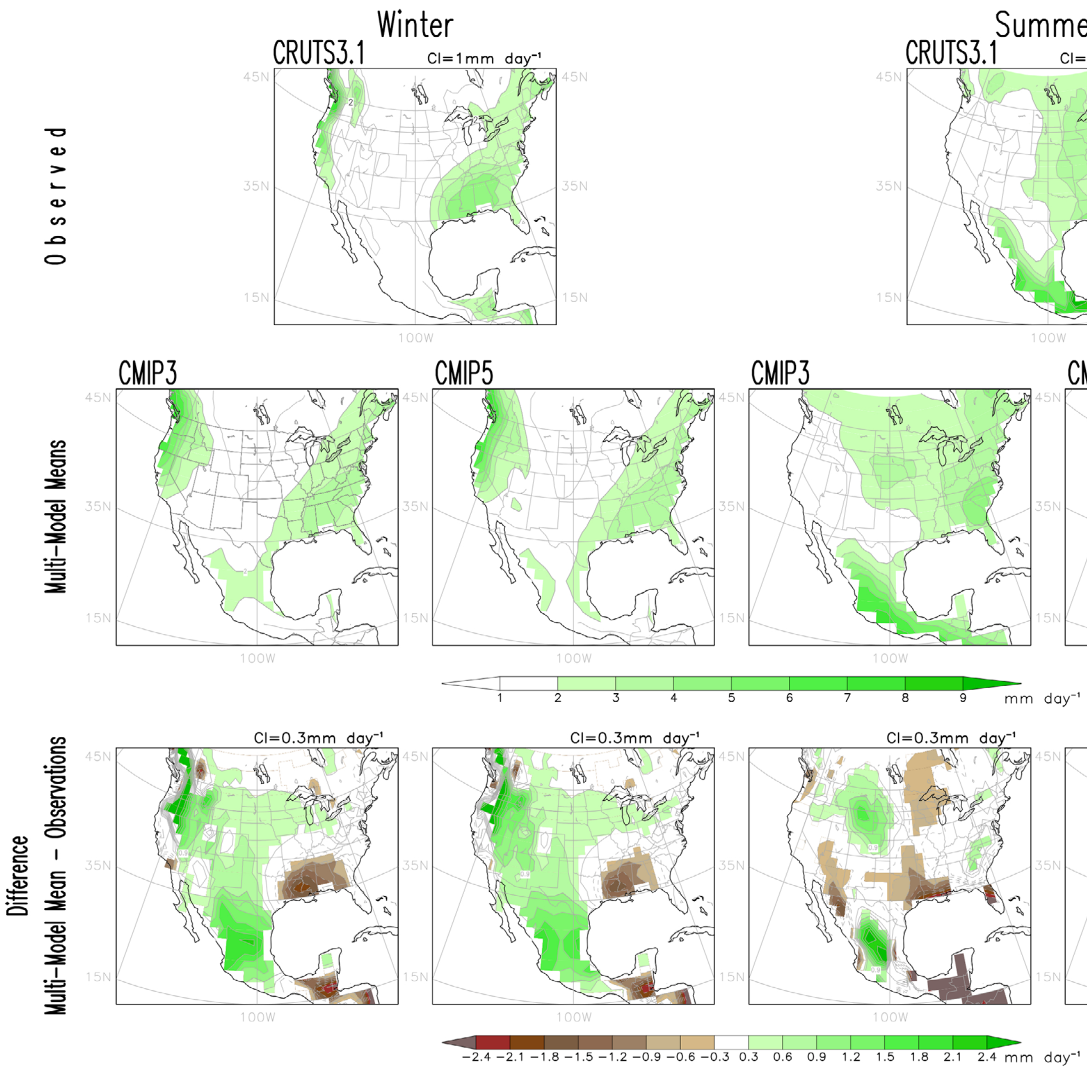
3.2. PRECIPITATION

The CMIP5 MME mean shows no improvements over the corresponding mean of CMIP3 models of mean winter and summer P over North America (Fig. 5). Mean winter P is characterized by maximum P over the Pacific Northwest and the East South Central regions of the US, as well as over Central America and southern Mexico; the Mountain and Midwest regions of the US as well as the majority of Mexico receive less P in this season. The CMIP3 and CMIP5 MME means of winter P agree in their general structure with the observed P over the US, but the maximum in the Pacific Northwest extends toward the Mountain region and the maximum over the South Central region is absent in both sets of models. Differences with observations show that the MME mean P is larger than observations over practically the whole of North America except over the East South Central region of the US and Central America and southern Mexico where P is less than observed. On the other hand, mean summer P is characterized by a wetter eastern US with maximum values along the narrow coastal region of the South Atlantic states of the US, and Mexico and Central America; western US and northern Mexico, along the Texas border, have less P in this season. The general structure of the CMIP3 and CMIP5 MME means of summer P are similar to each other and to the observed summer P, however differences are apparent over the Mountain, West North Central and South Atlantic states of the US and over the high plains of Mexico. Excessive P is present in the CMIP3 and CMIP5 MME mean over the Mountain and inland of the South Atlantic states of the US and the high plains of Mexico. Conversely, reduced P is simulated over the central, western south central and the coastal band of the southern states of the US (particularly Florida and along the Gulf of Mexico coast), as well as over the mountain regions of Mexico and Central America.

When normalized to global temperature change, the spatial patterns of precipitation change over North America are similar between CMIP3 and CMIP5

Future projections of wintertime P for 2070-2099 in CMIP3 (SRESA2) and CMIP5 (RCP8.5) models show a similar large-scale pattern of increases in mid- to high latitudes and decreases in the subtropics (Maloney et al., 2014). There are projected increases along the western coast of North America from California northwards and the eastern coast from the mid Atlantic states northward with good model agreement on the sign of the changes. However, the boundary between the increases and decreases is shifted slightly south in the CMIP5 ensemble, to give increases over parts of California (Neelin et al., 2013; see section 6.3 for more details). During the summer, higher precipitation is projected in Alaska and the Yukon, with complete model agreement on the sign, and along the entire Arctic coast. Summertime reductions in precipitation are focused on the east Pacific warm pool and the Caribbean (Maloney et al. 2014). The model agreement on these summertime decreases has increased from CMIP3 to CMIP5, with all CMIP5 models examined by Maloney et al. (2014) in agreement for major parts of the region (see section 6.1). When normalized to global temperature change, the spatial patterns of precipitation change over North America are similar between CMIP3 and CMIP5, with significant differences in the magnitude where there is a northward shift in the boundary between wetting and drying in the southwestern US, and also across parts of Mexico, possibly reflecting the differences in scaled temperature change (Collins et al., 2013).

Figure 5. Mean winter and summer climatological precipitation for the period 1971-1999. Observations (upper row, from GPCP), CMIP3 and CMIP5 multi-model means (middle row) and difference of multi-model mean minus observations (lower row). Green/brown shading denotes positive/negative differences in precipitation (lower row); contour interval is 1mm day^{-1} for the mean values and 0.3 mm day^{-1} for the differences. Statistically significant differences calculated at the 5% level using a two-tailed Student's t-test are shaded. Fields have been interpolated to a common 1.5°×1.5° grid. Figure adapted from Ruiz-Barradas et al (2013).

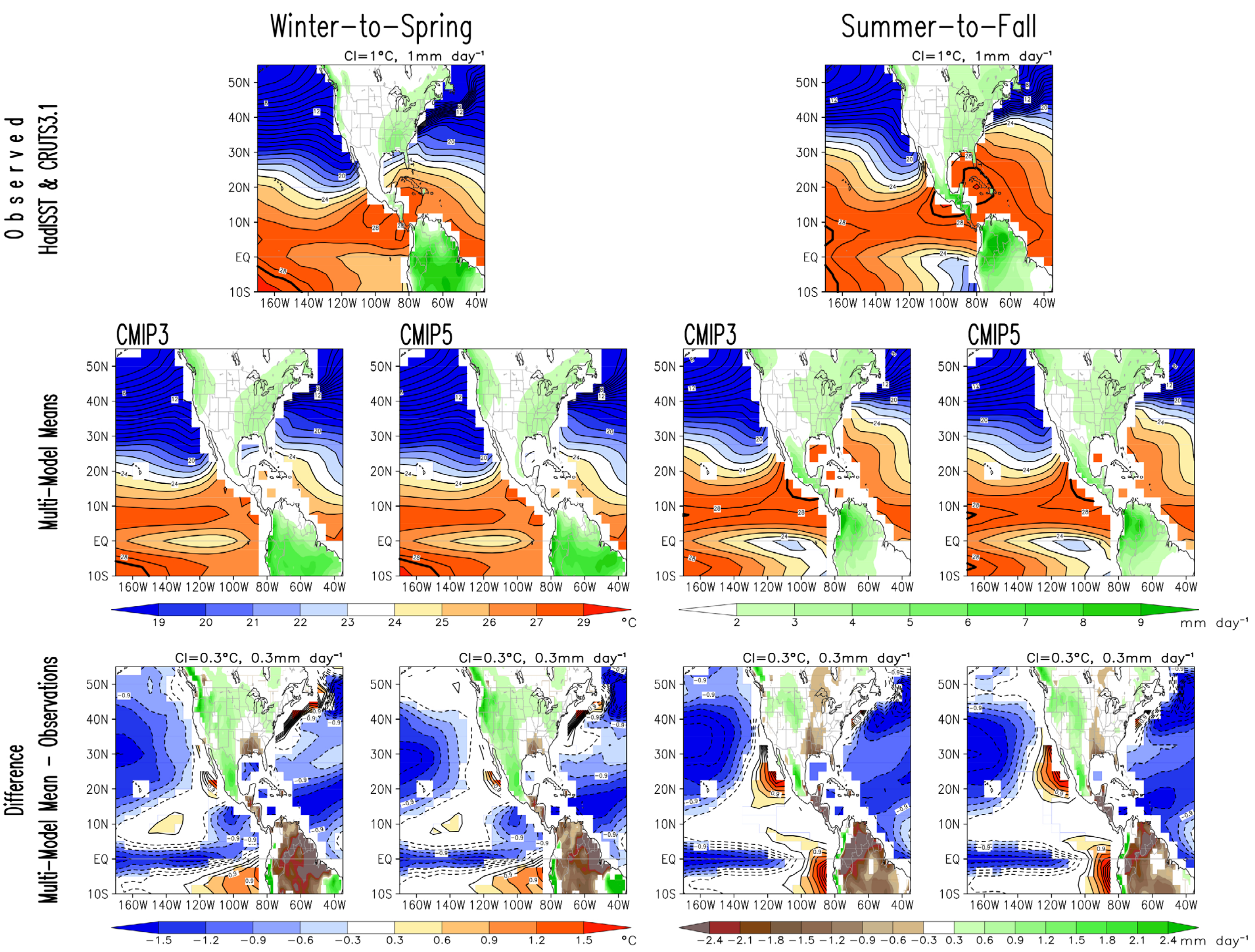


3.3. SEA SURFACE TEMPERATURE

The multi-model ensemble mean of CMIP5 models for sea surface temperature shows no improvement over the corresponding mean of CMIP3 models of winter-to-spring and summer-to-fall sea surface temperature in the adjacent oceans of North America

The MME mean of CMIP5 models for SST shows no improvement over the corresponding mean of CMIP3 models of winter-to-spring (DJFMAM) and summer-to-fall (JJASON) SST in the adjacent oceans of North America (Fig. 6). The Western Hemisphere Warm Pool (WHWP), where SST is equal to or larger than 28.5°C (thick black line on maps), usually is absent from December to February and appears in the Pacific from March to May, while it is present in the Caribbean and Gulf of Mexico, or Intra-Americas Sea, from June to November (Wang and Enfield, 2001). The cooler part of the year (winter to spring) is characterized by the small extension of SST in excess of 27°C and a suggestion of a cold tongue in the eastern equatorial Pacific, while during the warmer part of the year (summer to fall) the extension of SST in excess of 27°C is at a maximum and the cold tongue is well defined over the eastern Pacific. High P along the Mexican coasts, Central America, the Caribbean Islands and the central-eastern US is associated with tropical SST in excess of 27°C during the warm half of the year; a decrease in the regional precipitation south of the equator is also evident in this warm half of the year. The structure of the CMIP3 and CMIP5 MME means agree with each other and with the general structure of the observed winter-to-spring SST, except that they do not show the high SST in the eastern Pacific and Intra-Americas Sea around Central America, or the weak cold tongue in the equatorial eastern Pacific off the coasts of Ecuador and Peru. The errors are particularly characterized by warm biases off the northeast coast of the US, especially during the cold season; off California and the Pacific side of northwestern Mexico, especially during the warm half of the year; and off the coasts of Ecuador and Peru, a problem that may be related to the models' poor ability to simulate stratus clouds and transport by Ekman currents (Zheng et al. 2011). Cold biases are extensive over both oceans. The cold SST bias in the equatorial Pacific seems to agree with the wet/dry bias over the US (resembling La Niña conditions), while the warm SST bias off the coasts of Ecuador and Peru agree with the wet coastal bias over these countries and the dry bias over the countries to the east (resembling El Niño conditions). The structures of the CMIP3 and CMIP5 MME means also agree with each other and with the general structure of the observed summer-to-fall SST, except that the SST in the Atlantic Warm Pool (AWP) region is cooler than observed (Liu et al., 2013). The AWP SST bias in CMIP5 is more modulated by an erroneous radiation balance due to misrepresentation of high-level clouds rather than misrepresentation of low-level clouds as was the case for CMIP3 (Liu et al. 2013). The CMIP3 and CMIP5 MME mean errors highlight similar areas of cold/warm bias to those in the winter-to-spring part of the year except that the cooling bias over the Pacific side of Central America is no longer present.

Figure 6. Mean winter-to-spring and summer-to-fall climatological sea surface temperature and precipitation for the period 1971-1999. Observations (upper row from HadISST and GPCP), CMIP3 and CMIP5 multi-model means (middle row) and difference of multi-model mean minus observations (lower row). Red/blue shading denotes positive/negative differences in temperature while green/brown shading denotes positive/negative differences in precipitation (lower row); contour interval is 1°C (1 mm day^{-1}) for mean values for temperature (precipitation) and 0.3 mm day^{-1} (0.3°C) for the differences. Statistically significant differences calculated at the 5% level using a two-tailed Student's t-test are shaded. The thick black line is the 28.5°C isotherm which is used as a marker for the Western Hemisphere Warm Pool. Temperature (precipitation) fields have been interpolated to a common 5°×2.5° (1.5°×1.5°) grid. Figure adapted from Ruiz-Barradas et al (2013).



A comprehensive evaluation of temperature and precipitation extremes is given in Sun et al. (2014). Here we present additional analysis and literature review on the projected changes in extremes and partial attribution of differences between CMIP3 and CMIP5.

4.1. TEMPERATURE EXTREMES

Sillman et al., (2013a) provide an initial global evaluation of CMIP3 and CMIP5 models in their ability to simulate the climatology and trends for a range of extreme temperature indices as well as precipitation extreme indices. These indices include the maximum and minimum daily maximum and minimum temperatures, the number of cold nights/days and warm nights/days, the duration of warm/cold spells and the number of frost days, ice days, summer days and tropical nights. For the temperature indices they found that the performance of the CMIP5 ensemble is similar to the CMIP3 ensemble, but that the spread amongst the CMIP3 models is larger than that for CMIP5. This is despite the larger number of CMIP5 models examined (31 versus 18). Although Sillmann et al. (2013a) found no clear relationship between model resolution and performance, higher resolution CMIP5 models tend to be closer to the ensemble median for some temperature extreme indices, suggesting that resolution plays a role. Both sets of models are also able to simulate the long-term trends in temperature indices and, in the case of CMIP5 (and some CMIP3 models with volcanic forcing), simulate the short-term changes related to volcanic eruptions.

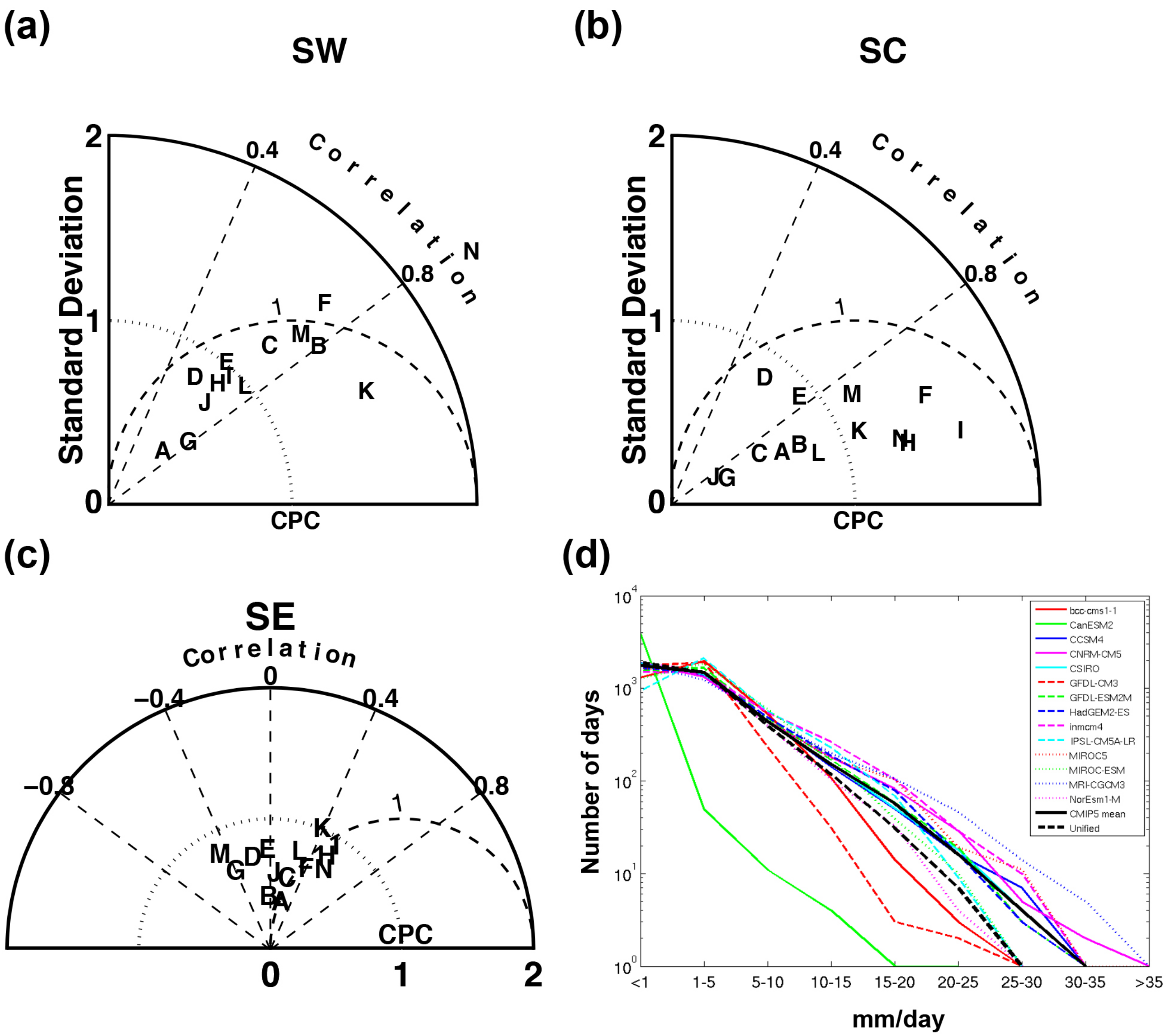
Projected changes in temperature extremes are expected across North America, with the number of 90-degree days increasing by 50-100% across the Midwest and Northeast by end of century in CMIP5 models (Maloney et al., 2014). Projected changes under the CMIP3 SRES scenarios are generally comparable to those of the CMIP5 RCP scenarios, although projected changes under RCP4.5 are larger than for SRES B1 despite similar radiative forcing (Sillman et al., 2013b). RCP8.5 changes exceed changes under any SRES scenario as expected because of the higher radiative forcing.

4.2. PRECIPITATION EXTREMES

An analysis of extreme precipitation in CMIP5 for North America by Sheffield et al. (2013a) (Fig. 7) showed that models tend to underestimate the frequency of precipitation extremes in the southeast and south central regions of the US, but overestimate extremes in the Southwest (although the number of events is generally small), and do better at representing the spatial variability in the southwest and south central regions, than the Southeast (Fig. 7). Some models do well across all regions, although it is unclear why. In the Northeast, mean precipitation biases are driven mainly by biases in heavy precipitation, and the set of models spans the observed heavy precipitation frequency. Sheffield et al. (2013a) noted that these biases could not be explained by individual model biases in the representation of extra-tropical cyclones.

DeAngelis et al. (2013) evaluated daily precipitation statistics for CMIP3 over North America and found robust underestimation across models (at least ¾ of models) of the intensity of heavy and extreme precipitation along the Pacific coast, southeastern United States, and southern Mexico. Overestimation of light precipitation offsets these biases to give generally realistic mean precipitation (Stephens et al., 2010). Biases appear to be associated with the representation of local forcing rather than large-scale circulation. DeAngelis et al. (unpublished manuscript) extended their original

Figure 7. Comparison of regional precipitation extremes between CMIP5 models and the CPC daily observational dataset. (Adapted from Sheffield et al., 2013a). (a)-(c) Taylor diagrams of the spatial pattern of annual number of days when precipitation > 10 mm day^{-1} over the (a) southwest, (b) south central, and (c) southeast US. The standard deviations have been normalized relative to the observed values. (d) Frequency distribution of daily average precipitation (in mm day^{-1}) for the Northeast region. Region definitions are given in Sheffield et al. (2013a). (A: CanESM2, B: CCSM4, C: GFDL-CM3, D: GFDL-ESM2G, E: GFDL-ESM2M, F: GISS-E2R, G: HadCM3, H: HadGEM2-CC, I: HadGEM2-ES, J: IPSL-CM5A-LR, K: MIROC4h, L: MIROC5, M: MPI-ESM-LR, N: MRI-CGCM3). Adapted from Sheffield et al. (2013a).



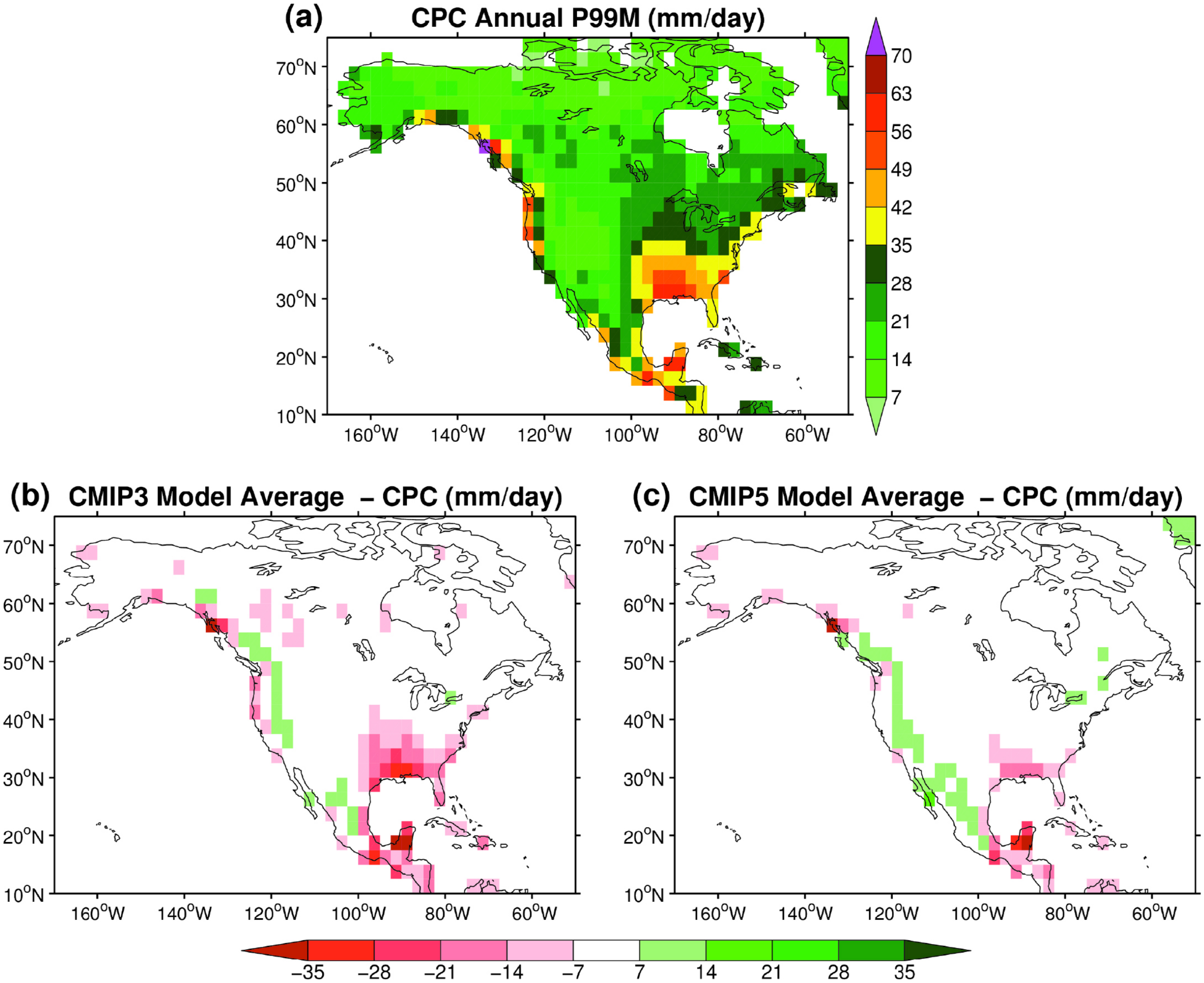
evaluation to CMIP5 models, which show smaller biases than CMIP3 in heavy precipitation over the southeastern United States (i.e., the magnitude of heavy to extreme daily precipitation is somewhat higher and more realistic; Fig. 8). They hypothesize that improvements in CMIP5 are due to increased resolution (although the increase in resolution from CMIP3 to CMIP5 is modest) and perhaps better parameterization of convective precipitation in some of the models, as the large-scale circulations during extreme events in the southeastern US are very similar between CMIP3 and CMIP5. For the summer in the Southeast, the circulation anomalies during extreme events are slightly stronger in CMIP5, and the relationship between horizontal resolution and biases is weaker than in winter. Thus, larger magnitudes of summer heavy precipitation in CMIP5 may be from a combination of higher resolution, better convective parameterizations, and slightly stronger circulation features.

Overall, the CMIP5 models tend to underestimate the frequency of heavy and extreme daily precipitation events for North America, but show a slight improvement over CMIP3 (Sillman et al., 2013a; DeAngelis et al. (unpublished manuscript)), especially in the southeastern US. Increasing resolution appears to help (e.g. Wehner et al. 2010; Li et al., 2011; DeAngelis et al., 2013; Kinter et al. 2013) but further investigation is required to determine whether there is a limit to the increase in skill. The general underestimation of extreme daily precipitation events by CMIP5 and CMIP3 models may be, in part, due to the underestimation of observed increasing trends over the past century (Min et al., 2011), which is notable in the Northeast, Midwest, and upper Great Plains (Karl et al., 2009; Walsh et al., 2014). A more comprehensive analysis of extremes across regions is required; some such analysis is described in Sun et al. (2014). Additional analysis should include attribution of events (e.g. to tropical cyclones (TC), extra-tropical storms, convective activity, topography, weather types, atmospheric circulation versus local processes, moisture sources, ENSO, etc.). For example, TCs contribute up to 30% of annual precipitation and up to 20-25% of extreme precipitation in the southeast (Knight and David, 2009) with implications for modeled extremes, as models severely underestimate the number and intensity of landfalling TCs. Furthermore, account needs to be taken of scale mismatches between models and observations that may hamper evaluations (see e.g. King et al. 2013), and the limitation of using small sample sizes (e.g. single model ensemble member) to detect and compare extreme events.

> The CMIP5 models tend to underestimate the frequency of heavy and extreme daily precipitation events for North America, but show a slight improvement over CMIP3, especially in the southeastern US.

Projected future changes in heavy precipitation in CMIP5, over North America (DeAngelis and Broccoli, 2013; Maloney et al., 2014) and globally (Scoccimarro et al., 2013), appear to be qualitatively consistent with earlier studies using CMIP3 data (DeAngelis et al., 2013). For CMIP5, DeAngelis and Broccoli (2013) note that heavy precipitation (defined as the precipitation from the 99th percentile and above) increases over much of North America by the end of the 21st century, with generally larger increases in higher latitudes and near the coasts, and mainly during the winter. The domain-averaged results from Maloney et al. (2013) indicate an increase of 20-30% for the late 21st century. Summertime increases are generally confined to very high latitudes, with only small

Figure 8. Comparison of annual extreme precipitation (average precipitation over all days when precipitation equals or exceeds the 99th percentile, P99M) from the CPC daily dataset and the CMIP3 and CMIP5 multimodel ensembles. Adapted from DeAngelis et al. (unpublished manuscript).



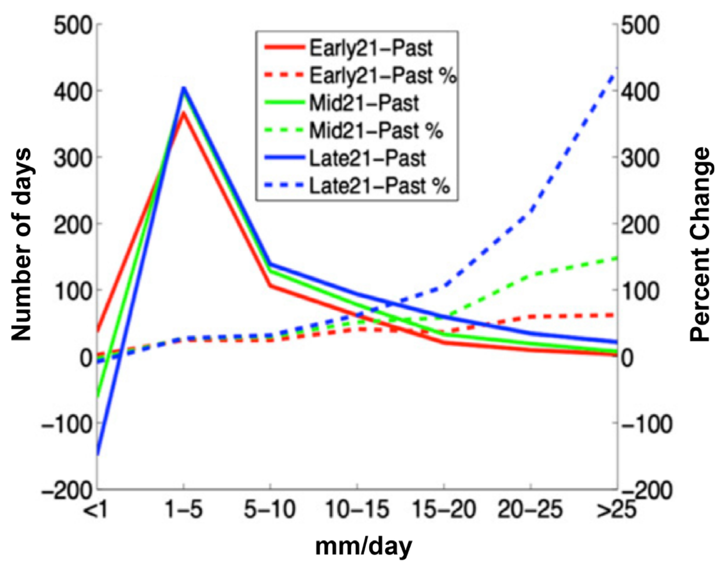
changes elsewhere. For low latitude regions, DeAngelis and Broccoli (2013) found that heavy precipitation decreases in intensity and frequency over regions such as western Mexico and the adjacent Pacific Ocean during winter and spring.

For the northeast US, the changes are on the order of 7%/K^{-1} in the winter (3-5%/K^{-1} in summer) in mean 99th percentile precipitation, which translates to about a 20-35% absolute increase across models (DeAngelis and Broccoli, 2013). Scoccimarro et al. (2013) indicate a 20-30% increase in moderate (75th percentile) to extreme (99.9th percentile) winter precipitation in CMIP5 for the eastern US (Florida to southeast Canada) when comparing the periods 1965-2005 and 2061-2100 for RCP8.5. The results from Maloney et al. (2013) indicate changes on the order of 4-5 times for events > 25 mm/day^{-1} (equivalent to about the 99th percentile) for the northeast US. However, it should be noted that this relates to a relatively small part of the domain (northeast US) where some of the largest changes are projected, and that the sample size (number of heavy events) is relatively small (change from 2 events to 8 events).

Overall, projected changes by the end of the 21st century (2070-2099) in moderate to extreme precipitation events show a 20-30% increase over the US, which is similar to CMIP3, with much higher increases in the northeastern US, especially in winter.

Overall, projected changes by the end of the 21st century (2070-2099) in moderate to extreme events show a 20-30% increase over the US, which is similar to CMIP3, with much higher increases in the northeastern US, especially in winter (Fig. 9). The wintertime changes are consistent with an upper limit for daily precipitation extremes as expected from the Clausius-Clapeyron relationship (Kharin et al., 2013). Min et al. (2011) note that a subset of CMIP3 models underestimate the observed wetting of extreme precipitation across the northern hemisphere during the second half of the 20th century, suggesting that the projections of future increases may be underestimated, although their analysis did not discriminate between warm and cool season changes, and other methodological choices may have muted the model trends. A more detailed regional analysis of the representation of extreme precipitation events and projected changes is required, including the mechanisms of projected changes and the relationship with warming.

Figure 9. Difference in the number of precipitation days and percentage change for each amount bin between the 2009-2038, 2038-2068, and 2069-2098 and the historical 1979-2004 period for the northeast land region define in Maloney et al., (2013). (Adapted from Maloney et al., 2013).



5.1. EL NIÑO AND THE SOUTHERN OSCILLATION (ENSO)

The representation of ENSO in CMIP5 has improved over CMIP3.

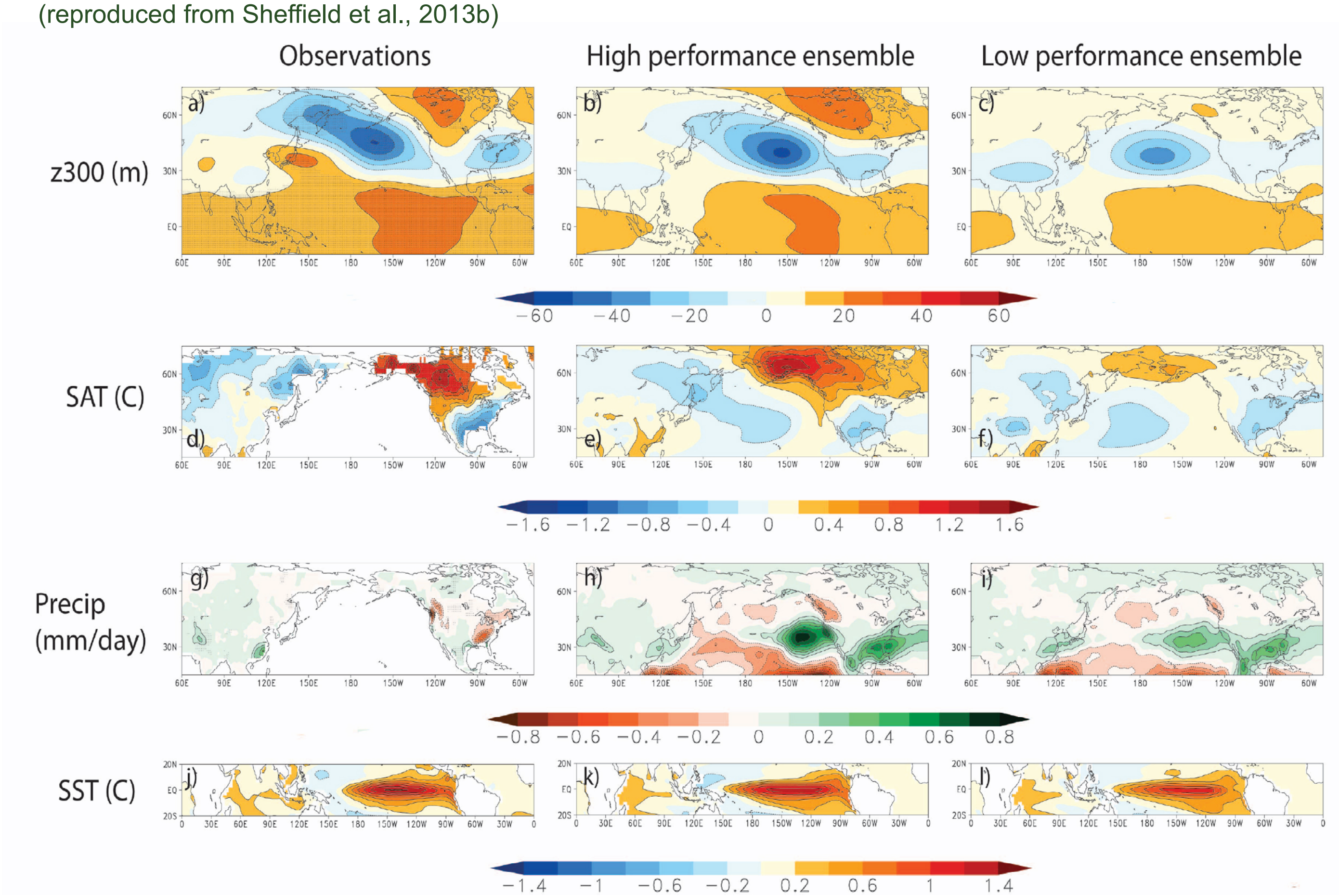
Guilyardi et al. (2012) and Bellenger et al. (2013) showed that, in general, the representation of ENSO in CMIP5 has improved over CMIP3. In particular, there is a 30% reduction of the cold bias in the west Pacific, and the large inter-model spread in the amplitude of ENSO has been reduced by a factor of 2, although it should be noted that model control simulations suggest multi-decadal to centennial modulations of ENSO amplitude (e.g. Wittenberg et al., 2009). Kim and Yu (2012) also compared ENSO simulations between CMIP3 and CMIP5 and concluded that the CMIP5 models show less inter-model spread, with the CMIP3 models more clearly separated into a group that produces strong ENSO intensities and a group that produces weak ENSO intensities, whereas the CMIP5 model ensemble has converged to a single group that is closer to observations. They find the reduction in spread to be different for the two types of ENSO (Eastern Pacific, EP; and Central Pacific, CP; Yu and Kao 2007 and Kao and Yu 2009). The reduction is particularly significant for the EP type, whose generation depends more on thermocline variations over the cold tongue region and their coupling with the atmosphere. The generation of the CP type has been suggested to be less sensitive to thermocline variations (e.g., Kao and Yu 2009, Yu et al. 2010, Yu and Kim 2011). It is possible that efforts aimed at improving model parameterizations (such as those associated with cumulus convection and ocean mixing) have enabled the CMIP5 models to more realistically simulate ocean-atmosphere coupling over the cold tongue region and the resulting EP ENSO variability in the region. Furthermore, the ENSO life cycle, as represented by the seasonal phase locking and the location of surface temperature anomalies is slightly improved since CMIP3 (Guilyardi et al. 2009). Sheffield et al. (2013b) examined the phase locking in CMIP5 models and showed that 68% of El Niño and 65% of La Niña episodes have peak amplitudes in fall or winter in the models, compared to 90% and 89%, respectively, in the observations, with implications for the representation of teleconnections with land surface climate. They note, however, that several of the models (CanESM2, CNRM-CM5, HadCM3 and NorESM1-M) do have fall/winter peak frequencies exceeding 80% for both El Niño and La Niña episodes. Bellenger et al. (2013) also note that fundamental characteristics of ENSO, such as its spectrum and central Pacific precipitation anomalies, are still poorly represented.

CMIP5 models with better ENSO performance did well in representing teleconnection patterns in temperature and precipitation over North America.

Sheffield et al. (2013b) examined teleconnections with North American climate in CMIP5 models, and noted that model skill was related to mean state biases in SST (Figure 10). Performance based on the representation of composite (El Niño/ La Niña) 300hPa geopotential height patterns versus NCEP/NCAR reanalysis indicated that the better-performing models (top half of models) did well in representing teleconnection patterns in T and P (with the exception being the failure to capture the negative P anomaly in the Tennessee and Ohio valleys). The lower-performance models had much weaker teleconnection patterns and a westward shift in height anomalies. This may be explained by stronger climatological (rather than interannual variability) SST cool biases (> 1.5C) in the low-performance models that drive lower P/convection, and weaker and more westward-shifted teleconnections. Langenbrunner and Neelin (2013) note that there has been little improvement in model performance for P teleconnections relative to CMIP3.

Furthermore, there is evidence that model errors are related to the different teleconnection patterns associated with the two types of ENSO (EP and CP), notably for winter temperatures (Sheffield et al., 2013b). Zou et al. (2014) show that CMIP5 models are more capable of simulating the El Niño impact on US winter climate for the traditional Eastern-Pacific type but not for the emerging Central-Pacific type. They offer the following explanation for this: during the EP El Niño, the largest SST anomalies are located in the eastern equatorial Pacific and influence the strength of the Walker circulation to give rise to basin-wide outgoing longwave radiation (OLR) anomalies. The modeled atmospheric responses to the EP El Niños are thus not sensitive to the SST anomaly structure and can be well simulated by most of the CMIP5 models. In contrast, the SST anomalies of the CP El Niño are in the central equatorial

Figure 10. Composites of (a)–(c) 300-hPa height (z300; m), (d)–(f) surface air temperature (SAT) (°C), (g)–(i) precipitation (mm day^{-1}), and (j)–(l) SST (oC) anomalies during DJF El Niño episodes (left) in observations and in (middle) high and (right) low performance CMIP5 ensembles. The observational SAT and precipitation composites are based on the CRU TS3.1 land near-surface temperature and precipitation datasets for 1901–2009. The z300, SAT, and precipitation composites are normalized by the Niño-3.4 SST anomaly. Stippling in the observed (a) z300, (d) SAT, and (g) precipitation composites indicates anomalies that are statistically significant at the 5% level. (reproduced from Sheffield et al., 2013b)



Pacific and induce only local OLR anomalies to the west of the SST anomalies. The modeled atmospheric responses to the CP El Niño are different among the models depending on the simulated magnitudes and locations of the CP El Niño SST anomalies. Kug et al. (2012) showed that CMIP5 models on average are slightly better at representing the two types of ENSO and that this was caused by the precipitation response to cold tongue El Niño, which is closely related to a dry bias over the equatorial eastern Pacific. Yu and Zou (2013) showed that the CP El Niño has a tendency to enhance the dry impacts and weaken the wet impacts produced by the EP El Niño on US winter precipitation.

Projected changes in ENSO are difficult to detect because of the large internal variability (Wittenberg, 2009; Maloney et al., 2014). However, changes in the base state of tropical SSTs may affect the balance of ENSO cycle feedbacks, changing the average amplitude of events (Guilyardi et al., 2012), and impact on ENSO teleconnections (Maloney et al., 2014). Projected changes for CMIP5 models under RCP8.5 show no change in ENSO event frequency and a slight increase in ENSO amplitude, although the inter-model spread is high enough, relative to the magnitude of the increase, that these changes are not robust (Maloney et al., 2014). Comparison of CMIP3 and CMIP5 (Guilyardi et al., 2012) shows that both sets of projections display a similar diverse range of changes in ENSO characteristics with the different model responses due to differences in the balance of feedbacks in each model. A robust finding, in CMIP5 at least (because of improved model historical performance and lower inter-model spread in the projections relative to CMIP3), may be an increase and decrease, respectively, of the intensity of the CP and EP ENSO types under RCP4.5 (Kim and Yu, 2012). Teleconnections with North American temperature and precipitation in CMIP5 models were found to strengthen, associated with increases in tropical convective forcing driven by the overall increase in SSTs (Maloney et al., 2014), but may also be related to the changes in the intensity of CP and EP ENSO types (Zou et al., 2014). In particular, the emerging CP El Niño may produce an overall drying effect on the US winter, particularly over the Ohio–Mississippi Valley, Pacific Northwest and Southeast (Yu and Zou, 2013).

5.2. PACIFIC DECADAL VARIABILITY

Sheffield et al. (2013b) showed that CMIP5 models replicate the basic horseshoe SST pattern of the Pacific Decadal Oscillation (PDO; defined as the leading empirical orthogonal function of extended winter (November–April) monthly-mean SST anomalies in the North Pacific poleward of 20oN), which is similar to CMIP3 models (Oshima and Tanimoto, 2009; Furtado et al., 2011). However, the CMIP5 models show a westward shift of the North Pacific center of action in models with respect to observations. Figure 11 shows an updated analysis comparing the PDO North Pacific SST patterns in CMIP3 and CMIP5 using Taylor diagrams. The left figure is taken from Oshima and Tanimoto (2009) for CMIP3 and a similar calculation is made on the right with CMIP5 models from Sheffield et al (2013b). The two PDO definitions are slightly different so the comparison is not exact, but it does suggest that the CMIP5 models do slightly better, with higher pattern correlations between model and observed regressions. The CMIP3 PDO pattern from Oshima and Tanimoto (2009) is based on linear regression against the mean SST over a box in the central North Pacific. The CMIP5 PDO pattern definition is the EOF-based regression pattern definition from Sheffield et al. (2013b). The CMIP5 regression was performed with unfiltered monthly

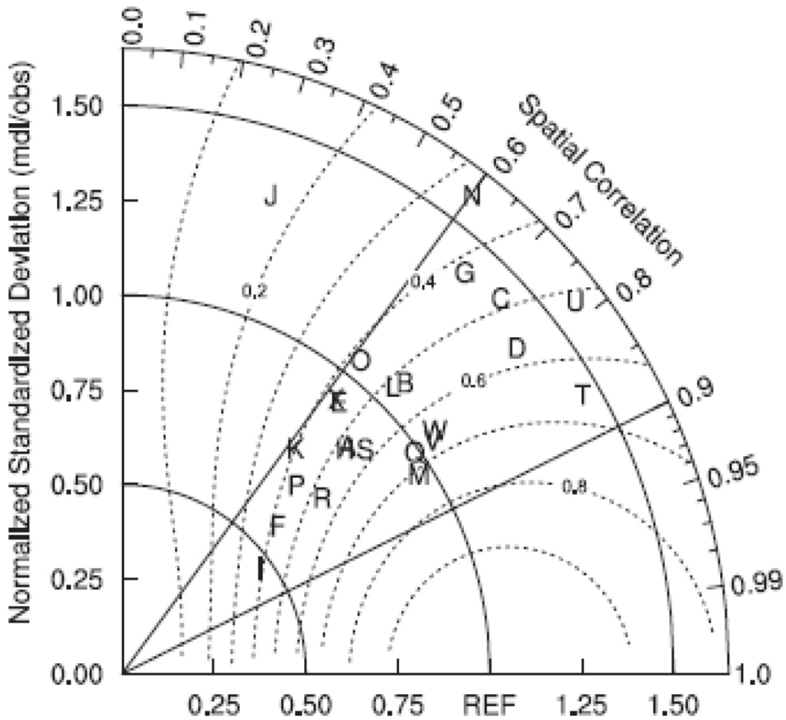
SST data, whereas Oshima and Tanimoto (2009) first applied a 5-year running mean smoother to the time series. It is unlikely that either difference in definition would have a significant impact on the results, because the leading EOF of North Pacific SST is dominated by variability in the region that Oshima and Tanimoto (2009) use, and the additional interannual variability in the CMIP5 definition would not impact the pattern much. Although the CMIP5 pattern correlation appears to be better than in CMIP3, the spatial standard deviation is overestimated. In the CMIP5 regression, the SST anomalies in the Kuroshio-Oyashio extension (KOE) region are too low, which exaggerates the SST anomaly difference between the eastern and western North Pacific. It is unclear from Oshima and Tanimoto (2009) whether this is a common feature in CMIP3 models, but this general feature is discernible in the CMIP3 PDO analysis of Overland and Wang (2007) (http://www.pmel.noaa.gov/foci/publications/2007/over0633.pdf). Therefore the hypothesis is that more models are getting the basic PDO structure correct, but these models generally suffer from a bias in exaggerated SST variability over the KOE region.

CMIP5 models do well in representing the Pacific Decadal Oscillation's influence on North American surface air temperature while not reproducing the precipitation teleconnection patterns over large parts of North America. Performance for precipitation is generally better in western North America (especially the Southwest), and worse in the east.

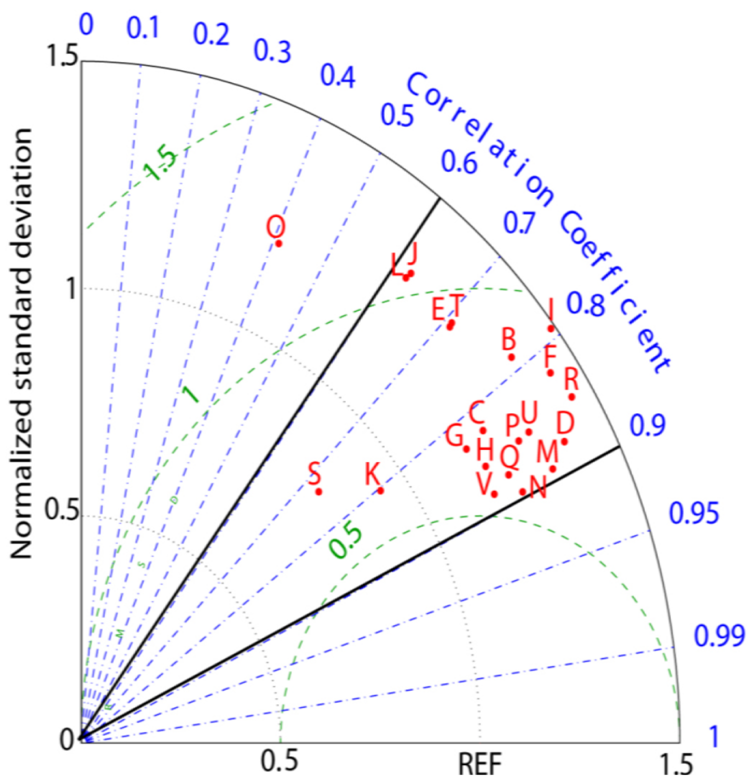
In terms of PDO teleconnections, Sheffield et al. (2013b) showed that the CMIP5 models do well in representing the PDO influence on North American surface air temperature, with positive (negative) anomalies in the Northwest (Southeast) during the positive phase of the PDO. However, the models cannot reproduce the precipitation teleconnection patterns over large parts of North America, with performance generally better in western North America (especially the Southwest), and worse in the east. In particular, Sheffield et al. (2013b) and Polade et al. (2013) note the deficiency over the Ohio Valley. There is little literature on comparisons between CMIP3 and CMIP5 for PDO teleconnections. Leinert (2011) discussed the results from CMIP3 and concluded that the models do reasonably well at reproducing the influence of both the tropical-Pacific-related and the extratropical part of the PDO on North American surface temperature, but only some of the influence of the PDO on precipitation that is mainly related to the tropical Pacific connection. Leinert (2011) shows that CMIP3 models tend to have a delayed response of the North Pacific to ENSO forcing due to model biases in the mixed layer depth and air-sea feedbacks, and they tend to overestimate the lower frequency variability due to model errors in the tropics and extra-tropics. The Polade et al. (2013) study notes that 9 of the 14 CMIP5 models evaluated showed improvement in simulating the PDO-like mode and its North American precipitation teleconnections, which they attribute generally to improved resolution and model physics. Despite the generally good performance for temperature teleconnections and possible improvements for precipitation in CMIP5, the deficiencies in the observational record in representing the full spectrum of Pacific variability should be noted. Any conclusions on model performance must be made with respect to the possible non-stationarity of PDO patterns and the modulation of teleconnections depending on the phase of ENSO (e.g. Wise et al., 2014; Kam et al., 2014).

Figure 11. Evaluation of the PDO sea surface temperature (SST) patterns. Taylor diagrams for CMIP3 (Oshima and Tanimoto 2009) and CMIP5 (Sheffield et al. 2013b) PDO SST patterns in various climate models, each designated by a letter (the letters are unrelated in the two plots). Although the PDO definitions vary slightly in the two studies, the CMIP5 models seem to perform somewhat better than CMIP3 models, with most pattern correlations less than 0.8 in CMIP3 but above 0.8 in CMIP5 models.

a) CMIP3



b) CMIP5



In terms of future projections, there has been no systematic comparison of CMIP3 and CMIP5 results for either a shift in the spatial pattern or intensity of the PDO, or its teleconnections with North American climate. Lapp et al. (2012) showed a weak shift in the CMIP3 MME mean to more occurrences of the negative phase of the PDO during the 21st century for three SRES scenarios (B1, A1B and A2), although the models were split among those showing a shift towards more negative conditions and those towards more positive conditions. For CMIP5 models, there is no equivalent study of changes in the PDO. However, there is some consistency with the results of Yeh et al. (2013), who showed that the negative phase of the PDO shifts to cooler SSTs for 9 CMIP5 models for RCP4.5 for 2100-2200. This shift may be associated with a shift towards more CP type El Niño events due to a shallowing of the thermocline (Yeh et al., 2009). For teleconnections, there is a statistically significant strengthening in CMIP5 models of teleconnections with North American temperature for all seasons by the end of the century for the model ensemble mean with the largest strengthening over western North America, which is consistent with an increase in tropical convective forcing (Maloney et al., 2014).

5.3. ATLANTIC MULTIDECADAL VARIABILITY

The representation of the AMV has generally improved in CMIP5 models (Sheffield et al., 2013b; Zhang and Wang 2013; Wang and Zhang 2013) compared to CMIP3 (Medhaug and Furevik, 2011), particularly after 1960. Sheffield et al. (2013b) speculated that this might be due to higher resolution models, improved parameterizations and the addition of time-evolving land cover. Furthermore, the representation of AMV in terms of the standard deviations (0.09 to 0.19°C) is comparable to, or slightly weaker than, the observations (~0.18°C) (Zhang and Wang 2013), which is an improvement from CMIP3 models (Ting et al., 2009). The models have also improved in terms of persistence (defined as the maximum time lag when the autocorrelation first crosses the significance line at the 90% level), which varies from 5 to 25 years in CMIP5, with an average of about 12 years, compared to an average of about 5 years in CMIP3 (Medhaug and Furevik 2011). Similar to the PDO, any conclusions must be tempered by the relatively short observational record with respect to the time scale of variability, although proxy records (e.g. Knudsen et al., 2011) and long-term climate simulations (Knight et al., 2005) support the existence of AMV over the past millennia.

> The representation of the Atlantic Multidecadal Variability has generally improved in CMIP5 models compared to CMIP3 with the improvement potentially due to higher resolution, improved parameterizations and time-evolving land cover.

The mechanism for AMV continues to be unclear despite its importance, and there exist the possibilities that it is forced to some degree or it is just noise. One potential mechanism is the role of variations in the Atlantic Meridional Overturning Circulation (AMOC) and associated heat transport fluctuations (Delworth and Mann 2000; Knight et al. 2005). The analysis of historical simulations for most of the CMIP5 models shows that an interaction between the AMV and AMOC can produce a multidecadal oscillation in the North Atlantic Ocean (Zhang and Wang 2013) consistent with the delayed advective oscillation mechanism proposed by Lee and Wang (2010). Modeling studies have also suggested that solar variability and/or volcanoes are important (Hansen et al. 2005; Ottera et al. 2010), or that aerosols can be a primary driver (Dunstone et al., 2013; Booth et al. 2012), although the robustness of the latter has been questioned because of discrepancies in the representation of observed ocean states (Zhang et al. 2013). A recent observational study suggests that a positive feedback between SST and dust aerosols in the North Atlantic via Sahel rainfall variability may be a mechanism (Wang et al. 2012).

More recent work has shown that the speed up (slow down) of the AMOC favors the generation of the warm (cold) phase of the AMV via the anomalous northward (southward) transport of heat in the upper ocean, which conversely leads to a weakening (strengthening) of the AMOC through changes in the meridional density gradient after a delayed period of ocean adjustment (Zhang and Wang 2013; Wang and Zhang 2013). Kumar et al. (2014) evaluated North Atlantic variability (and connections with the “warming hole” in the central US) and suggested that anthropogenic forcing plays a role in the AMV by increasing its persistence slightly more than expected based on CMIP5 pre-industrial control simulations. The detailed mechanisms of the AMV and the separation of natural and anthropogenic components remain challenging issues.

Several studies have highlighted the importance of the AMV in forcing precipitation variability over North America (e.g. Enfield et al. 2001, Sutton and Hodson 2005, Wang et al. 2006, Schubert et al. 2009, Nigam et al. 2011). However, the CMIP5 models do not simulate the observed teleconnection patterns well (Sheffield et al., 2013b), in part because of the poor representation of the spatial footprint of AMV SST variability -- a situation that has not improved since CMIP3 (Ruiz-Barradas et al. 2013). The AMO has a much stronger and clearer signal in precipitation in the southwest US and northwest Mexico (Hu and Feng, 2008) and so examining this relationship may provide a clearer way to examine how well the models simulate AMV remote effects on North America. The poor performance is also related to the effect of AMV on the lower level circulation, which modulates the Great Plains low-level jet (LLJ) and the convergence/divergence of moisture fluxes. AMV has a strong effect on upper troposphere circulation as well, by creating a favorable upper level front that nurtures convection and precipitation in the central US as shown in the modeling study of Hu et al. (2011). This is consistent with the result that the LLJ is less of a key factor for future summer precipitation in the Great Plains (see below). Improved understanding of teleconnections may be sought by examining the upper troposphere circulation in the models, and whether this is related to skill in simulating summer precipitation in the Great Plains. Furthermore, AMV effects on North American summer precipitation appear to be primarily to change the interannual variability of rainfall as opposed to providing a persistent decadal forcing, as suggested by Hu and Feng (2012).

There appears to be no literature on future projected changes in AMV in CMIP3 and CMIP5 simulations. In terms of future research, analysis of long control runs (1000 years or longer) would be helpful in clarifying the role of anthropogenic forcing versus natural variability for AMV and the relationship with the AMOC. Improved understanding of teleconnections may be sought by examining the upper troposphere circulation in the models, and whether this is related to skill in simulating summer precipitation in the Great Plains. Furthermore, AMV has a much stronger and clearer signal in precipitation in the southwest US and northwest Mexico (Hu and Feng, 2008), and examining this relationship may provide a clearer way to examine how well the models simulate AMV remote effects on NA.

6.1. PROJECTIONS OF DRYING FOR THE CARIBBEAN AND MEXICO

Drying of sub-tropical regions is evident in CMIP3 and CMIP5 model projections, in particular for the Caribbean and parts of Mexico. Specifically, Maloney et al. (2013) found that CMIP5 model projections showed reduced summertime precipitation in the east Pacific warm pool and the Caribbean, with agreement among all models for several regions including the major Caribbean islands, the Yucatan Peninsula and southwestern Mexico. This intermodel agreement is even higher than for CMIP3 (e.g. Neelin et al. 2006) and may be related to improved representation of regional precipitation. Ryu and Hayhoe (2013) showed that precipitation biases over the Caribbean region have decreased in CMIP5, and therefore more models are able to simulate the two rainy peaks that characterize the annual cycle of regional precipitation. Furthermore, the better models realistically reproduce summer changes in the meridional gradient of SST and westward extension of the North Atlantic Sub-tropical High (NASH). Further research is needed to identify the reasons for the drying signal in the region and whether this is related to model performance in representing the seasonal dynamics of precipitation and its relationships with the changes in SSTs and the NASH.

6.2. NORTH AMERICAN MONSOON

The seasonal cycle of monthly precipitation in the core monsoon region of northwest Mexico (23.875°-28.875°N, 108.875°-104.875°W) has been evaluated in CMIP5 models and compared to CMIP3 in Sheffield et al. (2013a). This region was selected specifically for its uniformity in the seasonal cycle of precipitation, reducing spatial variability in the calculation of the mean. This region is also at the core of the monsoon, which separates model errors in determining the northern boundary of the monsoon from errors in monsoon seasonality and intensity. The large majority of models are biased wet, with an average bias of 51.3%. Positive biases occur throughout the year with the largest biases seen from July through December (Geil et al. 2013). Similar monthly errors are seen in CMIP3 models (Liang et al., 2008), indicating that there has been no improvement in the magnitude of the simulated annual cycle of monthly precipitation. On the other hand, there does seem to be improvement in the timing of seasonal precipitation shifts, with 13 out of 21 (62%) CMIP5 models having a phase lag of zero months as compared to 6 out of 17 (35%) CMIP3 models in Liang et al. (2008).

> There has been no improvement in the magnitude of the simulated annual cycle of monthly precipitation for the North American monsoon, although there appears to be an improvement in the timing of seasonal precipitation shifts.

Several studies have shown that the North American monsoon region is expected to become drier on an annual basis over the 21st century (Maloney et al., 2013; Cook and Seager, 2013). However, this decrease is not evenly distributed throughout the year, with the majority of models showing decreases from November through August and increases for September through October. This increase in late season rainfall suggests that the monsoon is arriving later and ending later than during the historical period, although these results should be interpreted with caution as the models have a wet bias in the fall during the historical period (Geil et al. 2013) which may simply be growing in time. On the other hand, a drier July and August is particularly concerning since the result is more robust in terms of model consensus on the sign of the change and these months tend to be overestimated for the model's historical period (Geil et al., 2013; Sheffield et al., 2013a). Analysis of CMIP3 data suggest that, in general, precipitation in monsoon regimes such as the SW will be redistributed in the future with reduced spring rainfall and increased late season rainfall (Seth et al. 2010, 2011; Biasutti and Sobel 2009). The CMIP5 results are qualitatively similar to those from CMIP3 with high model agreement on reduced winter through early rainy season rainfall (Dec-Jul) but less of a consensus on increased late season rainfall (Maloney et al., 2013).

Overall, current analyses indicate early season drying and late season wetting that is seen in the monsoons of other regions of the Americas and Africa although the changes are less clear over Asia (Seth et al., 2013). However, these CMIP5 results and previous CMIP3 studies are generally based on analysis of monthly data and so the question of whether the SW monsoon (and other global monsoons) will have an increasingly later onset is difficult to determine. Sheffield et al. (2013a) note that more CMIP5 models show zero lag in the start of the monsoon than CMIP3 models. These zero-lag (unbiased) models were analyzed for future changes in the monsoon by Maloney et al. (2013) at monthly scale, which showed no evidence of a shift in the monsoon, based on this monthly analysis. A daily analysis of the historic CMIP5 simulations by Geil et al. (2013) showed that the models on average tend to start early compared to the daily TRMM Multi-satellite Precipitation Analysis (TMPA) satellite precipitation dataset for the core North American Monsoon (NAM) region, which is not reflected in the monthly analysis of Sheffield et al. (2013a).

An updated analysis based on daily precipitation data from 18 CMIP5 models (Y. Serra, unpublished results; Table 1) indicates a slight shift to a later onset date (6-8 days in the mean or median) in the SW monsoon by 2070-2099 for RCP8.5, but no late retreat date, so the duration of the monsoon is shortened somewhat. The change in retreat date is much harder to determine because of the higher variability in the models at the end of the season (Fig. 11), which is similar to the historical period (Geil et al., 2013). Note that the standard deviation of the MME mean is larger than the shift for both the onset and retreat, that the observed duration as estimated from the TMPA satellite data is 101 days, and that the models tend to start earlier and end earlier in comparison.

Further work is required on several fronts, including 1) to understand if the changes for broader monsoon regions are similar, such as into the southwest US and down into Mexico, and whether these results are similar to CMIP3, 2) to determine the physical reason for the shift in the monsoon timing and therefore its robustness, and 3) to understand changes in seasonal precipitation and its timing in other regions of North America. For the second area where extra work is required, the locations of the North Atlantic and North Pacific Sub-tropical Highs (NASH; NPSH) are relevant for the historical simulations and may help determine the shift. For example, a change in their positions and structure associated with warmer SSTs may be relevant.

Table 1. CMIP5 projected changes in the relative onset and retreat of the monsoon in the core region in terms of the median of the model ensemble.

	Historic (1979-2005)	RCP8.5 (2070-2099)	Difference (RCP8.5 – Hist.)
Onset date	8 Jun +- 3 days	16 Jun +- 4 days	8 +- 5 days *
Retreat date	25 Sep +- 3 days	24 Sep +- 3 days	0 +- 4 days
Monsoon duration	109 +- 4 days	101 +- 4 days	-8 +- 6 days *

* statistically significant

6.3. FUTURE PRECIPITATION IN THE SOUTHWEST

All generations of CMIP model simulations agree on the basic response of the mean hydrological cycle to global warming. For example, Maloney et al. (2013) examined changes in precipitation patterns over North America based on CMIP5 models and showed that the overall patterns of change are similar to CMIP3, with large scale increases in mid- to high-latitudes and decreases in the sub-tropics. This is the generic ‘wet-get-wetter, dry-get-drier’ change in which the subtropics dry and the deep tropical and mid- to high-latitude regions get wetter. While Mexico fits within the ‘dry-get-drier’ region, the southwest US has, within the history of model simulations, always been at the poleward edge of the region of subtropical drying with areas of the US to the north projected to get wetter. One notable difference between the CMIP3 and CMIP5 results is that the boundary between these two regions has shifted southward slightly, resulting in increases in projected precipitation over parts of California (Neelin et al., 2013).

Figure 12 examines this in more detail at regional and seasonal scales using a larger set of CMIP5 models (28) for 1900 to 2100. For the wintertime, the CMIP5 model projections of precipitation across the southwest vary in space and by season.

> One notable difference between the CMIP3 and CMIP5 results is that the boundary between the projected ‘wet-get-wetter’ and ‘dry-get-drier’ regions has shifted southward slightly, resulting in increases in projected precipitation over parts of California.

> When projected changes in precipitation minus surface evapotranspiration (P-ET, i.e. the net flux of water at the Earth’s land surface which sustains soil moisture and surface and subsurface flows) are examined, the boundary between drying to the south and wetting to the north is farther north than that for precipitation alone.

In the mid-winter season, almost all of California, Nevada, Utah and Colorado are projected to have increased precipitation, although there is no consensus on the changes in the southern part of this region, while Mexico, most of Arizona, New Mexico and Texas have decreased precipitation. In the spring season, almost all regions of southwest North America are projected to dry except for Texas, and the models project a notable reduction of precipitation across California, Nevada, Colorado and Utah. Projections for the summer season changes in the southwest are quite weak and not consistent across models. For the fall season, the models project decreases in precipitation for California, Mexico and southern and western parts of Nevada, Arizona, New Mexico and Texas, but with weaker model agreement, while there are projected declines in precipitation across all of southwest North America with model agreement in the regions of strongest drying.

When projected changes in precipitation minus surface evapotranspiration ($P\text{-}E_T$, i.e. the net flux of water at the Earth’s land surface which sustains soil moisture and surface and subsurface flows) are examined, the boundary between drying to the south and wetting to the north is farther north than that for precipitation alone. This is because warming drives increased evapotranspiration so that more of what falls as precipitation is recycled to the atmosphere with less available to replenish soil moisture and streamflow.

Winter precipitation is very important to water resources in the southwest US Winter precipitation is primarily delivered by storms propagating into

Figure 12. Multimodel mean seasonal changes in P and P-E between 1979-2005 and 2021-2040 for the RCP8.5 scenario based on 37 CMIP5 models. The stippling is where 3/4 or more of the model ensemble means agree with the sign of the multimodel mean change.

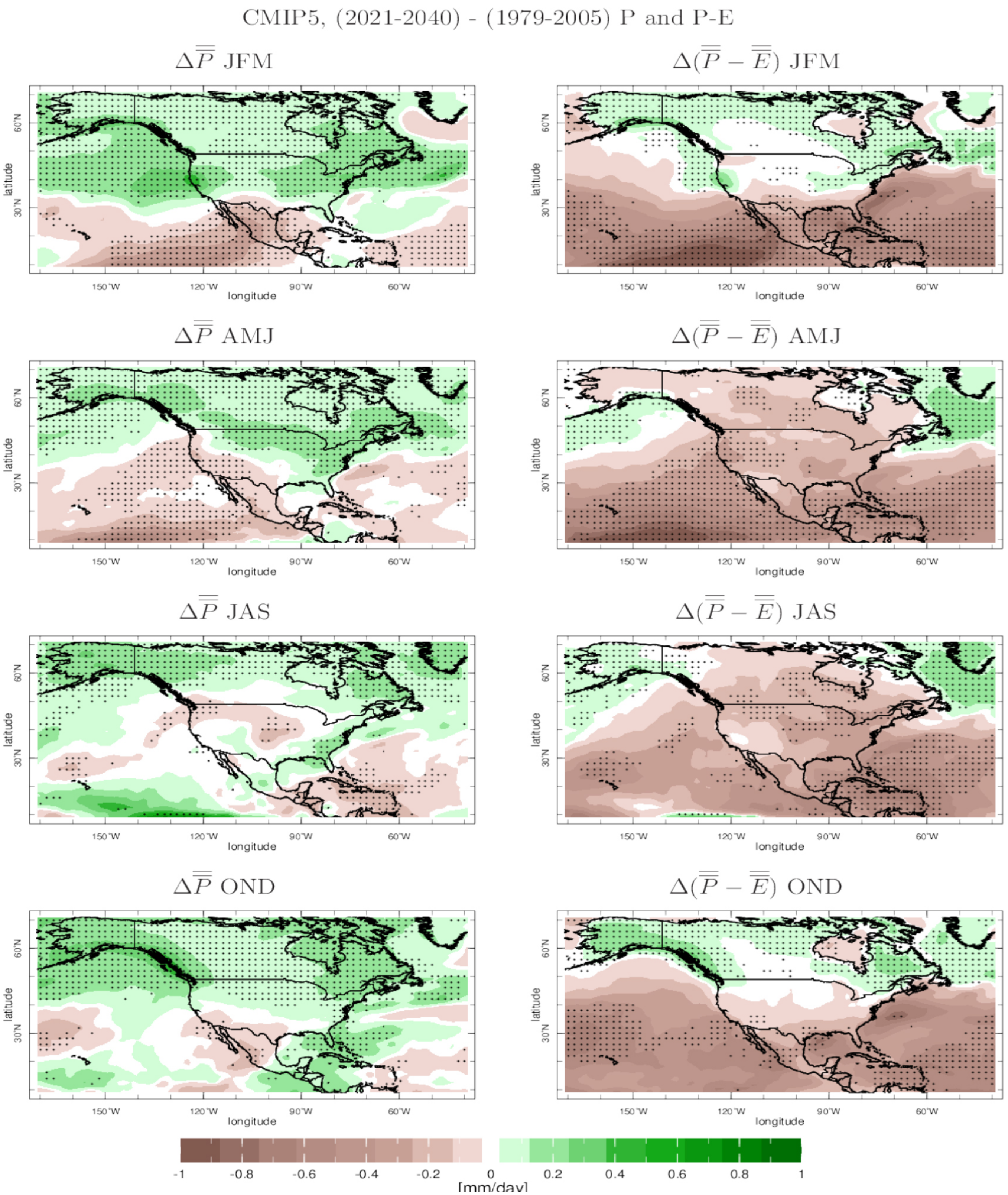
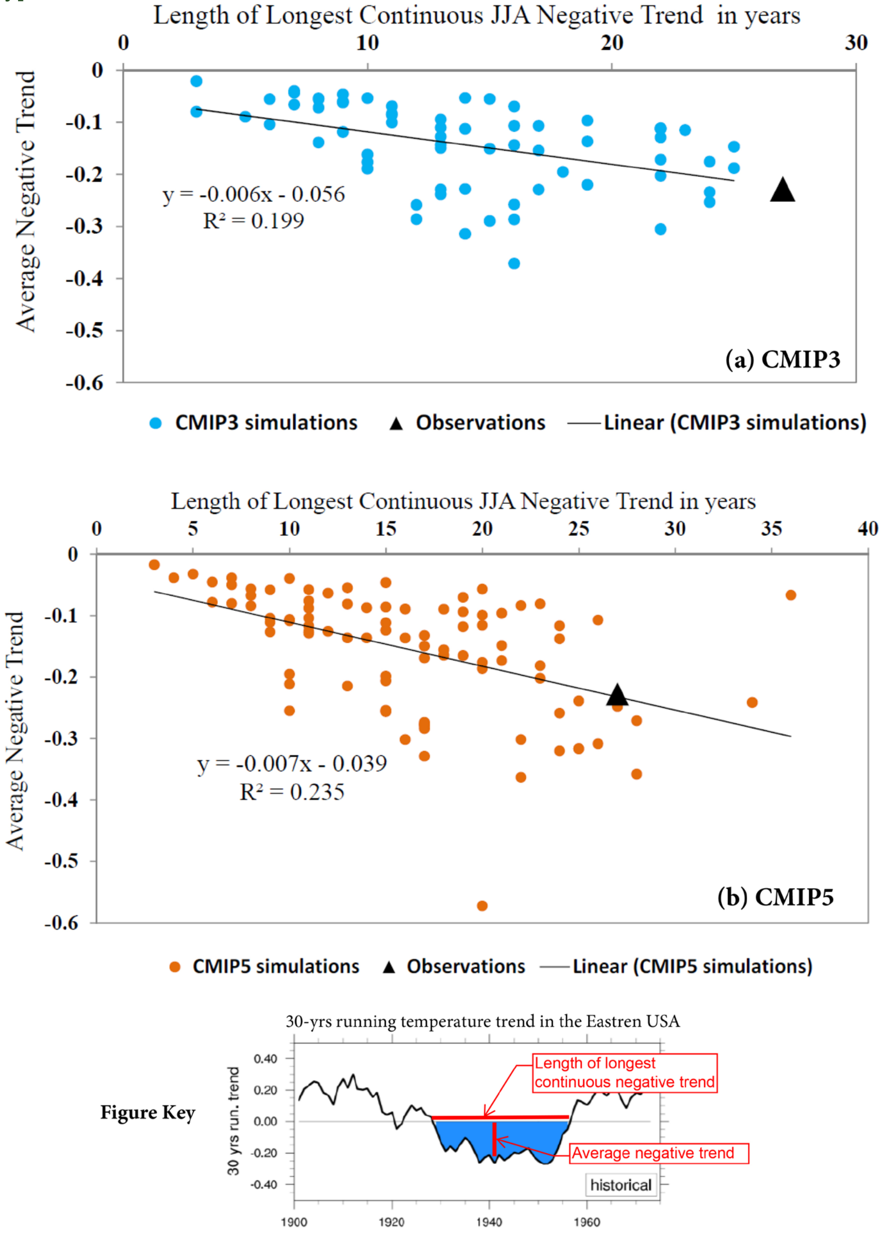


Figure 13. Evaluation of the simulation of a “warming hole” as a decadal climate variability signal in (a) CMIP3 and (b) CMIP5 models. Thirty-year running temperature trends are calculated for each year from 1901 to 1975, i.e., a trend at 1901 represents the trend for 1901-1930, and similarly a trend at 1975 represent the trend for 1975 to 2004. The length of the longest continuous negative temperature trend and the average negative temperature trend (ºC/decade) during that period are shown (see Figure Key). In the model simulations the negative temperature trend can occur during any part of the 20th century and does not necessarily coincide with the observed negative temperature trend period. We use 66 historical climate simulations from 22 CMIP3 models, and 92 simulations from 20 CMIP5 models. The CMIP3 models are: bccr_bcm2_0, cccma_cgcm3_1_t63, cnrm_cm3, csiro_mk3_0, csiro_mk3_5, gfdl_cm2_0, gfdl_cm2_1, giss_aom, giss_model_e_h, giss_model_e_r, iap_fgoals1_0_g, ingv_echam4, inmcm3_0, ipsl_cm4, miroc3_2_hires, miroc3_2_medres, mpi_echam5, mri_cgcm2_3_2a, ncar_ccsm3_0, ncar_pcm1, ukmo_hadcm3, and ukmo_hadgem1. The CMIP5 models are: CanESM2, CCSM4, CESM1-CAM5, CNRM-CM5, CSIRO-Mk3-6-0, GFDL-CM3, GFDL-ESM2G, GFDL-ESM2M, GISS-E2-H, GISS-E2-R, HadGEM2-CC, HadGEM2-ES, inmcm4, IPSL-CM5A-LR, IPSL-CM5A-MR, MIROC5, MIROC-ESM, MPI-ESM-LR, MRI-CGCM3, NorESM1-M. All available ensemble members are considered individually. Trends are calculated at the grid cell and then averaged for the Eastern United States [30º-47º N, and 80º-100º W , land only].



the region within the Pacific storm track, with the storms steered by the Pacific jet stream. In global warming simulations, the mid-latitude jet stream commonly shifts poleward and upward and strengthens. However, this rule-of-thumb description applies primarily to the changes averaged around latitude circles, while projected shifts also have important dependence on longitude and season. Over the eastern Pacific Ocean in winter, the strengthening of the Pacific jet stream in the current CMIP5 models occurs where the climatological storm track veers northward. Here models predict a northward shift of the storm track, as measured by upper troposphere eddy meridional velocity variance, but a southward shift of the jet stream (Neelin et al. 2013, Simpson et al. 2014, Seager et al. 2014). The winter wetting in central to northern California in the CMIP5 models appears to be related not to a storm track shift but due to increased mean flow moisture convergence and a wet area getting wetter (Seager et al. 2014). This should not be viewed as contradicting previous CMIP3 model results, but as a local southward revision of the boundary between subtropical drying and mid-latitude precipitation increases. Despite model agreement in the CMIP5 ensemble, confidence in this change is not high because it depends on regional jet stream dynamics that is not yet fully understood. Longitudinally and seasonally varying changes in the jet stream and storm track location and intensity are important for regional hydroclimate projections and for determining the mechanisms responsible. Further assessment of fidelity of the model simulations of these features should be a research priority.

In summary, the boundary between winter wetting and drying in the southwest US is projected to be farther south in CMIP5 relative to CMIP3 results, although the changes are likely not statistically significant, highly dependent on the region and season, and the fidelity of the models in reproducing jet stream dynamics. In spring almost all regions are expected to dry, while summer changes are weak. The boundary between wetting and drying moves further northward when P-E_T is considered. Given that the CO_2 sensitivity of CMIP5 models is similar to CMIP3 models (see section 2.1), the shift in the boundary between wetting and drying is likely to be related to the treatment of aerosols and differences in GHG scenarios in CMIP5.

6.4. THE US WARMING HOLE

Several studies have explored reasons for the warming hole (WH) – a regional cooling or lack of warming in the central and southeast US relative to the general warming nearly everywhere else. The identified mechanisms include connections with internal climate variability (Christidis et al., 2010) and large-scale decadal variability such as the PDO and AMV (Robinson et al., 2002, Kunkel et al., 2006, Wang et al., 2009; Meehl et al., 2012; Weaver, 2013; Kumar et al., 2013a,b), regional scale hydrological processes (Pan et al., 2004), land surface interactions (Liang et at. 2006), secondary organic aerosols during the growing season (Portmann et al., 2009), and aerosol effects (Leibensperger et al., 2012). The WH has lessened in intensity since the 1980s, which may be related to the reduction in aerosols since that time (Leibensperger et al., 2012), although a change in trend over just 30 years is unlikely to be statistically significant.

Analysis of CMIP5 model simulations shows that they generally do not replicate the magnitude and timing of the WH, which was also found for the CMIP3 ensemble (Kunkel et al., 2006; Pan et al.

Analysis of CMIP5 model simulations shows that they generally do not replicate the magnitude and timing of the Warming Hole, which was also found for the CMIP3 ensemble.

2013). Figure 13 analyzes the WH as a decadal climate variability phenomenon in CMIP3 and CMIP5 simulations. All models show a negative 30-year running trend over some part of the 20th century in the eastern US, although most models underestimate the magnitude as well as the persistence (length) of the observed WH. The observed trend is contained within the CMIP5 ensemble (but at the upper end) in contrast with the CMIP3 models, for which the observed trend falls outside of the ensemble. This suggests that the inclusion of non-greenhouse gas forcings in CMIP5, e.g. aerosols and land use change (a negative radiative forcing), have increased the probability of simulating long-term negative trends in temperature. A forced response would tend to coincide in time with the observed signal and a small fraction of the CMIP5 models do show some skill in the timing of the WH (Kumar et al., 2013b; Pan et al. 2013). However, this is not a robust response across models and further investigation is required to determine the role of forcings and the ability of the models to translate this into a response. Most CMIP5 model projections (95% range from individual simulations) do not show a WH in the first half of the 21st century for annual or JJA means under RCP4.5 and RCP8.5 (Kumar et al. 2013b, Pan et al. 2013, and Maloney et al. 2014). While a WH does not emerge in the climate projections of the second half of the 21st century under RCP8.5, there is a slight chance (~10%) of a WH under RCP4.5 that may be related to CO_2 stabilization after 2050 in this emission scenario (Kumar et al. 2013b; Maloney et al. 2014).

6.5. EXTRA-TROPICAL CYCLONE ACTIVITY

Future projections of North American storm track activity show a much more significant decrease in CMIP5 models for the RCP8.5 scenario than CMIP3 for SRES A2, with the largest decrease in summer and smallest decrease in spring (Chang, 2013). While most models agree on the sign of the change, there is significant model-to-model spread, with the models projecting large storm track decreases also projecting negative precipitation changes over much of eastern North America and the southern U.S. Across the combined CMIP5 and CMIP3 ensemble, model-to-model differences in projected storm track change are found to correlate significantly with model-to-model differences in projected change in locally defined mean available potential energy (MAPE), suggesting that differences in the projected change in local MAPE can partly account for not only model-to-model differences, but also differences between CMIP5 and CMIP3 projections.

Future projections of North American storm track activity show a much more significant decrease in CMIP5 models for the RCP8.5 scenario than CMIP3 for SRES A2.

Based on preliminary analyses by Edmund Chang (Stony Brook U.), part of the difference between the projected storm track change by CMIP3 and CMIP5 can be explained by differences between the projected change in meridional temperature gradient between the lower and higher latitudes (polar amplification), not only at the surface, but also in the mid-troposphere. CMIP5 models as an ensemble project a larger decrease in both the surface and mid-tropospheric temperature gradient than CMIP3 models. The reasons for the differences in gradients are unclear, but may be related to the more comprehensive treatment of aerosols in CMIP5 and/or different GHG forcing scenarios. It should be noted, however, that there is large spatial variability in the future extra-tropical cyclone changes. Many regions show future cyclone decreases as noted above, but in some areas of the storm track, the decreases are smaller or there are increases, for example over the northeastern US, as found by Colle et al. (2013). Some of this variability may be an artifact of how cyclones are defined and tracked (Chang, 2014).

7.1. SUMMARY

Overall, the multi-model ensemble (MME) mean performance has not improved substantially in CMIP5 relative to CMIP3 for climatological variables (precipitation, sea surface temperature) over North American regions, except for a slight improvement for near surface air temperature over land. Some models have improved considerably, while others have surprisingly gotten worse. Projected changes in moderate to extreme precipitation events show a 20-30% increase over the US, which is similar to CMIP3, with much higher increases in the northeastern US, especially in winter. However, CMIP5 models tend to underestimate the frequency of heavy and extreme daily precipitation events, despite a slight improvement over CMIP3, especially in the southeastern US. Further work is required to understand the regional variability in model performance and the attribution of extreme events.

Generally, the CMIP5 models show better skill for basic attributes of ENSO with performance related to the mean SST state. It is unclear whether the representation of teleconnections with North American climate has improved. It is likely that the structure of Pacific Decadal Variability (PDV) as indexed by the PDO is slightly better simulated, albeit with larger biases in variability in the northwest Pacific. Teleconnections for precipitation also are improved slightly but remain poor overall. Atlantic Multidecadal Variability (AMV) as represented by the AMO is better represented in CMIP5 models in terms of decadal variability and persistence than CMIP3 models, but its SST footprint and teleconnections with North American climate are still poorly represented. The "warming hole" observed in the southeast US during the course of the 20th century is not replicated in the CMIP5 ensemble, and appears to be related to multi-decadal variability in the north Atlantic rather than a forced signal. Model projections show no warming hole in the future, although reduced warming is related to the negative phase of the AMO.

In southwest North America, projections of changes in precipitation from CMIP5 tend to be more robust overall than CMIP3 based on model consensus on the sign of the change. In particular, a more robust signal of summer drying in the Caribbean and southwest Mexico is evident in CMIP5 models. For the Southwest monsoon, previous studies of CMIP3 and CMIP5 models indicate a redistribution of Southwest monsoon rainfall with reduced early season amounts and increased late season amounts, although the latter is less robust because of large inter-model variability. New analysis based on CMIP5 daily data shows a slight shift to a later onset of about 6-8 days but no late retreat date. The boundary between winter wetting and drying in the Southwest US is projected to move southward in CMIP5 relative to CMIP3 results, although the changes are highly dependent on the region and season, and the fidelity of the models in reproducing jet stream dynamics. In spring, almost all regions are expected to dry, whilst summer changes are weak. The boundary between wetting and drying moves farther northward when precipitation minus evapotranspiration is considered. The CMIP5 models project a more significant decrease in extra-tropical storm track activity than CMIP3 that may be related to a larger projected decrease in the temperature gradient between lower and higher latitudes.

7.2. FUTURE RESEARCH DIRECTIONS

This report provides an overview of current understanding of differences in CMIP3 and CMIP5 model performance and projections, but by no means is comprehensive in terms of the literature reviewed and analyses carried out.

A more comprehensive evaluation is needed, especially at regional scales and in terms of direct comparisons of CMIP3 and CMIP5 ensembles, for several climate features and processes. Furthermore, there is considerable scope for improving our understanding of the attribution of the differences, which may ultimately help understand errors in current models and therefore the robustness of future projections. There are also several aspects of the model evaluation process that need to be improved or better understood, in particular for extremes, including observational uncertainties and their impact on evaluations, scale mismatches between models and observations (e.g. King et al. 2013), and the impact of small sample sizes (e.g. single model ensemble member) to detect and compare extreme events.

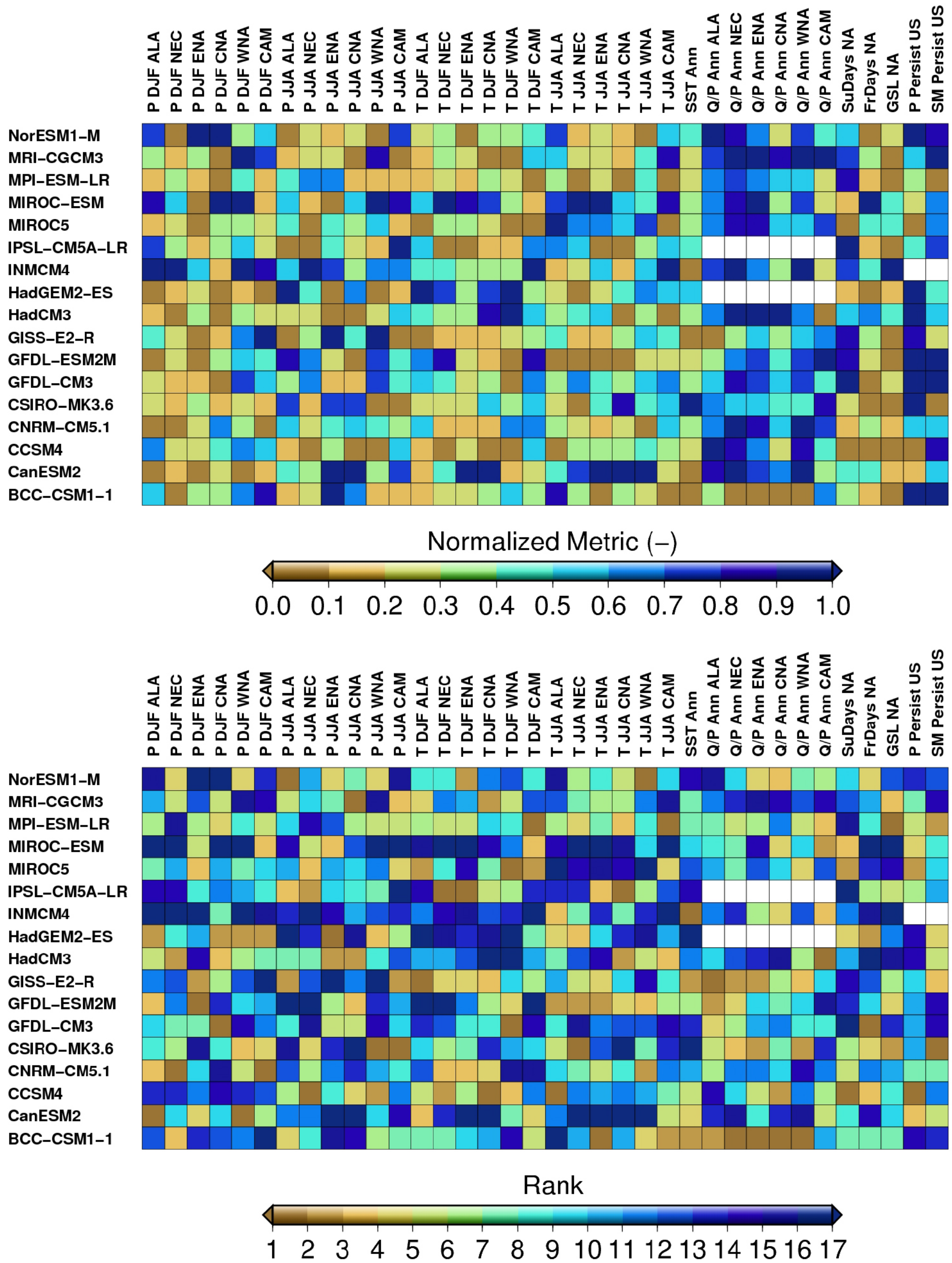
In terms of specific climate processes, there are several avenues of future research that are recommended in the individual sections above, for example, in understanding the attribution of extreme events and the relationship with biases in larger scale climate processes. There are hints that increasing resolution appears to help (e.g. Wehner et al. 2010; Li et al., 2011; DeAngelis et al., 2013) but as the increases in model resolution from CMIP3 to CMIP5 have been modest and performance is influenced by changes in other factors, such as forcings and model parameterizations, conclusions on the impact of resolution are not possible. Further investigation is required to determine the impact of resolution and whether there is a limit to the increase in skill, which has been documented for some models as resolution approaches 30km, after which skill plateaus or even degrades, likely due to the breakdown of the assumptions in the convection parameterizations (Kinter et al., 2013). Understanding the influence of increasing resolution versus improved parameterizations through controlled experiments should be a research priority.

Differences between CMIP3 and CMIP5 projected changes for several aspects of climate may be related to the updated treatment of aerosols in CMIP5, given that the CO_2 sensitivity of CMIP5 models is similar to CMIP3 models. For example, the shift in the boundary between wetting and drying in the southwestern US and the projected decrease in extra-tropical storm track activity is possibly related to the treatment of aerosols and differences in GHG scenarios in CMIP5. However, this needs to be confirmed across a range of models – something that is not possible with the handful of CMIP5 models that have run the full set of forcing experiments.

Future Research Directions:

- Regional-scale evaluation of CMIP3/CMIP5 differences.
- Evaluation of model performance for extremes needs improvement taking into account observational uncertainties, scale mismatches, and small sample sizes.
- Further evaluation of the impact of increased resolution and thresholds of model parameterization performance and suitability.
- Better understanding of the impact of new GHG scenarios, aerosol treatments, and enhanced land surface processes on the projections.
- Continued review of model evaluations based on historical simulations and implications for uncertainty in future projections.
- Comprehensive analyses including diverse metrics and evaluations of means, extremes, variability, and teleconnections.

Figure 14. Comparison of 17 CMIP5 models across a set of North American continental performance metrics based on bias values. (top) Biases normalized relative to the range of bias values across models, with lower values indicating lower bias. (bottom) Models ranked according to bias values, with 1 indicating the model with the lowest bias and 17 the model with the highest bias. Results for models without available data are indicated in white. The bias metrics shown (in order from left to right) are for regional precipitation (P) for DJF and JJA, regional temperature (T) for DJF and JJA, annual SSTs for surrouding oceans, annual runoff ratios (Q/P), the annual number of summer days (SuDays), frost days (FrDays) and growing season length (GSL), and east-west gradient in the number of persistent precipitation (P Persist) and soil moisture (SM Persist) events. (from Sheffield et al., 2013a).



The impact of vegetation on climate and their feedbacks have not been well evaluated, but may have important implications in the context of increasing CO_2 that may contribute to CMIP3/CMIP5 differences and the uncertainties in future projections Global coupled climate-carbon models have shown that the reduction in ET due to elevated atmospheric CO_2 can reduce surface relative humidity, increase the depth of the atmospheric boundary layer, reduce low cloud cover and increase runoff (Sellers et al. 1996, Gedney et al. 2006, Betts et al. 2007). These physiological-response-induced land surface and cloud feedbacks can lead to a 10-20% increase in surface solar radiation and about 0.5-0.7K surface temperature increase over global land under the double CO_2 concentrations (e.g. Boucher et al. 2008; Andrews et al. 2011; Pu and Dickinson 2014). These effects are particularly strong over the eastern US relative to global land (e.g., Doutriau-Boucher et al. 2009, Betts and Chiu 2010). The impact on precipitation is highly uncertain. For example, simulations based on the HadCM3LC model suggest a 3% decrease of rainfall over global land due to a drier atmospheric boundary layer (Andrews et al. 2011), whereas simulations of the CESM1.0.2 model suggest a 0.03 mm/day increase of rainfall over global land due to enhanced moisture transport from the ocean as a result of stronger land warming (Ping and Dickinson 2013). A clearer understanding is therefore required of the impacts of plant physiological response to elevated atmospheric CO_2 on the surface water and energy budgets, and its contribution to the observed and projected changes of temperature, humidity, precipitation and clouds, over vegetated land regions such as eastern US.

One aspect that has not been well evaluated is the performance of individual models across different climate processes and whether their skill has any bearing on the plausibility of the future projections. A limited analysis of model errors across different metrics was carried out in Sheffield et al. (2013a) for regional/seasonal precipitation and near surface air temperature, annual sea surface temperature, regional/annual runoff over precipitation, summer days/frost days, growing season length, and precipitation/soil moisture persistence (Fig. 14). Sheffield et al. (2013a) note that no model stands out as being better than the others, although there are indications that higher resolution models tend to perform better for some climate features, especially for the regional features as expected, but not universally so and not for basic climate variables. The historical performance of a climate model depends on the variable examined and the metrics used, and may not necessarily give an indication of the robustness of its future projections (Glecker et al., 2008; Knutti and Sedláček, 2013). Further review is required on model evaluations and the relationship with future projections. A more comprehensive analysis is required across metrics, including representation of means, extremes, variability and teleconnections, and in the context of impacts of interest (e.g. regional precipitation). Potential benefits may be gleaned from analysis of whether an improved mean state leads to better representation of variability, for example in ENSO, and therefore to improved teleconnections and less uncertainty in future projections.

7.3. IMPLICATIONS FOR DESIGN AND EXECUTION OF CMIP6

The results of this report indicate a slight overall improvement in the representation of historic climate in CMIP5 relative to CMIP3 and some aspects of regional climate change that are more robust in terms of model skill and model consensus.

However, there remain many climate features that continue to be poorly represented (e.g. precipitation extremes and teleconnections) and large uncertainties in future projections overall and especially relative to their skill in replicating relevant processes. The next CMIP phase (CMIP6) is currently in the initial design phase (Meehl et al., 2014) and affords the opportunity to make progress towards understanding these ongoing issues through targeted experiments and analyses. The design of CMIP6 is set against the background of the World Climate Research Programme (WCRP) Grand Challenges that cover questions related to clouds, circulation and climate sensitivity; changes in the cryosphere; climate extremes; regional climate information; regional sea level rise; and water availability; plus an additional theme covering questions related to biospheric forcings and feedbacks. In addition to core experiments that provide continuity from previous CMIPs (such as AMIP type experiments; pre-industrial control runs; and climate sensitivity experiments), CMIP6 allows for the proposal of model intercomparison projects (MIPs) that target a specific process. Given the main results in this paper, the following experiments and targeted analyses may warrant exploration:

1. Climate Variability and Surface Climate Teleconnections. AMIP-type runs may help understand the performance of models, especially in the context of biases in the mean SST state and observed non-stationarities in teleconnections. For example, teleconnections of North American surface climate with ENSO are non-stationary, in part due to the influence of decadal variability in the north Pacific and Atlantic oceans (Kam et al., 2014a,b). Proposed CMIP6 AMIP type runs are for 1979-2010, but longer-term runs (~100-years) would be necessary. Regionally coupled experiments would also have merit to isolate the influence of different ocean basins on land surface climate, for example by allowing full coupling in the ENSO region to determine the role of the north Pacific and north Atlantic Oceans.

2. Temperature and Precipitation Extremes. Coordinated analysis of attribution of types of extreme events is required (e.g. to tropical cyclones, extra-tropical storms, convective activity, topography, weather types, atmospheric circulation versus local processes, moisture sources, ENSO, etc.). Coordinated intercomparisons on identified key processes in models perhaps complemented by experiments across sets of models related to model resolution versus physics parameterizations may help researchers understand biases. Larger ensembles will be useful to improve the robustness of calculated statistics of extremes. Issues related to observational uncertainties and scale mismatches need to be addressed, for example, through improved representation of scaling behavior of extremes.

ACKNOWLEDGMENTS

This report was compiled the CMIP5 Task Force under the auspices of the National Oceanic and Atmospheric Administration (NOAA) Modeling, Analysis, Predictions and Projections (MAPP) program, which is managed by the Climate Program Office (CPO) in the Office of Oceanic and Atmospheric Research (OAR). The authors acknowledge financial support from the CPO for this work, and thank four reviewers who provided comprehensive and useful comments.

Andrews, Timothy, et al. "A regional and global analysis of carbon dioxide physiological forcing and its impact on climate." Climate dynamics 36.3-4 (2011): 783-792.

Andrews, T., J. M. Gregory, M. J. Webb, and K. E. Taylor (2012), Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models, Geophys. Res. Lett., 39, L09712, doi:10.1029/2012GL051607.

Betts RA, Boucher O et al (2007) Projected increase in continental runoff due plant response to increasing carbon dioxide. Nature 448:1037–1042. doi:10.1038/nature06045

Biasutti, M., and A. H. Sobel, 2009: Delayed Sahel Rainfall and Global Seasonal Cycle in a Warmer Climate. Geophys. Res. Letts., 36:L23707. doi:10.1029/2009GL041303.

Bellenger H., E. Guilyardi, J. Leloup, M. Lengaigne, J. Vialard (2013). ENSO representation in climate models: from CMIP3 to CMIP5. Clim. Dyn., DOI: 10.1007/s00382-013-1783-z

Boucher, O., A. Jones, and R. A. Betts (2009), Climate response to the physiological impact of carbon dioxide on plants in the Met Office Unified Model HadCM3, Clim. Dyn., 32, 237–249, doi:10.1007/s00382-008-0459-6.

Chang, E. K. M., 2013: CMIP5 projection of significant reduction in extratropical cyclone activity over North America. J. Climate, 26, 9903-9922, DOI: 10.1175/JCLI-D-13-00209.1 in review.

Chang, E. K. M., 2014: Impacts of background field removal on CMIP5 projected changes in Pacific winter cyclone activity. J. Geophys. Res. Atmos., 119, 4626–4639, doi:10.1002/2013JD020746.

Christidis, N., P. A. Stott, F. W. Zwiers, H. Shiogama, and T. Nozawa, 2010: Probabilistic estimates of recent changes in temperature: A multi-scale attribution analysis. Climate Dyn., 34, 1139–1156.

Colle, Brian A., Zhenhai Zhang, Kelly A. Lombardo, Edmund Chang, Ping Liu, Minghua Zhang, 2013: Historical Evaluation and Future Prediction of Eastern North American and Western Atlantic Extratropical Cyclones in the CMIP5 Models during the Cool Season. J. Climate, 26, 6882–6903. doi: http://dx.doi.org/10.1175/JCLI-D-12-00498.1

Collins, M., R. Knutti, J. Arblaster, J.-L. Dufresne, T. Fichefet, P. Friedlingstein, X. Gao, W.J. Gutowski, T. Johns, G. Krinner, M. Shongwe, C. Tebaldi, A.J. Weaver and M. Wehner, 2013: Long-term Climate Change: Projections, Commitments and Irreversibility. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Cook, B. I., and R. Seager, 2013: The response of the North American Monsoon to increased greenhouse gas forcing, J. Geophys. Res. Atmos., 118, 1690–1699, doi:10.1002/jgrd.50111.

Corti, S., A. Weisheimer, T. Palmer, F. Doblas-Reyes, and L. Magnusson, 2012: Reliability of decadal predictions. Geophys. Res. Lett., 39, L21712, doi:10.1029/2012GL053354.
Dai, A., 2013: Increasing drought under global warming in observations and models. Nat. Clim. Change, 3, 52-58. DOI: 10.1038/NCIMATE1633.

DeAngelis, Anthony M., Anthony J. Broccoli, Steven G. Decker, 2013: A Comparison of CMIP3 Simulations of Precipitation over North America with Observations: Daily Statistics and Circulation Features Accompanying Extreme Events. J. Climate, 26, 3209–3230. doi: http://dx.doi.org/10.1175/JCLI-D-12-00374.1

DeAngelis, A. M., and A. J. Broccoli, 2013: Projected Changes in Heavy Precipitation over North America in CMIP5 Climate Model Simulations, 25th Conference on Climate Variability and Change, AMS Annual Meeting, Austin, TX, USA.

Delworth, T. L., and M. E. Mann, 2000: Observed and simulated multidecadal variability in the Northern hemisphere. Clim. Dynam., 16, 661–676.

Deser, C., Phillips, A., Bourdette, V. & Teng, H. Uncertainty in climate change projections: the role of internal variability. *Clim. Dynam*. 38, 527–547 (2012).

Doutriaux Boucher, M., M. J. Webb, J. M. Gregory, and O. Boucher (2009), Carbon dioxide induced stomatal closure increases radiative forcing via a rapid reduction in low cloud, Geophys. Res. Lett., 36, L02703, doi:10.1029/2008GL036273.

Dunstone, N. J., D. M. Smith, B. B. B. Booth, L. Hermanson, and R. Eade, 2013: Anthropogenic aerosol forcing of Atlantic tropical storms. *Nature Geoscience*, 6, 534–539. doi:10.1038/ngeo1854

Enfield, D. B., A. M. Mestas-Nunez, and P. J. Trimble, 2001: The Atlantic multidecadal oscillation and its relationship to rainfall and river flows in the continental US. Geophys. Res. Lett., 28, 2077–2080.

Furtado, J. C., E. Di Lorenzo, N. Schneider, and N. A. Bond, 2011. North Pacific decadal variability and climate change in the IPCC AR4 models. J. Climate, 24, 3049-3067.

Gedney N, Cox PM, Betts RA, Boucher O, Huntingford C, Stott PA (2006) Detection of a direct carbon dioxide effect in continental river runoff records. Nature 439:835–838. doi:10.1038/nature04504

Geil, K. L., Y. L. Serra and X. Zeng, 2013: Assessment of CMIP5 Model Simulation of the North American Monsoon System. J. Climate, 26, 8787–8801. doi: http://dx.doi.org/10.1175/JCLI-D-13-00044.1.

Gleckler, P. J., K. E. Taylor, and C. Doutriaux (2008), Performance metrics for climate models, J. Geophys. Res., 113, D06104, doi:10.1029/2007JD008972.

Guilyardi E., H. Bellenger, M. Collins, S. Ferrett, W. Cai & A. Wittenberg (2012). A first look at ENSO in CMIP5. CLIVAR Exchanges, 58, 29-32

Hu, Q., and S. Feng, 2008: Variation of the North American summer monsoon regimes and the Atlantic Multidecadal Oscillation. J. Climate, 21, 2371-2383.

Hu, Q., S. Feng, and R.J. Oglesby, 2011: Variations in North American summer precipitation driven by the Atlantic Multidecadal Oscillation. J. Climate, 24, 5555-5570.

Hu, Q., and S. Feng, 2012: AMO- and ENSO-driven summertime circulation and precipitation variations in North America. J. Climate, 25, 6477-6495.

Kam, J., J. Sheffield, and E. F. Wood, 2014: Changes in drought risk over the contiguous United States (1901–2012): The influence of the Pacific and Atlantic Oceans. Geophys. Res. Lett., 41, 5897–5903, doi:10.1002/2014GL060973.

Kao, H.-Y. and J.-Y. Yu, 2009: Contrasting Eastern-Pacific and Central-Pacific Types of ENSO. Journal of Climate, 22, 615-632.

Kavvada, A., A. Ruiz-Barradas, and S. Nigam, 2013: AMO's structure and climate footprint in observations and IPCC AR5 climate simulations. Climate Dynamics, accepted.

Kim, H.-M., P. J. Webster, and J. A. Curry, 2012: Evaluation of short-term climate change prediction in multi-model CMIP5 decadal hindcasts. Geophys. Res. Lett., 39, L10701, doi:10.1029/2012GL051644.

Kim, S. T., and J.-Y. Yu (2012), The two types of ENSO in CMIP5 models, Geophys. Res. Lett., 39, L11704, doi:10.1029/2012GL052006.

King, A. D., Alexander, L. and Donat, M. G. (2013) The efficacy of using gridded data to examine extreme rainfall characteristics: a case study for Australia. *International Journal of Climatology*, 33, 2376-2387 http://dx.doi.org/10.1002/joc.3588

Kinter III, J. L., B. Cash, D. Achuthavarier, J. Adams, E. Altshuler, P. Dirmeyer, B. Doty, B. Huang, L. Marx, J. Manganello, C. Stan, T. Wakefield, E. Jin, T. Palmer, M. Hamrud, T. Jung, M. Miller, P. Towers, N. Wedi, M. Satoh, H. Tomita, C. Kodama, T. Nasuno, K. Oouchi, Y. Yamada, H. Taniguchi, P. Andrews, T. Baer, M. Ezell, C. Halloy, D. John, B. Loftis, R. Mohr, and K. Wong, 2013: Revolutionizing Climate Modeling – Project Athena: A Multi-Institutional, International Collaboration. *Bull. Amer. Meteor. Soc.*, 94, 231-245.

Knight, J. R., R. J. Allan, C. K. Folland, M. Vellinga, M. E. Mann, 2005: A signature of persistent natural thermohaline circulation cycles in observed climate. Geophys. Res. Lett., 32, L20708, doi:10.1029/2005GL024233.

Knudsen, M. F., M.-S. Seidenkrantz, B. H. Jacobsen, and A. Kuijpers, 2011: Tracking the Atlantic Multidecadal Oscillation through the last 8,000 years. *Nature Communications,* 2, 178. doi:10.1038/ncomms1186

Knutson, T.R., Sirutis, J.J., Vecchi, G.A., Garner, S., Zhao, M., Kim, H.S., Bender, M., Tuleya, R.E., Held, I. M., and Villarini, G., 2013: Dynamical downscaling projections of twenty-first-century atlantic hurricane activity: CMIP3 and CMIP5 model-based scenarios. J. Climate, 26 (17), 6591-6617

Knutti R. and J Sedláček 2013: Robustness and uncertainties in the new CMIP5 climate model projections. Nat. Clim. Change, 3, 369-373. DOI: 10.1038/NCLIMATE1716.

Kug, J.-S., Y.-G. Ham, J.-Y. Lee, and F.-F. Jin, 2012: Improved simulation of two types of El Niño in CMIP5 models, Environ. Res. Lett., 7(3), 034002, doi:10.1088/1748-9326/7/3/034002.

Kumar S., V. Merwade, J. Kinter III, D. Niyogi, 2013a: Evaluation of Temperature and Precipitation Trends and Long-term Persistence in CMIP5 20th Century Climate Simulations. J. Climate, 26 (12), 4168-4185.

Kumar, S., J. Kinter, P. A. Dirmeyer, Z. Pan, J. Adams, 2013b: Multidecadal Climate Variability and the "Warming Hole" in North America: Results from CMIP5 Twentieth- and Twenty-First-Century Climate Simulations. J. Climate, 26, 3511–3527. doi: http://dx.doi.org/10.1175/JCLI-D-12-00535.1

Kumar, S., J. L. Kinter III, P. A. Dirmeyer, and D. M. Lawrence, 2014: Climate Processes in CMIP5: The "Warming Hole" simulations in CMIP5 models – role of natural climate variability versus anthropogenic effects. AMS 94th Annual Meeting, 2-6 February 2014, Atlanta, Georgia.

Kunkel, K. E., X.-Z. Liang, J. Zhu, and Y. Lin, 2006: Can CGCMs simulate the twentieth-century ''warming hole'' in the central United States? J. Climate, 19, 4137–4153.

Langenbrunner, B., and J. D. Neelin, 2013: Analyzing ENSO Teleconnections in CMIP Models as a Measure of Model Fidelity in Simulating Precipitation. J. Climate, 26, 4431–4446. doi: http://dx.doi.org/10.1175/JCLI-D-12-00542.1

Lapp, S. L., St. Jacques, J.-M., Barrow, E. M. and Sauchyn, D. J. (2012), GCM projections for the Pacific Decadal Oscillation under greenhouse forcing for the early 21st century. Int. J. Climatol., 32: 1423–1442. doi: 10.1002/joc.2364

Lee, S.-K., and C. Wang, 2010: Delayed advective oscillation of the Atlantic thermohaline circulation. *J. Climate*, 23, 1254-1261.

Liang, X.-Z., J. Pan, J. Zhu, K. E. Kunkel, J. X. L. Wang, and A. Dai, 2006: Regional climate model downscaling of the U.S. summer climate and future change. J. Geophys. Res., 111, D10108, doi:10.1029/2005JD006685.

Liang, X.-Z., J. Zhu, K. E. Kunkel, M. Ting, and J. X. L. Wang, 2008: Do CGCMs simulate the North American monsoon precipitation seasonal-interannual variations. J. Climate, 21, 3755-3775.

Lienert, F., 2011: Simulation and Prediction of North Pacific Sea Surface Temperature. Thesis (Ph.D.), University of Victoria (Canada), 131 p., ISBN: 9780494824276.

Liu, H., C. Wang, S.-K. Lee, and D. B. Enfield, 2013: Atlantic warm pool variability in the CMIP5 simulations. J. Climate, 26, 5315-5336.

Maloney, E. D., S. J. Camargo, E. Chang, B. Colle, R. Fu, K. L. Geilw, Q. Hu, X. Jiang, N. Johnson, K. B. Karnauskas, J. Kinter, B. Kirtman, S. Kumar, B. Langenbrunner, K. Lombardo, L. Long, A. Mariotti, J. E. Meyerson, K. Mo, J. D. Neelin, Z. Pan, R. Seager, Y. Serraw, A. Seth, J. Sheffield, J. Thibeault, S.-P. Xie, C. Wang, B. Wyman, and M. Zhao, 2013: North American Climate in CMIP5 Experiments: Part III: Assessment of 21st Century Projections. J. Climate, in revision.

McSweeney, C. F., R. G. Jones, and B. B. B. Booth, 2012: Selecting Ensemble Members to Provide Regional Climate Change Information. *J. Climate*, 25, 7100–7121. doi: http://dx.doi.org/10.1175/JCLI-D-11-00526.1

Medhaug, I., and T. Furevik, 2011: North Atlantic 20th century multidecadal variability in coupled climate models: Sea surface temperature and ocean overturning circulation. Ocean Sci., 7, 389–404.

Meehl, G. A., J. M. Arblaster, and G. Branstator, 2012: Mechanisms contributing to the warming hole and the consequent U.S. east–west differential of heat extremes. J. Climate, 25, 6394–6408.

Meinshausen, M., S. C. B. Raper and T. M. L. Wigley (2011). "Emulating coupled atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6: Part I – Model Description and Calibration." Atmospheric Chemistry and Physics 11: 1417-1456. doi:10.5194/acp-11-1417-2011.

Min, S.-K., X. Zhang, F. W. Zwiers, and G. C. Hegerl, 2011: Human contribution to more-intense precipitation extremes. *Nature*, 470, 378-381, doi:10.1038/nature09763.

Nakićenović, N., et al. (2000) IPCC Special Report on Emissions Scenarios. Cambridge, UK and New York, NY: Cambridge University Press.

Neelin, J. D., M. Munnich, H. Su, J. E. Meyerson, and C. E. Holloway, 2006: Tropical drying trends in global warming models and observations. Proc. Nat. Acd. Sci., 103, 6110-6115.

Neelin, J. D., B. Langenbrunner, J. E. Meyerson, A. Hall, and N. Berg, 2013: California winter precipitation change under global warming in the Coupled Model Intercomparison Project 5 ensemble. J. Climate, 26, 6238-6256.

Nigam, S., B. Guan, and A. Ruiz-Barradas, 2011: Key role of the Atlantic multidecadal oscillation in 20th century drought and wet periods over the Great Plains. Geophys. Res. Lett., 38, L16713, doi:10.1029/2011GL048650.

Oshima, K. and Y. Tanimoto, 2009: An Evaluation of Reproducibility of the Pacific Decadal Oscillation in the CMIP3 Simulations. Journal of the Meteorological Society of Japan, 87, 755-770.

Ottera, O. H., Bentsen, M., Drange, H., and Suo, L.: External forcing as a metronome for Atlantic multidecadal variability, *Nat. Geosci.*, 3, 688–694, doi:10.1038/NGEO955, 2010.

Pan, Z., R. W. Arritt, E. S. Takle, W. J. Gutowski Jr., C. J. Anderson, and M. Segal, 2004: Altered hydrologic feedback in a warming climate introduces a "warming hole." Geophys. Res. Lett., 31, L17109, doi:10.1029/2004GL020528.

Pan, Z., X. Liu, S. Kumar, Z. Gao, and J. Kinter, 2013: Intermodel variability and mechanism attribution of central and southeastern U.S. anomalous cooling in the 20th century as simulated by CMIP5 models. J. Climate, 26 (17), 6215–6237. doi: http://dx.doi.org/10.1175/JCLI-D-12-00559.1

Polade, S. D., A. Gershunov, D. R. Cayan, M. D. Dettinger, and D. W. Pierce, 2013: Natural climate variability and teleconnections to precipitation over the Pacific-North American region in CMIP3 and CMIP5 models. Geophys. Res. Lett., 40, 2296–2301, doi:10.1002/grl.50491.

Portmann, R. W., S. Solomon, and G. C. Hegel, 2009: Spatial and seasonal patterns in climate change, temperatures, and pre- cipitation across the United States. Proc. Natl. Acad. Sci. USA, 106, 7324–7329.

Pu, Bing, and Robert E. Dickinson. "Diurnal spatial variability of Great Plains summer precipitation related to the dynamics of the low-level jet." Journal of the Atmospheric Sciences 2014 (2014).

Robinson, W. A., R. Reudy, and J. E. Hansen, 2002: General circulation model simulations of recent cooling in the east-central United States. J. Geophys. Res., 107, 4748, doi:10.1029/2001JD001577.

Ruiz-Barradas, A., S. Nigam, and A. Kavvada, 2013: The Atlantic Multidecadal Oscillation in 20th century climate simulations: uneven progress from CMIP3 to CMIP5. Climate Dynamics, in review

Ruiz-Barradas A., S. Nigam and A. Kavvada, 2013: Assessment of CMIP3 and CMIP5 20th century climate simulations over North America: similar climatologies. In preparation.

Ryu J.-H., K. Hayhoe, 2013: Understanding the sources of Caribbean precipitation biases in CMIP3 and CMIP5 simulations. Clim. Dyn, Online First paper. DOI 10.1007/s00382-013-1801-1.

Schubert, S. D., and Coauthors, 2009: A U.S. CLIVAR 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.

Scoccimarro, E. S. Gualdi, A. Bellucci, M. Zampieri, and A. Navarra, 2013: Heavy precipitation events in a warmer climate: results from CMIP5 models. J. Climate, doi: http://dx.doi.org/10.1175/JCLI-D-12-00850.1

Seager R., M. Ting, C. Li, N. Naik, B. Cook, J. Nakamura and H. Liu, 2013: Projections of declining surface-water availability for the southwestern United States. Nat. Clim. Change, 3, 482-486. DOI: 10.1038/NCLIMATE1787.

Seager, R., D. Neelin, I. Simpson, H. Liu, N. Henderson, T. Shaw, Y. Kushnir, M. Ting, and B. Cook, 2014: Dynamical and Thermodynamical Causes of Large-Scale Changes in the Hydrological Cycle over North America in Response to Global Warming. *J. Climate*, 27, 7921–7948. doi: http://dx.doi.org/10.1175/JCLI-D-14-00153.1

Sellers, P. J., et al. (1996), Comparison of radiative and physiological effects of doubled atmospheric CO_2 on climate, Science, 271, 1402 –1406, doi:10.1126/science.271.5254.1402.

Sen Gupta, A., N. C. Jourdain, J. N. Brown, and D. Monselesan, 2013: Climate Drift in the CMIP5 Models. *J. Climate*, 26, 8597–8615. doi: http://dx.doi.org/10.1175/JCLI-D-12-00521.1

Seth, A., M. Rojas, S. A. Rauscher, 2010: CMIP3 projected changes in the annual cycle of the South American Monsoon. Climatic Change, 98, 331{357, doi:10.1007/s10584-009-9736-6.

Seth, A., S. Rauscher, M. Rojas, S. Camargo, A. Giannini, 2011: Enhanced spring convective barrier for monsoons in a warmer world? Climatic Change Let., 104, 403{414, doi:10.1007/s10584-010-9973-8.

Seth, A., S. Rauscher, M. Biasutti, A. Giannini, S. Camargo, and M. Rojas, 2013: CMIP5 Projected Changes in the Annual Cycle of Precipitation in Monsoon Regions. J. Climate, 26, 7328–7351. doi:10.1175/JCLI-D-12-00726.1.

Sheffield, J., A. Barrett, B. Colle, R. Fu, K. L. Geil, Q. Hu, J. Kinter, S. Kumar, B. Langenbrunner, K. Lombardo, L. N. Long, E. Maloney, A. Mariotti, J. E. Meyerson, K. C. Mo, J. D. Neelin, Z. Pan, A. Ruiz-Barradas, Y. L. Serra, A. Seth, J. M. Thibeault, and J. C. Stroeve, 2013a: North American Climate in CMIP5 Experiments. Part I: Evaluation of historical simulations of continental and regional climatology. J. Climate, 26, 9209–9245. doi: http://dx.doi.org/10.1175/JCLI-D-12-00592.1.

Sheffield, J., S. J. Camargo, B. Colle, Q. Hu, X. Jiang, N. Johnson, S. Kumar, K. Lombardo, B. Langenbrunner, E. Maloney, J. E. Meyerson, J. D. Neelin, Y. L. Serra, D.-Z. Sun, C. Wang, S.-P. Xie, J.-Y. Yu, T. Zhang, 2013b: North American Climate in CMIP5 Experiments: Part II: Evaluation of historical simulations of intra-seasonal to decadal variability. *J. Climate*, 26, 9247–9290. doi: http://dx.doi.org/10.1175/JCLI-D-12-00593.1.

Sillmann, J., V. V. Kharin, X. Zhang, F. W. Zwiers, and D. Bronaugh (2013a), Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate, J. Geophys. Res. Atmos., 118, 1716–1733, doi:10.1002/jgrd.50203.

Sillmann, J., V. V. Kharin, F. W. Zwiers, X. Zhang, and D. Bronaugh (2013b), Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections, J. Geophys. Res. Atmos., 118, 2473–2493, doi:10.1002/jgrd.50188.

Simpson, I. R., T. A. Shaw, and R. Seager, 2014: A Diagnosis of the Seasonally and Longitudinally Varying Midlatitude Circulation Response to Global Warming. J. Atmos. Sci., 71, 2489–2515. doi: http://dx.doi.org/10.1175/JAS-D-13-0325.1

Stephens, G. L., T. L'Ecuyer, R. Forbes, A. Gettlemen, J.-C. Golaz, A. Bodas-Salcedo, K. Suzuki, P. Gabriel, and J. Haynes (2010), Dreary state of precipitation in global models, J. Geophys. Res., 115, D24211, 13.

Sun, L., K. E. Kunkel, L. E. Stevens, G. Dobson, A. Buddenberg, and D. R. Easterling, 2014: Regional Surface Climate Conditions in CMIP3 and CMIP5 for the United States: Differences, Similarities, and Implications for the U.S. National Climate Assessment. NOAA Technical Report NESDIS ???

Sutton, R. T., and D. L. R. Hodson, 2005: Atlantic ocean forcing of North American and European summer climate. Science, 309, 115–117, doi:10.1126/science.1109496.

Thober, S., and L. Samaniego (2014), Robust ensemble selection by multivariate evaluation of extreme precipitation and temperature characteristics, J. Geophys. Res. Atmos., 119, 594–613, doi:10.1002/2013JD020505.

Ting, M., Y. Kushnir, R. Seager, and C. Li, 2009: Forced and internal twentieth-century SST trends in the North Atlantic. J. Climate, 22, 1469–1481.

van Vuuren et al (2011) The Representative Concentration Pathways: An Overview. Climatic Change, 109 (1-2), 5-31

Villarini, G., and G. A. Vecchi, 2012: Twenty-first-century projections of North Atlantic tropical storms from CMIP5 models. Nature Climate Change, 2 (8), 604-607.

Villarini, G., and G. A. Vecchi, 2013: Projected increases in North Atlantic tropical cyclone intensity from CMIP5 models. J. Climate, 26 (10), 3231-3240.

Walsh, J., D. Wuebbles, K. Hayhoe, J. Kossin, K. Kunkel, G. Stephens, P. Thorne, R. Vose, M. Wehner, J. Willis, D. Anderson, S. Doney, R. Feely, P. Hennon, V. Kharin, T. Knutson, F. Landerer, T. Lenton, J. Kennedy, and R. Somerville, 2014: Ch. 2: Our Changing Climate. *Climate Change Impacts in the United States: The Third National Climate Assessment*, J. M. Melillo, Terese (T.C.) Richmond, and G. W. Yohe, Eds., U.S. Global Change Research Program, 19-67. doi:10.7930/J0KW5CXT.

Wang, C., and D. B. Enfield (2001), The tropical western hemisphere warm pool. Geophys. Res. Lett., 28, 1635–1638.

Wang, C. D. B. Enfield, S.-K. Lee, and C. W. Landsea, 2006: Influences of the Atlantic warm pool on Western Hemisphere summer rainfall and Atlantic hurricanes. J. Climate, 19, 3011–3028.

Wang, H., S. Schubert, M. Suarez, J. Chen, M. Hoerling, A. Kumar, and P. Pegion, 2009: Attribution of the seasonality and regionality in climate trends over the United States during 1950–2000. J. Climate, 22, 2571–2590.

Wang, C., and L. Zhang, 2013: Multidecadal ocean temperature and salinity variability in the tropical North Atlantic: Linking with the AMO, AMOC and subtropical cell. J. Climate, 26, 6137-6162.

Weaver, S. J., 2013: Factors associated with decadal variability in Great Plains summertime surface temperatures. J. Climate, 26, 343–350.

Wittenberg, A. T., 2009: Are historical records sufficient to constrain ENSO simulations? Geophys. Res. Lett., 36, L12702, doi:10.1029/2009GL038710

Yeh, S. W., J. S. Kug, B. Dewitte, M. H. Kwon, B. P. Kirtman, and F. F. Jin, 2009: El Niño in a changing climate. Nature, 461: 511–570.

Yeh, S.-W., Kim, H., Kwon, M. and Dewitte, B. (2013), Changes in the spatial structure of strong and moderate El Niño events under global warming. Int. J. Climatol.. doi: 10.1002/joc.3876

Yu, J.-Y. and H.-Y. Kao, 2007: Decadal Changes of ENSO Persistence Barrier in SST and Ocean Heat Content Indices: 1958-2001. *Journal of Geophysical Research*, 112, D13106, doi:10.1029/2006JD007654.

Yu., J.-Y., H.-Y. Kao and T. Lee, 2010: Subtropics-Related Interannual Sea Surface Temperature Variability in the Equatorial Central Pacific. *Journal of Climate*, 23, 2869-2884.

Yu., J.-Y. and S. T. Kim, 2011: Relationships between Extratropical Sea Level Pressure Variations and the Central-Pacific and Eastern-Pacific Types of ENSO, *Journal of Climate*, 24, 708-720.

Yu, J.-Y. and Y. Zou, 2013: The enhanced drying effect of Central-Pacific El Niño on US winter, Environmental Research Letters, 8, doi:10.1088/1748-9326/8/1/014019.

Zhang, L., and C. Wang, 2013: Multidecadal North Atlantic sea surface temperature and Atlantic meridional overturning circulation variability in CMIP5 historical simulations. J. Geophys. Res., 118, 5772–5791, doi:10.1002/jgrc.20390.

Zheng, Y., T. Shinoda, J.-L. Lin, G. N. Kiladis, 2011: Sea Surface Temperature Biases under the Stratus Cloud Deck in the Southeast Pacific Ocean in 19 IPCC AR4 Coupled General Circulation Models. J. Climate, 24, 4139–4164.

Zou, Y., J.-Y. Yu, T. Lee, M.-M. Lu, and S. T. Kim (2014), CMIP5 model simulations of the impacts of the two types of El Niño on the U.S. winter temperature, J. Geophys. Res. Atmos., 119, 3076–3092, doi:10.1002/2013JD021064.

NOAA Technical Report OAR CPO-2

REGIONAL CLIMATE PROCESSES AND PROJECTIONS FOR NORTH AMERICA: CMIP3/CMIP5 DIFFERENCES, ATTRIBUTION AND OUTSTANDING ISSUES

HOW TO CITE:

Sheffield, J., A. Barrett, D. Barrie, S.J. Camargo, E.K.M. Chang, B. Colle, D.N. Fernando, R. Fu, K.L. Geil, Q. Hu, X. Jiang, N. Johnson, K.B. Karnauskas, S.T. Kim, J. Kinter, S. Kumar, B. Langenbrunner, K. Lombardo, L.N. Long, E. Maloney, A. Mariotti, J.E. Meyerson, K.C. Mo, J.D. Neelin, S. Nigam, Z. Pan, T. Ren, A. Ruiz-Barradas, R. Seager, Y.L. Serra, A. Seth, D.-Z. Sun, J.M. Thibeault, J.C. Stroeve, C. Wang, S.-P. Xie, Z. Yang, L. Yin, J.-Y. Yu, T. Zhang, M. Zhao (2014), Regional Climate Processes and Projections for North America: CMIP3/CMIP5 Differences, Attribution and Outstanding Issues, NOAA Technical Report OAR CPO-2