

© Copyright 2014 American Meteorological Society (AMS). For permission to reuse any portion of this work, please contact permissions@ametsoc.org. Any use of material in this work that is determined to be "fair use" under Section 107 of the U.S. Copyright Act (17 U.S. Code §107) or that satisfies the conditions specified in Section 108 of the U.S. Copyright Act (17 USC § 108) does not require the AMS's permission. Reproduction, systematic reproduction, posting in electronic form, such as on a website or in a searchable database, or other uses of this material, except as exempted by the above statement, requires written permission or a license from the AMS. All AMS journals and monograph publications are registered with the Copyright Clearance Center (<https://www.copyright.com>). Additional details are provided in the AMS Copyright Policy statement, available on the AMS website (<https://www.ametsoc.org/PUBSCopyrightPolicy>).

How Well Are the Distribution and Extreme Values of Daily Precipitation over North America Represented in the Community Climate System Model? A Comparison to Reanalysis, Satellite, and Gridded Station Data

MELISSA GERVAIS, JOHN R. GYAKUM, EYAD ATALLAH, AND L. BRUNO TREMBLAY

Department of Atmospheric and Oceanic Science, McGill University, Montreal, Quebec, Canada

RICHARD B. NEALE

National Center for Atmospheric Research, Boulder, Colorado

(Manuscript received 31 May 2013, in final form 7 April 2014)

ABSTRACT

An intercomparison of the distribution and extreme values of daily precipitation between the National Center for Atmospheric Research Community Climate System Model, version 4 (CCSM4) and several observational/reanalysis data sources are conducted over the contiguous United States and southern Canada. The use of several data sources, from gridded station, satellite, and reanalysis products, provides a measure of errors in the reference datasets. An examination of specific locations shows that the global climate model (GCM) distributions closely match the observations along the East and West Coasts, with larger discrepancies in the Great Plains and Rockies. In general, the distribution of model precipitation is more positively skewed (more light and less heavy precipitation) in the Great Plains and the eastern United States compared to gridded station observations, a recurring error in GCMs. In the Rocky Mountains the GCMs generally overproduce precipitation relative to the observations and furthermore have a more negatively skewed distribution, with fewer lower daily precipitation values relative to higher values. Extreme precipitation tends to be underestimated in regions and time periods typically characterized by large amounts of convective precipitation. This is shown to be the result of errors in the parameterization of convective precipitation that have been seen in previous model versions. However, comparison against several data sources reveals that model errors in extreme precipitation are approaching the magnitude of the disparity between the reference products. This highlights the existence of large errors in some of the products employed as observations for validation purposes.

1. Introduction

From a socioeconomic perspective, precipitation is one of the most important variables to predict in future climate, because of its implications for water resources and natural disasters. For these purposes, it is not simply the average precipitation that is important but rather the spatial and temporal distribution of precipitation intensities. One aspect of this distribution of crucial importance to society is extreme precipitation. For example, heavy precipitation can have large impacts on crops both because of the direct impact of flooding as

well as the negative consequences of excess soil moisture (Rosenzweig et al. 2002; Gornall et al. 2010). A recent study by Soccimarro et al. (2013) showed a 2%–5% increase in the intensity of extreme events over land in the future, using an ensemble average of models from phase 5 of the Coupled Model Intercomparison Project (CMIP5). The use of global climate models (GCMs) for prediction of changes in the distribution of precipitation requires continual assessment of the ability of these GCMs to represent the distribution in the current and historical climate. GCMs are well known to have significant errors in the distribution of precipitation, historically precipitating too frequently and with too low of an intensity when compared to observations (Dai et al. 1999; Iorio et al. 2004; Sun et al. 2006), even when the mean precipitation is well represented (Chen et al. 1996; DeMott et al. 2007).

Corresponding author address: Melissa Gervais, Department of Atmospheric and Oceanic Sciences, McGill University, 805 Sherbrooke Street West, Montreal QC H3A 0B9, Canada.
E-mail: melissa.gervais@mail.mcgill.ca

The convective parameterizations in GCMs tend to be associated with large errors in model precipitation, in particular in the extreme precipitation. Several studies have been conducted examining the diurnal cycle of precipitation in previous versions of the National Center for Atmospheric Research (NCAR) Community Climate System Model, version 3 (CCSM3) and have shown that the convective parameterization triggers convection too early and frequently, which does not allow for the buildup of convective available potential energy (CAPE) necessary for heavier rainfall events (Dai and Trenberth 2004; DeMott et al. 2007). Although increasing resolution does aid in many aspects of the representation of precipitation, improvements to the parameterization are necessary to solve this problem of timing and consequently intensity (Iorio et al. 2004; Dirmeyer et al. 2012). In CCSM4, a more recent version of the NCAR GCM, improvements have been made to the parameterization of deep convection to include convective momentum transport and dilution due to entrainment in the calculation of CAPE (Gent et al. 2011; Neale et al. 2013). Gent et al. (2011) compared the simulation of daily precipitation frequency over land between 20°N and 20°S in the CCSM4 to the CCSM3. The CCSM4 showed significant improvement, even when run at the CCSM3's lower resolution, which they attribute to improvements to the deep convection scheme (Gent et al. 2011). However, there are significant differences in the convection that occurs in the tropics versus that which occurs in the midlatitudes.

In this study, we conduct a validation of the precipitation distribution and extreme values in the more recent versions of the NCAR GCM: the fully coupled CCSM4; the Community Atmosphere Model, version 4 (CAM4); and CAM5. Since these newer versions of the NCAR GCM are improved over the previous version (CCSM3) both in terms of their parameterizations and resolution, their representation may be closer to observations than previous versions. As the ability of GCMs to produce accurate precipitation fields increases, we need to consider the validity of the common assumption used in the validation of GCMs that observational errors are smaller than model errors. It is therefore important to consider the errors within the reference datasets used for validation, as well as the methods used to compare models to reference data. To this end, the validation will be made against three observational or reanalysis products to help constrain the extent of observational errors, which we may think of as the disparity between various reference data sources. A remapping method is used to regrid between various resolutions instead of an interpolation to be consistent with the interpretation of a GCM being an area average of precipitation (Chen

and Knutson 2008; Gervais et al. 2014). Furthermore, all statistics are computed after the remapping procedure to ensure that there are no mismatches in scales as can occur when an index at a point location is remapped to a different resolution (Kursinski and Zeng 2006), an issue that Sillmann et al. (2013) cited in their analysis. Focusing on the CCSM allows for more detailed error analysis both spatially and seasonally, which helps to elucidate the abilities and limitations of the models in their representation of various precipitation mechanisms.

2. Data

a. North American Amalgamated Precipitation

We employ two currently available datasets: the Climate Prediction Center daily Unified Precipitation Dataset (UPD), provided by the National Oceanic and Atmospheric Administration/Earth System Research Laboratory (Xie et al. 2007), and the daily 10-km Gridded Climate Dataset for Canada (GCDC), provided by the National Land and Water Service (Hutchinson et al. 2009), to create a gridded precipitation dataset over a contiguous region in North America. The native grid types and spacings of the two datasets differ, with the UPD being on a 0.25° latitude–longitude grid and the GCDC being on a 10 km \times 10 km Cartesian grid. We amalgamate these two datasets over their common time period of 1961–2006 to create the North American Amalgamated Precipitation (NAAP) dataset, where the UPD covers the contiguous United States and the GCDC covers Canada south of 60°N.

The methods used to grid station data differed between the UPD and the GCDC. For the reader's reference, we will briefly outline these gridding methods used by the data creators of the UPD and the GCDC. The UPD was created through the optimal interpolation of 24-h precipitation accumulations from gauge-based measurements in the continental United States (Xie et al. 2007). The method of interpolation was conducted in two steps. First, a daily precipitation climatology was created by summing the first 6 harmonics of station data time series for stations with reporting rates over 80% during the period 1978–97 (Xie et al. 2007). This daily precipitation climatology was then interpolated using the Shepard (1968) method, onto the 0.25° latitude–longitude analysis grid (Xie et al. 2007). An orographic correction was conducted on the daily climatology due to a general bias toward lower precipitation in mountainous regions, which results from a bias in station locations toward lower elevations in these regions (Xie et al. 2007). The Parameter-Elevation Regressions on Independent Slopes Model (PRISM) monthly precipitation climatology (Daly et al. 2002), which is adjusted

for orographic effects using empirical relationships that are established locally and are available over the continental United States, was used to conduct an orographic adjustment of the daily climatology. The correction to the daily UPD climatology was done by scaling the daily climatology so that its monthly accumulations closely match that of the PRISM monthly accumulations while preserving the variability of the daily climatology (Xie et al. 2007). In the second step, the ratio of the daily station data over the uncorrected gridded daily climatology was calculated at each station location (Xie et al. 2007). Using the interpolated climatological field for the ratio allows stations to be used even if their observation length is too short to be included in the climatology (Xie et al. 2007). This ratio was then interpolated using the optimal interpolation of Gandin (1965) to the analysis grid (Xie et al. 2007). The interpolation of this ratio was conducted because this field is smoother in space than the daily station data themselves and thus results in less errors when interpolated (Xie et al. 2007). Multiplying the interpolation ratio by the orographically adjusted daily climatology at each analysis grid point then yields the desired interpolated daily precipitation field (Xie et al. 2007).

The GCDC used a trivariate thin-plate smoothing spline (Hutchinson 1995) to interpolate 24-h precipitation accumulations from Environment Canada, creating a $10\text{ km} \times 10\text{ km}$ gridded precipitation dataset for Canada south of 60°N . In this method, the elevation was defined using a digital elevation model, and a scaling factor was applied to increase precipitation with elevation. First, a binary field of precipitation occurrence was created from which grid points with and without precipitation were determined. Second, a precipitation surface was created through the interpolation of station data that had precipitation. In this step, the square root of the precipitation value was interpolated instead of the full value as it is more normally distributed. The final interpolated precipitation field was equal to the precipitation surface for grid points that were determined to be precipitating and zero where it was deemed nonprecipitating. Hutchinson et al. (2009) find that the errors in the GCDC dataset are modest for seasonal and annual averages but are relatively large for daily precipitation and extremes, even in the southern portion of the data where the station density is highest. This was attributed in part to the high spatial variability and low data coverage (Hutchinson et al. 2009).

To combine the UPD and the GCDC into a single dataset we linearly interpolate the GCDC from a $10\text{ km} \times 10\text{ km}$ grid to a 0.25° latitude–longitude grid. The grid spacing of the GCDC is much higher than the UPD (>2.5 times) and so the distances over which the interpolation is conducted are short, and as a result we

expect that additional errors associated with using a simple linear interpolation over a more complex method should be small. The two datasets are then combined for the years 1961–2006 (when both datasets are available), where the GCDC covers Canada and the UPD covers the United States. In the Great Lakes regions, the coverage is split between the GCDC and UPD datasets. Attempts were not made to smooth the boundary between the two datasets, as this could introduce errors in the higher-order statistics of the combined dataset. The NAAP is the final product of this merger between the GCDC and the UPD.

b. Global Precipitation Climatology Project One-Degree Daily

Global Precipitation Climatology Project One-Degree Daily (GPCP 1DD) is a satellite-derived precipitation dataset at 1° latitude–longitude resolution from 1997 to 2008. This dataset should not be confused with the two other GPCP products, the GPCP version 2 satellite–gauge monthly precipitation dataset or the GPCP satellite–gauge pentad dataset. The GPCP 1DD is based on two satellite products: namely, the threshold-matched precipitation index (TMPI) for the regions between 40°N and 40°S and the Television and Infrared Observation Satellite Operational Vertical Sounder (TOVS) Pathfinder Path A outside of this region (Huffman et al. 2001). The GPCP 1DD will subsequently be referred to as the GPCP.

The TMPI uses 3-hourly brightness temperatures determined from infrared radiometers mounted on geosynchronous satellites, where precipitation is deemed to be occurring if the brightness temperature is below a threshold value and it is assumed to occur at a constant specified rain rate. Low brightness temperatures are indicative of ice particles in clouds, which have a relatively weak relationship to precipitation occurrence (Huffman et al. 1997). This method is mostly useful in regions of deep convection between 40°N and 40°S (Huffman et al. 2001). The threshold brightness temperature and conditional rain rates vary on a monthly basis and are calculated using information from the Special Sensor Microwave Imager and the GPCP version 2 satellite–gauge monthly precipitation dataset (Huffman et al. 1997).

The TOVS dataset is based on a relationship between cloud-top pressure, fractional cloud cover, a profile of relative humidity, and precipitation that is empirically determined using collocated gauge measurements (Susskind et al. 1997). The TOVS data go through several processing steps to remove biases relative to the TMPI at the border between the two datasets. The TOVS data are rescaled by setting low values of precipitation to zero so that the value of the total number of rainy days at the border with the TMPI are equal

(Huffman et al. 2001). The precipitation amounts are rescaled during precipitating days so that the total amounts match the GPCP version 2 satellite–gauge dataset (Huffman et al. 2001). Finally, a smooth transition is created at the border of the TMPI and the TOVS on a daily basis by calculating the difference between the TOVS and TMPI data at the border and then adding a function to the TOVS that is this difference at the 40°N and 40°S border decreased linearly to 0 at 50°N and 50°S, respectively (Huffman et al. 2001).

There are many known issues with satellite data in reproducing accurate daily values, particularly with respect to the frequency of events and the magnitude of extremes (Sun and Barros 2010). Global daily precipitation products are rare and so the GPCP is an appealing dataset and has been used in numerous studies as a source of daily precipitation observations for the validation of GCM output (e.g., Emori et al. 2005; DeMott et al. 2007; Scoccimarro et al. 2013). The creators of the GPCP, however, suggest that the data only be used for time mean calculations because of errors in the daily amounts (Huffman et al. 2001). The GPCP will be included in this analysis since it often employed in the literature. However, the potential for errors will be indicated through comparison with the NAAP data, since they will be used without the spatial or temporal averaging advised by the data creators.

c. Climate Forecast System Reanalysis

The National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR) is a coupled global reanalysis spanning the period 1979–2010 at 0.5° latitude–longitude grid resolution (Saha et al. 2010). In this analysis, 6-hourly precipitation totals are summed into daily precipitation totals to be consistent with the NAAP dataset. The CFSR assimilates atmospheric variables and is at a higher resolution, which should improve its performance over a GCM. Regions with a dense sounding network, such as the United States, are likely to have better represented mass fields and therefore more accurate precipitation than other regions. The precipitation output from the CFSR, however, is produced solely by the CFSR background precipitation model and does not assimilate any precipitation observations (CFS Team 2011, personal communication). Consequently, it is likely to suffer from similar types of issues in the production of precipitation as GCMs.

d. Community Climate System Model, versions 3 and 4

The CCSM4 consists of four component models: the Community Atmosphere Model, version 4; the Community Land Model, version 4; the Parallel Ocean

Program, version 2; and the Los Alamos sea ice model (CICE). The component models are coupled at every atmospheric time step, except the ocean component model, which is coupled once per day. We use a pre-industrial control run of the CCSM4 with additional output (MOAR), forced with historical International Panel on Climate Change (IPCC) values for incoming solar radiation, carbon dioxide, and aerosols. The resolution of the atmosphere component model CAM4 in the CCSM4 control run is $0.9^\circ \times 1.25^\circ$ latitude–longitude and it has 26 levels in the vertical using hybrid sigma–pressure coordinates (similar to CCSM3) (Gent et al. 2011).

For comparison with the CCSM4, we also use a CCSM3 preindustrial control run from the CMIP3 experiment. The resolution of the CCSM3 model run at T85 spectral truncation is approximately 1.4° latitude–longitude and also has 26 levels in the vertical. More details regarding this version of the model can be found in Collins et al. (2006).

The deep convection scheme in the model follows Zhang and McFarlane (1995), consisting in general of an ensemble of entraining plumes and compensating downdrafts at various heights that occur when the atmosphere is unstable. In CAM4, this scheme was revised to include impacts of convective momentum transport and the calculation of CAPE to now be diluted through entrainment (Neale et al. 2013). The closure assumption of the deep convection scheme is that CAPE is consumed at an exponential rate. The inclusion of entrainment in the calculation of CAPE can reduce its value and improve the vertical moisture structure. The land model also received improvements that could aid in the representation of precipitation (Gent et al. 2011). More details on the physics, parameterization, and their improvements in this model version can be found in Gent et al. (2011) and Neale et al. (2013).

e. Community Atmosphere Model, versions 4 and 5

To evaluate the ability of the atmosphere only model, output from an Atmospheric Model Intercomparison Project (AMIP)-style control run of CAM4 and CAM5 will be used. These runs are conducted at the same resolution as the CCSM4 run but using time-varying SSTs and sea ice in addition to the IPCC forcings. Although the CAM5 was produced shortly after the CAM4, there are changes to the shallow convection scheme and in the representation of aerosol indirect effects. Further details can be found in Neale et al. (2010).

3. Methods

Following Chen and Knutson (2008), model output precipitation is interpreted in this study as an area

TABLE 1. Datasets, their acronyms, data types, native grid sizes, and native time periods. Note that for all intercomparisons the resolution is remapped to the CCSM4 resolution and the time period is limited to 1979–2005, except for the GPCP, which is limited to 1997–2005.

Dataset	Acronym	Data type	Native grid size	Time period
North American Amalgamated Precipitation	NAAP	Gridded station observations	0.25° latitude–longitude	1948–2006
Global Precipitation Climatology Project One-Degree Daily	GPCP 1DD	Satellite observations	1.0° latitude–longitude	1997–2007
NCEP Climate Forecast System Reanalysis	CFSR	Reanalysis	0.5° latitude–longitude	1981–2010
Community Atmosphere Model, version 4	CAM4	GCM atmospheric component model	0.9° × 1.25° latitude–longitude	1979–2010
Community Atmosphere Model, version 5.1	CAM5	GCM atmospheric component model	0.9° × 1.25° latitude–longitude	1979–2006
Community Climate System Model, version 4	CCSM4	Fully coupled GCM	0.9° × 1.25° latitude–longitude	1960–2005

average of precipitation within a model grid box. To make consistent comparisons between the various gridded observational, reanalysis, and model data, we ensure that all datasets are remapped from their native resolutions (Table 1) to the grid size of the lowest-resolution data used in the comparison, in a manner that is consistent with the area average view point of a grid box. In this study, unless the 1.4° latitude–longitude CCSM3 is also being compared, the lowest-resolution data are the CCSM4, CAM4, and CAM5 model outputs, which are all on a 0.9° × 1.25° latitude–longitude grid, referred to as 1°. The resolution change is accomplished here using a first-order conservative remapping method from the Spherical Coordinate Remapping and Interpolation Package (SCRIP) from the Los Alamos National Laboratory (Jones 1999). The method computes weights for each input grid point based on the area overlap between the input grid boxes and the output grid boxes. Multiplication of the input grid precipitation field by these weights regrids the dataset while conserving the total amount of precipitation.

Several metrics for the distribution of precipitation are utilized. Kursinski and Zeng (2006) showed that the order of operations of the computation of precipitation indices versus spatial averaging is important. In all cases, the statistics are calculated on the specified grid after the interpolation or remapping procedure, which is consistent with the method in Chen and Knutson (2008). If the order of this operation is reversed, results of errors in GCMs may be ambiguous, as was found for example in Sillmann et al. (2013). All statistics are calculated during the time period of interest (annual or bimonthly) for a single year; then, the value of each year in the time period is averaged. This climatological averaging time period is 1979–2005 for all datasets, except for any computation involving the GPCP, where it is 1997–2005.

There are small changes when all results are averaged over the 1997–2005 period; however, they do not impact the interpretation of our results. The bimonthly periods used are January–February (JF), March–April (MA), May–June (MJ), July–August (JA), September–October (SO), and November–December (ND). Given the non-Gaussian distribution of precipitation, median and percentile values will be used as metrics of precipitation. The 97th percentile of precipitation in a given period for each year averaged over all years (1979–2005 or 1997–2005) will subsequently be referred to as the climatological extreme precipitation.

The full distribution of precipitation is represented using two metrics. First, the empirical cumulative distribution function (CDF) for all days is computed for each year and then it is climatologically averaged. The CDF is a common metric used for validation of GCMs, which shows the cumulative probabilities of increasing daily precipitation amounts. The significance of differences between CDFs is determined using the Kolmogorov–Smirnov (KS) (Massey 1951; Stephens 1970) and Cramer–Von Mises (CvM) tests (Anderson and Darling 1952; Anderson 1962). These two tests check the null hypothesis that the CDF of the test data are from the same population as the NAAP using various metrics of the difference between the distributions of the full precipitation time series (not the annual distribution climatologically averaged). Second, we show the annual total precipitation versus bins of daily precipitation accumulation averaged climatologically, which we call the total mass distribution (TMD). This shows the contribution that each daily accumulation range has toward the total precipitation. The TMD is a physically intuitive metric that is useful in understanding the importance of heavy precipitation events. The probability of heavy events is low but they are more significant in terms of

their contribution toward the total precipitation, although in the CDF they represent relatively small changes. The skewnesses of the TMDs are used to demonstrate the spatial coherence of the differences in the shape of these functions between the various precipitation datasets. Generally, the TMDs of precipitation are positively skewed, since light precipitation occurs more frequently than heavier precipitation. A more positive skewness value translates to less high and more low daily precipitation amounts. A Monte Carlo method is used to test the hypothesis that the skewnesses of the TMDs are the same, relative to the NAAP. The data from the NAAP and the test dataset are randomly reassigned to two new datasets, 1000 times. The difference in skewness between the NAAP and test datasets are considered significant if outside the range of 5th to 95th percentiles of differences found using the Monte Carlo test.

4. Results

The three reference products, the NAAP, GPCP, and CFSR, each have advantages and disadvantages for use in the validation of GCMs. The NAAP data are based on direct station measurements and are thus considered to be the closest to the truth. The disadvantages of gridded station datasets like the NAAP are that they can have biases when the station density is low (Gervais et al. 2014); they can suffer from measurement error, particularly for solid precipitation (Goodison et al. 1998; Cherry et al. 2007); and they are only available over continents. The GPCP and the CFSR have the advantage that they are available globally. However, these datasets have the potential for large errors as the GPCP is created using indirect measurements and the precipitation in the CFSR is produced by a model. In this study we are using all of these datasets to evaluate precipitation in the CCSM model in North America. Focusing on North America allows for analysis of regional precipitation errors in more detail. Furthermore, satellite products like the GPCP and reanalysis products such as the CFSR are commonly used to validate models in regions where station data are not available (e.g., Sillmann et al. 2013; Shiu et al. 2012). Consequently, this study has the advantage of comparing these reference products for validation purposes in a region with typically good station observations.

To evaluate the potential for biases in the NAAP due to low station density, we apply results from an experiment by Gervais et al. (2014), who examined the impact of station density on precipitation statistics in the United States. Their experiment consisted of the interpolation and remapping of station data using the same methodology as in this study but conducting the gridding

repeatedly with successively fewer stations. In doing so, they are able to infer biases resulting from decreased station density. They normalize these results across a region to produce a distribution of potential biases with respect to station density. They found that the impact of station density on biases in precipitation statistics changed seasonally and regionally. In general, a small (large) decorrelation length scale of station data and a low (high) spatial homogeneity of station statistics result in larger (smaller) biases (Gervais et al. 2014). Experimentally derived upper and lower bound curves of potential biases in climatological median and extreme precipitation over the entire United States are created by fitting a curve of the form

$$y = ae^{bx} + ce^{dx} \quad (1)$$

to the 99th and 1st percentiles of this distribution (Gervais et al. 2014). This provides a measure of the upper and lower bound of biases that takes into account the full breadth of biases across the United States.

In this study, the experimentally derived upper and lower bound curves of potential bias in climatological precipitation versus station density in gridded station data in the United States of Gervais et al. (2014) are applied to the station density of the NAAP (Fig. 1). This is used to produce an estimate of biases in the NAAP due to station density (Fig. 2). According to Hutchinson et al. (2009), the GCDC dataset is the most complete daily dataset of Canadian precipitation for this time period, though the station density is smaller than any of the non-Canadian datasets that they came across. The density of stations used in the NAAP is very heterogeneous with the most stations located in more densely populated regions, fewer stations in mountainous regions, and scarce stations in northern Canada (Fig. 1). As a result, there are very large biases in much of Canada (Fig. 2). There are additional sources of bias in the NAAP due to station measurement error, which has been shown to be on the order of 10% for liquid (Adam and Lettenmaier 2003) and on the order of 100% for solid precipitation (Goodison et al. 1998; Cherry et al. 2007). In this analysis the NAAP is generally considered to be the closest representation of the true precipitation field; however, potential biases associated with lower station density will still be considered.

Examining the average annual median precipitation of the observational and reanalysis data provides a general idea of the magnitude and patterns of observational errors in these datasets. The GPCP has a similar pattern in the annual median precipitation as the NAAP, although there is a lack of detail in the western mountain ranges in the GPCP (Figs. 3b,c). The biggest difference

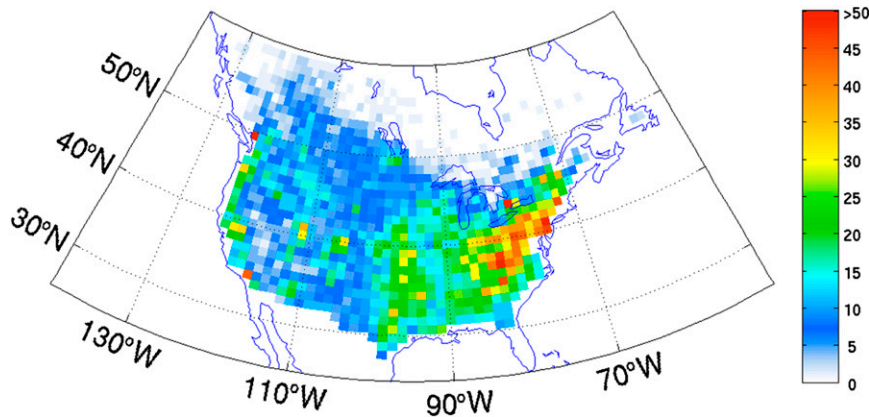


FIG. 1. Number of stations per CCSM4 grid box ($0.9^\circ \times 1.25^\circ$ latitude–longitude), averaged over the years 1975, 1985, and 1995. Values are rounded to the nearest integer.

between the GPCP and that NAAP is that generally the magnitude is higher in the GPCP and there is a discontinuous decrease in the median precipitation in the eastern United States when the input data source changes from the TMPI (south) to the TOVS (north) at 40°N (Figs. 3b,c). At this boundary, the percent errors in median precipitation in the GPCP relative to the NAAP drop from 40%–60% to 0%–20%. The pattern of error in the CFSR relative to the NAAP is very different than that of the GPCP (Figs. 3c,e). Differences in the CFSR relative to the NAAP include a reduction in area and eastward shift of the region of higher annual median precipitation in the southeastern United States (Figs. 3d,e). However, the errors in the eastern United States and Canada are typically less than 20%. The CFSR has higher annual median precipitation in the Rocky Mountains, stretching from the United States to Canada and in much of Canada, than the NAAP (Figs. 3b,c), with percent errors up to 60%. The potential of errors due to station density in the NAAP in these regions are high for the climatological median and the extreme precipitation; however, the errors in the CFSR relative to the NAAP in the Rockies are greater than the upper bound of these climatological errors (Fig. 2).

There are some large climatological errors in the average annual median precipitation between the models and the NAAP; however, the results are generally very promising. In certain regions, such as the East Coast, the magnitude of the percent climatological error relative to the NAAP in the average annual median precipitation is lower than that of the GPCP in the eastern United States and Canada (Fig. 3). If we consider observational error to be the difference between reference products, this implies that the models are within observational error in these regions. One area of substantial errors in all the model runs is an underestimation of the

median precipitation in the southeastern United States, with up to 40% difference relative to the NAAP, which is smaller in the CCSM4 than in the CAM4 (not shown) and CAM5 (Figs. 3f–i). There are some general biases toward higher median precipitation along the West Coast and interior mountain ranges in both models, which are higher in the CCSM4 than the CAM5 (Figs. 3f–i). In Canada, the CAM4 (not shown), CAM5 and CCSM4 generally perform well compared to the NAAP except for an overestimation over the Rocky Mountains and some higher values of median precipitation in the north that are similar to those of the CFSR and smaller than those of the GPCP (Figs. 3f–i).

In addition to errors in the median field, we are interested in how well the distribution of precipitation is represented. Specific locations are chosen to use as examples of these distributions, where the locations are geographically diverse and include sample points with different precipitation climatologies (Fig. 4). How representative these points are of the area around them depends on the spatial homogeneity of the precipitation distribution, which is typically higher in the east than in the west (not shown). The CDFs and TMFs are shown in Figs. 5 and 6. For the CDF, we can apply the KS and CvM tests of significance, which are used here to determine whether products have the same distribution as the NAAP. The CDFs at these locations are significantly different for all datasets, a result that is generally true apart from some isolated locations (not shown). For the TMDs we use a Monte Carlo method to determine whether differences in the skewness of the TMDs, relative to the NAAP, are significant. The skewnesses of the TMDs and their significances are shown for the entire region of study in Fig. 7.

The northern and southern West Coast points are within a coastal region with predominantly orographic

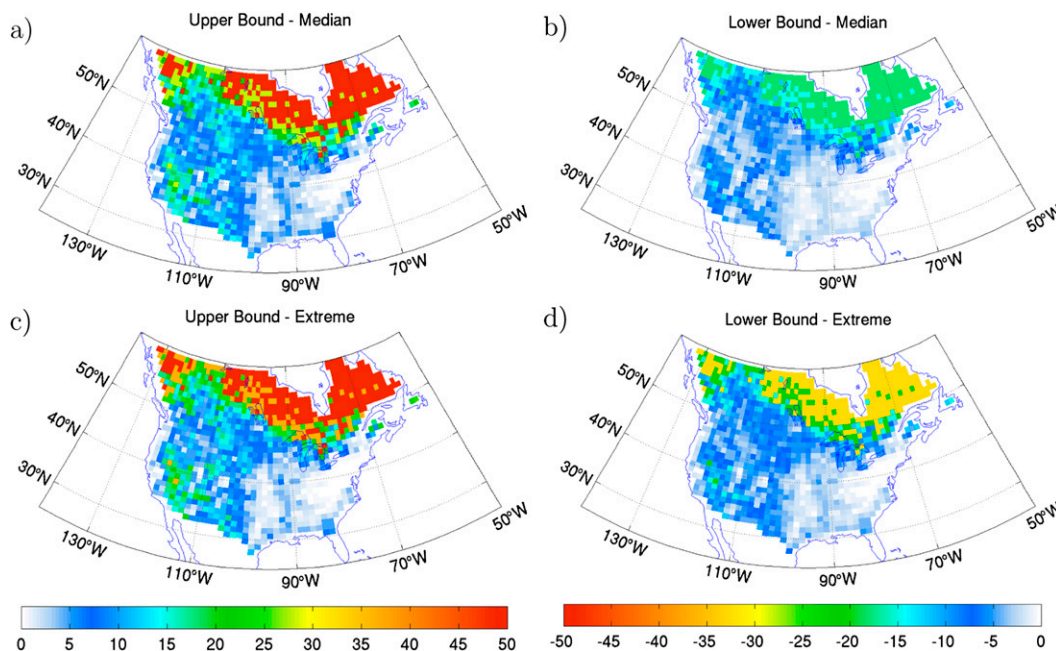


FIG. 2. (a),(c) Upper and (b),(d) lower bound on the percent bias in climatological annual (a),(b) median and (c),(d) extreme precipitation for the NAAP data using an experimentally derived relationship between upper and lower errors bounds and station density found in Gervais et al. (2014). Note that the color scales are reversed between the upper and lower bound maps such that the magnitude of the color schemes are identical but in opposing directions.

precipitation. The seasonality in these locations is tied to the intensity of the storm track, with higher precipitation in the winter when the storm track is more intense. The largest discrepancies in the CDFs of the northern West Coast point are that the GPCP and the CAM5 have too many nonprecipitating and light precipitation days (Fig. 5a). This results in precipitation totals in the GPCP being deficient throughout much of the range of daily amounts (Fig. 6a). For the CFSR, CCSM4, and CAM4 CDFs, we see somewhat higher probabilities beginning from the 1 mm day^{-1} until the 20 mm day^{-1} amounts or a shallower slope in probability over this time, which results in lower precipitation totals for these amounts compared to the NAAP (as seen in Fig. 6a). In general, for the northern West Coast points the TMF of the models are bracketed by reference products throughout the entire distribution, except the CAM5 for very low daily amounts (Fig. 6a). The southern West Coast point typically has much less precipitation compared to other locations (Fig. 6b). In relation to the CDF, the TMD can be thought of as an integration, over a range of daily amounts, of the change in probability times the amount and the number of days. The CDFs and the TMDs show that the CCSM4 has a higher number of precipitating days resulting in a general over production across all intensities for the southern West Coast (Fig. 6b). The CAM5 also has some notable issues at this location, with

too many nonprecipitating and light precipitating days until around the 5 mm day^{-1} amount.

The two Rockies points (western and eastern) are inland but are similarly in a region of predominantly orographic precipitation. The precipitation intensity in the Rockies is typically lighter and the annual total is very low compared to other locations. The CDFs for the two Rockies points show that the NAAP has more rain-free days than any of the other products (Figs. 5c,d), which results in their being more precipitation and a shift toward higher amounts in the other products (Figs. 6c,d). The exception is the GPCP for the eastern Rockies point where there are less precipitating days at amounts $>10 \text{ mm day}^{-1}$, where the slope of the CDF curve levels off (Figs. 5d, 6d). For the GPCP, the western Rockies point is located south of 40°N and the eastern Rockies point is located north of 40°N (Fig. 4), where the data source changes from the TMPI to the TOVS, respectively. The changes in the steepness of the CDF slope and the skewness of the TMD may be symptomatic of changes in data source for the GPCP, specifically since the distribution is skewed in the data processing and adjusted to produce a smoother boundary (Huffman et al. 2001). However, as discussed previously, the Rockies typically have a lower density of station observations and the NAAP precipitation is adjusted for orography. There is thus a possibility that the errors in

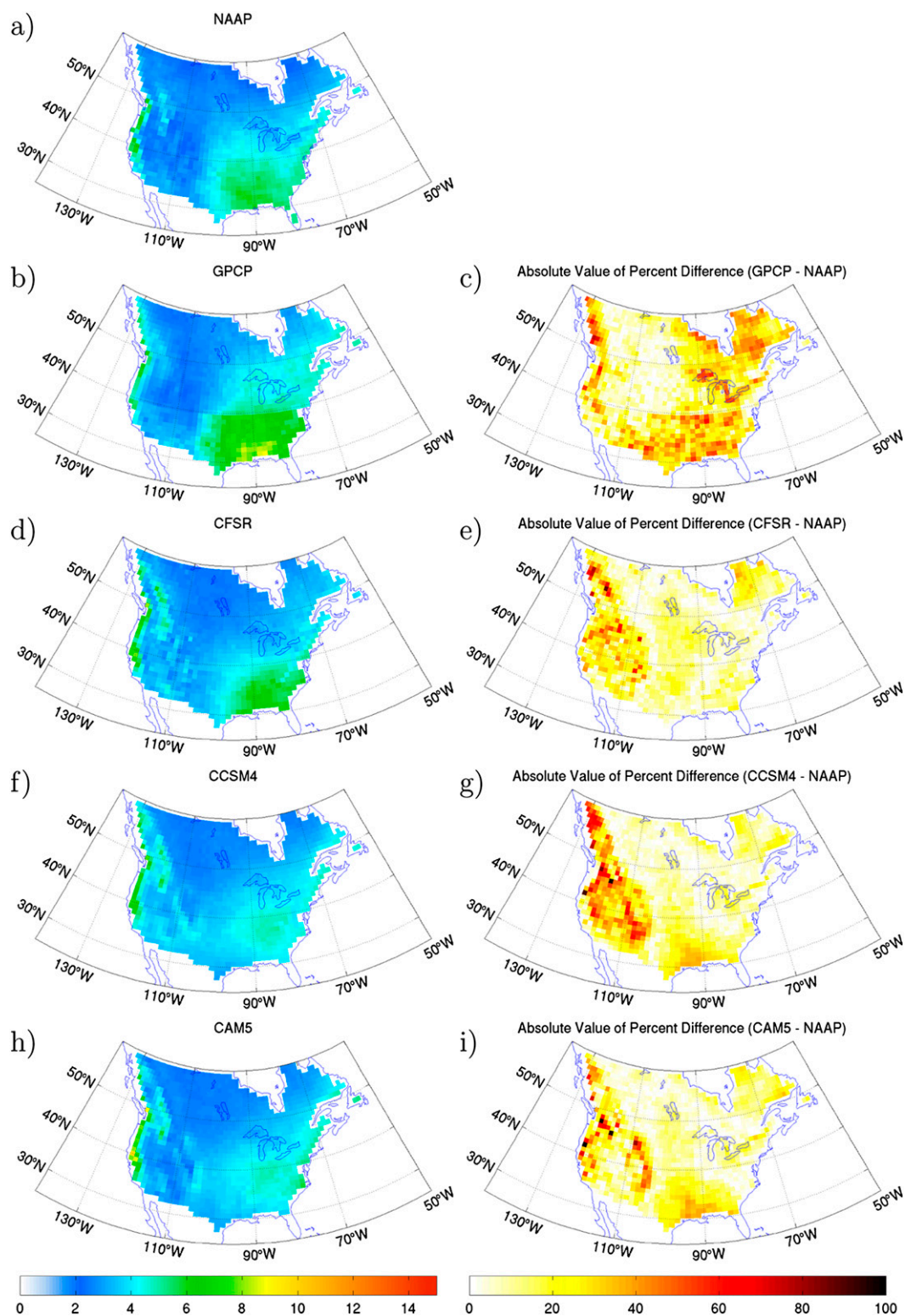


FIG. 3. (left) Climatological annual median precipitation (mm day⁻¹) for (a) NAAP, (b) GPCP, (d) CFSR, (f) CCSM4, and (h) CAM5 and (right) absolute value of percent anomaly relative to the NAAP {e.g., $|(GPCP - NAAP)/NAAP| \times 100\}$ in climatological annual median precipitation (mm day⁻¹) for (c) GPCP, (e) CFSR, (g) CCSM4, and (i) CAM5.

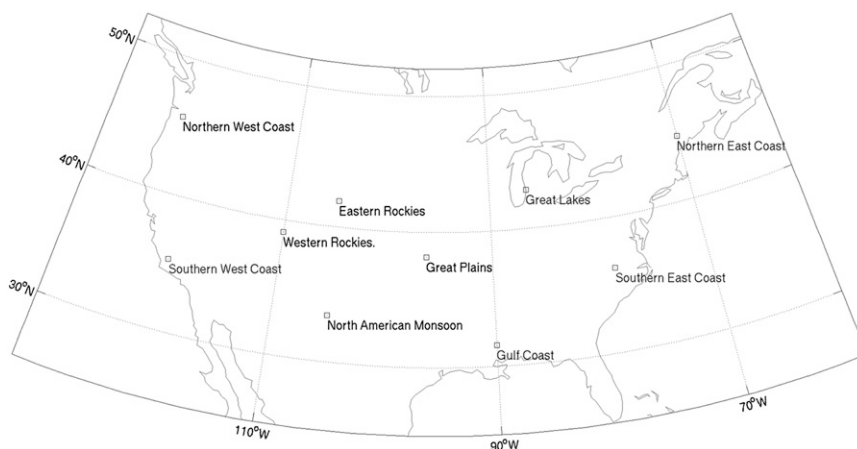


FIG. 4. Location of grid points used to examine the distribution of precipitation.

the median and distribution in the region are related to errors in the NAAP either because of station density or to the orographic adjustment. In such regions it is difficult to determine what is observational errors and what is model error; however, the general similarity between the NAAP and the GPCP for the western and the NAAP and the CFSR for the eastern Rockies points for the TMD suggests that the NAAP is performing well (Figs. 6c,d).

The North American monsoon point is in a region that has low annual precipitation amounts and is named for the North American monsoon, a period of enhanced precipitation during the late summer/early fall. The CDF reveals that there are more precipitating days in general in the models than the NAAP (Fig. 5e) and the opposite for the other reference products. This results in TMDs with amounts of precipitation that are consistently higher in the models than any of the reference datasets (Fig. 6e), where we can also see that the TMD is skewed toward higher amounts. The Great Lakes point, on the other hand, is better represented by the models than either the CFSR or the GPCP for the CDF and the TMD (except the CAM5 for some amounts; Figs. 5f, 6f). This is particularly notable in the CDF for nonprecipitating and low precipitation amounts (Fig. 5f). This implies that the CAM4 and CCSM4 models are within observational error at this location.

The east coast of North America is influenced by the proximity to the Atlantic Ocean and the Appalachian mountain range. In the summer, the region receives predominantly convective precipitation. During July–August and September–October, it can experience heavy precipitation from tropical cyclones that can impact anywhere along the coast. In the winter it tends to experience larger-scale precipitating systems, with the most notable being nor'easters. The greatest difference

in the CDFs of the northern East Coast point is that the GPCP has too few nonprecipitating days followed by a steeply sloping probability of amounts up to 20 mm day^{-1} (Fig. 5g), which results in an overabundance of daily amounts until the $20\text{--}25 \text{ mm day}^{-1}$ bin and an underproduction of heavier amounts (Fig. 6g). Similar to the eastern versus western Rockies where the two points are located on opposite sides of the 40°N change in GPCP data source, the TMD of the GPCP shows an excess of lighter and deficit of heavy events in the northern East Coast and the opposite in the southern East Coast (Figs. 6g,h), as well a change from an underabundance in the north to an overabundance in the south of days without rainfall ($<1 \text{ mm day}^{-1}$; Figs. 5g,h). For the model TMDs, there tends to be more high amounts than the other reference products; for the CAM5, there tends to be an underproduction of low precipitation amounts. The behavior of the models are also the opposite for the southern East Coast relative to the northern East Coast, where we see a positively skewed TMD at the southern East Coast point for all of the models relative to the reference products (Fig. 6h).

The Great Plains and Gulf Coast regions experience a great deal of convection. For the Gulf Coast this is true for most of the year, whereas convection is particularly important in the spring/summer seasons in the Great Plains. The Great Plains and Gulf Coast points also have the greatest errors in their precipitation distributions. For the Great Plains point, all model-based products (CFSR and GCMs) have too much light precipitation and not enough heavy precipitation, seen in both the CDFs and the TMDs (Figs. 5i, 6i). This could be associated with in the GCMs' inability to produce heavy convective precipitation events. The Gulf Coast point sees similar errors from the GCMs, but the CFSR performs much better than in the Great Plains in comparison

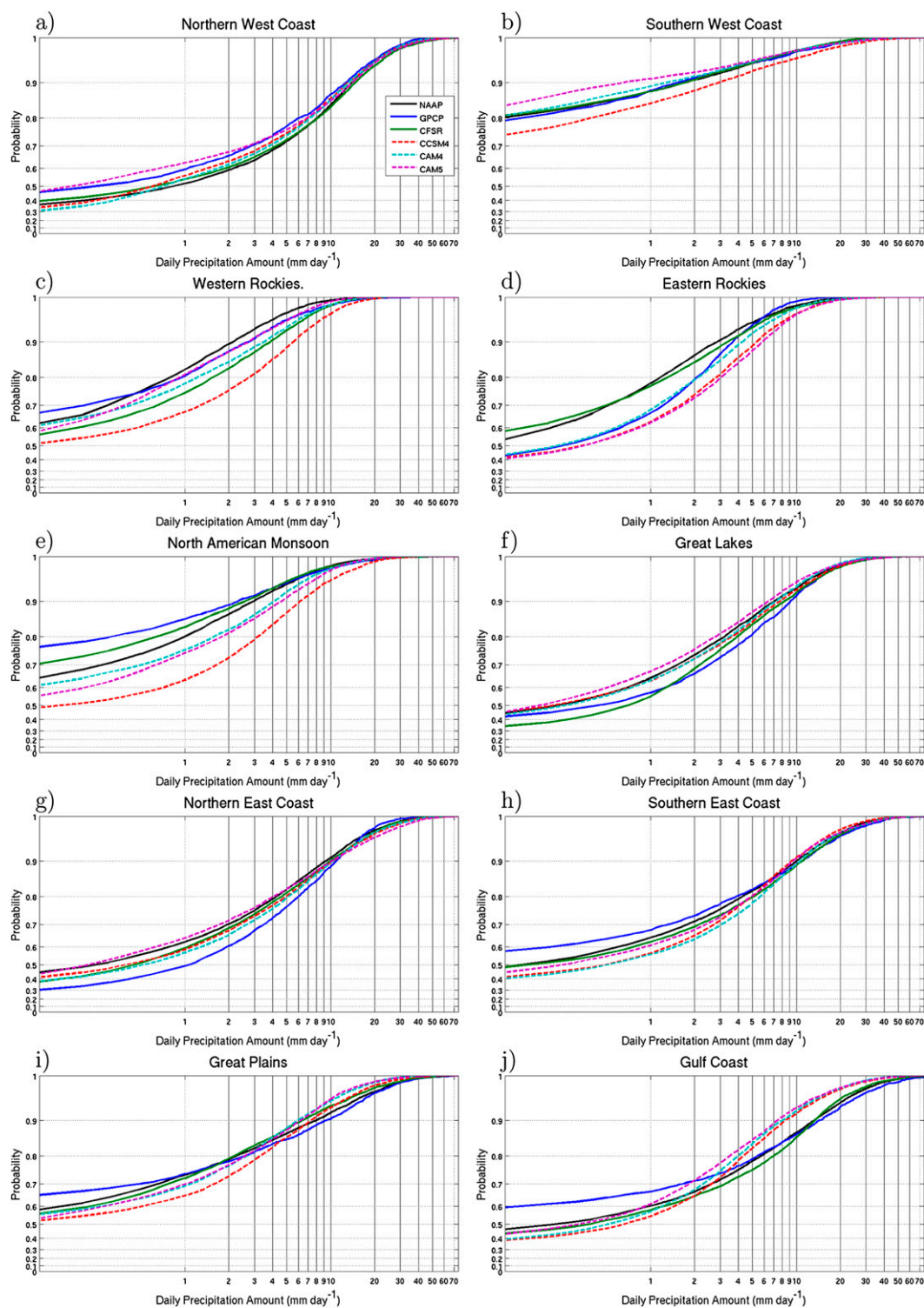


FIG. 5. Climatological annual empirical cumulative distribution function over all days at point locations as indicated in Fig. 4.

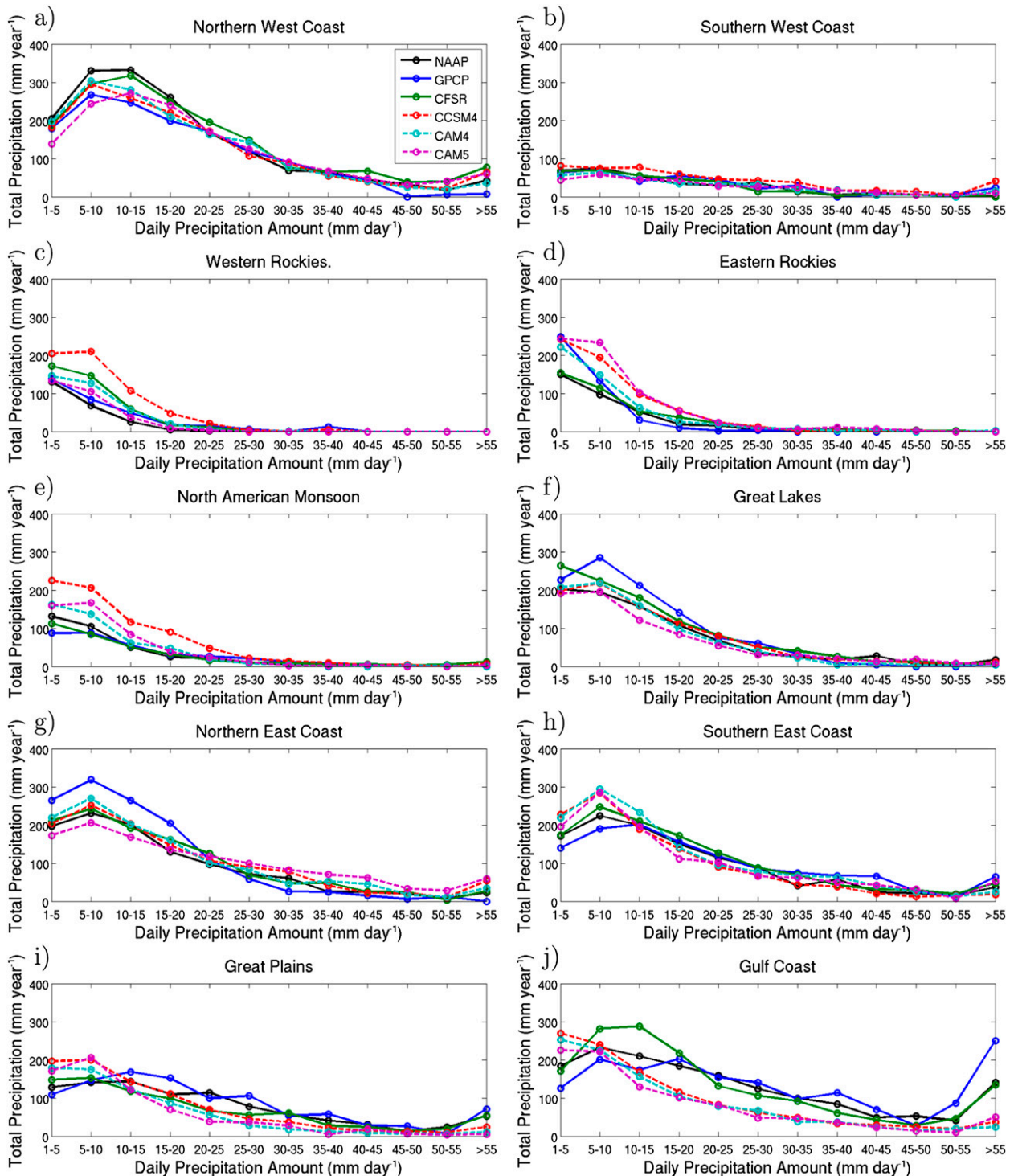


FIG. 6. Climatological annual total mass distribution at point locations as indicated in Fig. 4.

to the NAAP (Figs. 5j, 6j). For the Gulf Coast CDFs, the GPCP has errors in the opposite direction (Fig. 5j), which demonstrate that, if the GPCP were used solely to validate the GCMs, we would assume even larger errors

in the GCMs, while the opposite would be true if the CFSR were used. These two locations highlight the importance of using several data sources for validation and how the errors associated with reference data sources

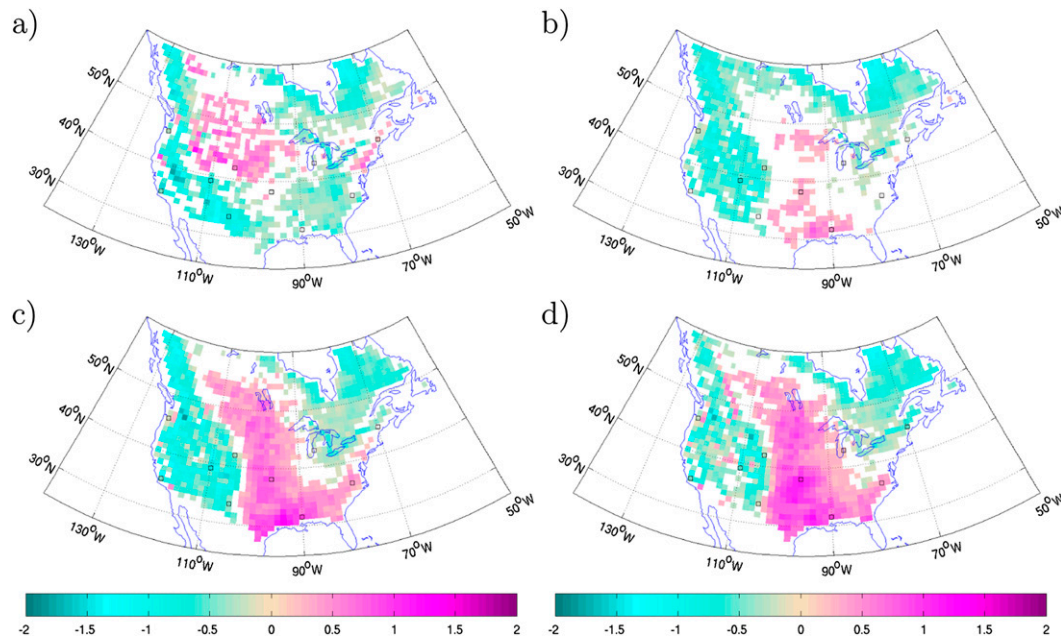


FIG. 7. Anomalies relative to the NAAP (e.g., GPCP–NAAP) of the skewness of the climatological annual total mass distribution, at each grid point, for the (a) GPCP, (b) CFSR, (c) CCSM4, and (d) CAM5. Locations shown in Fig. 4 are included for reference and shown as open squares.

can have a large influence on the interpretation of model results.

A comparison of the skewness of the TMDs relative to the NAAP reveals spatial coherency in many of the features mentioned above for the specific point locations (Fig. 7). There is a distinct shift in the difference in the skewness of the GPCP compared to the NAAP at 40°N, where it is less skewed south of 40°N and vice versa (Fig. 7a). This coincides with the change in data source from the TMPI (south) to TOVS (north) in the GPCP and is thus likely due to errors in the GPCP. This shift in the distribution was noted previously for the western versus eastern Rockies points and southern versus northern West Coast points. This seems to be a robust feature along the 40°N latitude, except along the West Coast. The CFSR is negatively skewed relative to the NAAP in the west, where there is high orography, and positively skewed in the east (Fig. 7b). The differences in skewness between the GPCP and the CFSR relative to the NAAP are of the same magnitude but the spatial patterns differ.

The bias patterns of skewness in the GCMs are very similar to that of the CFSR but more amplified (Figs. 7c,d). The CFSR, with higher resolution and mass fields based on assimilated data, is likely to have a somewhat better prediction of precipitation than the GCMs; however, it is still model based so the similar error patterns to the GCMs are expected. The skewness in the Great Plains is consistent with the idea that there is too much light and not enough heavy precipitation in the models,

particularly in regions that experience convection. It is interesting to note that the error in the skewness in the Great Plains is higher in the uncoupled models (Fig. 7d) (CAM4 not shown) than in the fully coupled model (Fig. 7c), and the error in the Rockies is lower in the uncoupled models (CAM4 not shown) and higher in the coupled model.

There are many regions in northern Canada where the bias in the skewness with respect to the NAAP is negative for all the other reference and model products (Fig. 7). These regions include northern Quebec, Ontario, Manitoba, and the entire west coast of British Columbia. It is unclear whether this is a result of a common bias in the GPCP, CFSR, CCSM4, and CAM5 or it is due to the sparseness of station data in these regions (Fig. 1) and the resulting potential for large errors in the NAAP data (Fig. 2).

The ability of the GCMs to reproduce observed extremes in different bimonthly periods is shown in the climatological extreme precipitation in each of these models with comparison to the NAAP (Fig. 8). In the NAAP, we see a maximum of extreme values along the West Coast whose seasonality is such that the extremes are greater in winter (JF, MA, and ND) and small or nonexistent in the summer (MJ and JA). The models produce the higher extremes in this region as well as the seasonality; however, during the bimonthly periods where this feature is largest (JF, MA, and ND), the magnitudes are on the order of 10–20 mm day^{−1} larger.

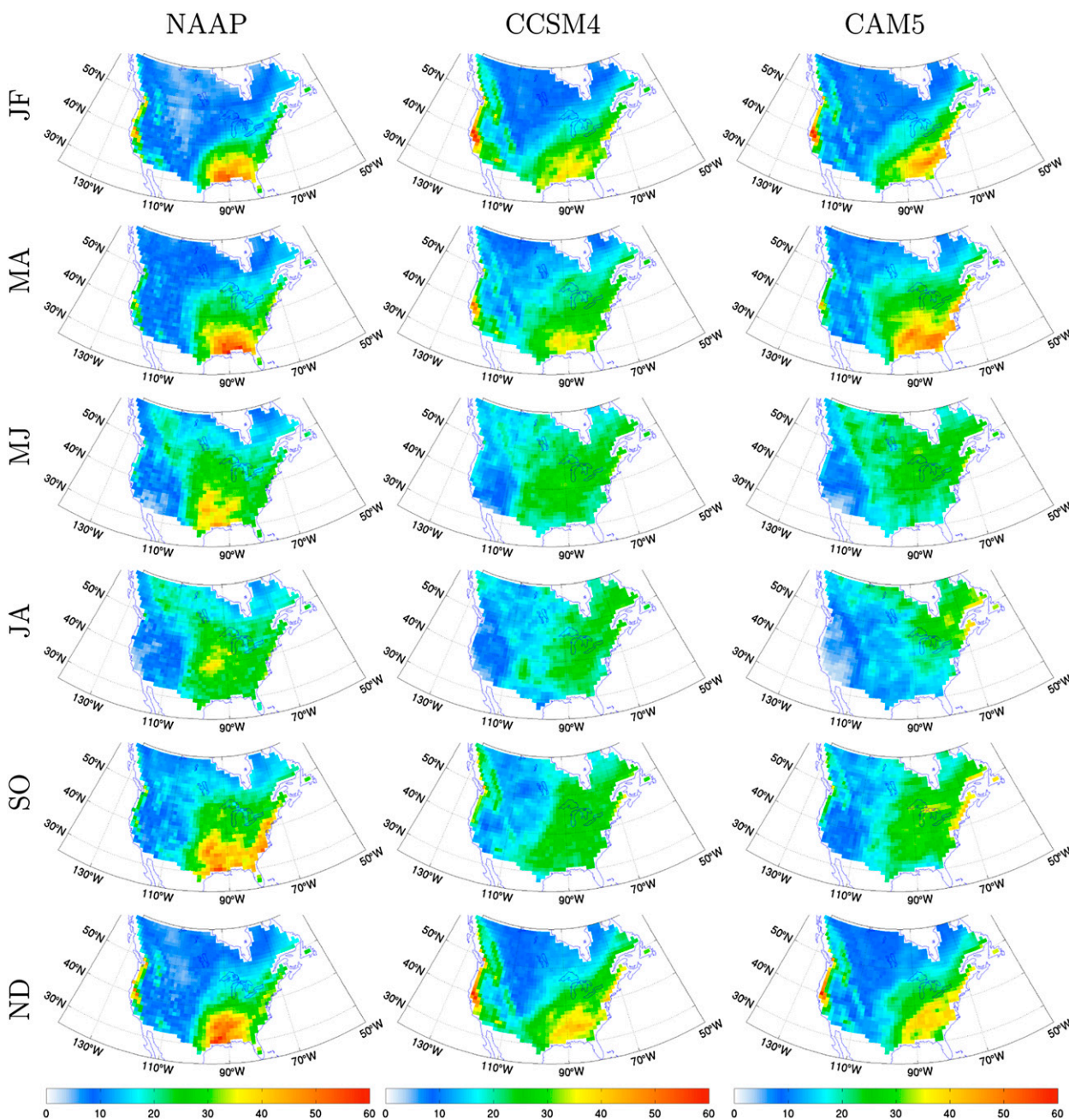


FIG. 8. Climatological (1979–2005) extreme precipitation (mm day⁻¹) for (left) NAAP, (center) CCSM4, and (right) CAM5 in bimonthly periods: January–February, March–April, May–June, July–August, September–October, and November–December.

Within the Rocky Mountains, the NAAP generally has finer scale structures, even though the models are remapped to the same resolution. Studies using an earlier version of the model saw that increasing the resolution resulted in much better representations of orographic precipitation (Gent et al. 2010; Iorio et al. 2004). Therefore, this may be a result of the model resolution being too coarse to fully capture details in the orography

that are important for precipitation. During the JA period, the models have lower values than the NAAP in the North American monsoon region. This may be indicative of issues representing the monsoon processes in the model. In general, however, the largest errors in the climatological extreme precipitation are located east of the Rockies, where heavier convective precipitation is frequent.

Focusing on errors in the eastern United States, one of the most notable features that the models are not able to produce is the heavy precipitation in the plains during the MJ and JA bimonthly periods (Fig. 8). During these bimonthly periods, the Great Plains experience heavy convection and propagating mesoscale convective systems (MCS), which contribute significantly to the total rainfall in these seasons (Fritsch et al. 1986). DeMott et al. (2007) discussed issues in creating heavy convection in the CAM3 model, particularly in the Great Plains. They found that the diurnal cycle in the CAM3 model was too strong and the timing of convection was too early. They used subdaily observations and model output to identify the mechanisms involved and found that the necessity for the model to consume CAPE within a 1-h-long convective adjustment time scale, as well as the condition that plumes must detrain at the level of minimum moist static energy, resulted in convective plumes that were too high and included too little entrainment. As a result, the convection occurred too quickly (DeMott et al. 2007). This then caused the model to be unable to build up moisture and larger CAPE values that would be necessary for heavy rainfall at a later time in the day (DeMott et al. 2007). The inclusion of entrainment in the calculation of CAPE in the newer model versions may have helped but not solved these issues with the parameterization scheme, resulting in an improved representation but still a continued bias in the model's heavy precipitation during the MJ and JA seasons.

To investigate whether this issue with convective parameterization is perpetuated in the newer model version, the phase and magnitude of the diurnal cycle are examined using 3-hourly precipitation from the CAM3, CAM4, and CAM5 models in comparison to 3-hourly satellite precipitation estimates from the Tropical Rainfall Measuring Mission version 7 (TRMM) from the National Aeronautics and Space Administration (Huffman et al. 2007). Results for the CAM5 model are also shown at a higher resolution, $0.23^\circ \times 0.31^\circ$ latitude–longitude, which is referred to as 0.25° . We expect representativeness errors will only impact the magnitude of the diurnal maximum but not the timing of maximum convection, so the 0.25° data are not remapped to the 1° resolution. Results are shown for the time period of 2001–10, except for the CAM5 0.25° model where a common time period was not available. The CAM5 0.25° model analysis is shown for 1996–2005, a time period of the same length and overlapping as many years as possible with the other products for 1996–2005. We found little change to the phase and amplitude when using alternate sets of earlier 10-yr analysis periods. The phase is computed as the time of

the peak in the first harmonic of the diurnal variation, and the magnitude is the mean precipitation over all days. Comparing the diurnal timing from CAM3 to CAM4 run on the same dynamical core as the 1° resolution shows a reduction in the magnitude of the diurnal cycle over the United States but no clear improvements in the timing of the diurnal maximum (Fig. 9). There are some minor improvements in the central United States in the CAM5 (Fig. 9), but the diurnal timing issue persists.

In CAM3, Dirmeyer et al. (2012) showed that increasing the resolution of the GCM did not resolve the diurnal timing problem so long as the precipitation was still parameterized. This is also seen in the CAM5, where increasing the resolution from 1° to 0.25° does not improve the timing issue over most regions, notably in the southeast (Fig. 9). Over the Rocky Mountain region, there is an area of later phase that develops at higher resolution; this error is found in many other orographic regions in the model at high resolution (Bacmeister et al. 2014).

Dirmeyer et al. (2012) did find that the CAM3 with an embedded two-dimensional cloud-resolving model, which is known as the superparameterized CAM3 model (SP-CAM3), did have better timing of their diurnal precipitation maximum (Dirmeyer et al. 2012). In a newer version of the model with embedded cloud-resolving model, SP-CAM3.5, Pritchard et al. (2011) were even able to produce propagating systems in the Great Plains much like MCS found in observations. These results are consistent with the idea that issues with heavy precipitation in the plains during the summer are due to issues with the parameterization of convection. Although the SP-CAM3.5 does have many other issues with the representation of precipitation, such as an overproduction of intense precipitation (Iorio et al. 2004) and some remaining timing issues (Dirmeyer et al. 2012), it is nonetheless illustrative and encouraging to see that a model with explicit convection has some ability to produce MCS.

Heavy precipitation in the Gulf Coast during the JF, MA, and ND bimonthly periods is simulated in both the CAM5 and the CCSM4, but the magnitude is much too low and often displaced farther toward the East Coast (Fig. 8). There are many different precipitation mechanisms responsible for extreme precipitation at this time of year and so it is difficult to speculate as to the source of the error.

Another glaring issue in the model representation of extremes occurs in the SO season in the Gulf Coast and along the U.S. East Coast (Fig. 8). Tropical cyclones are responsible for up to a quarter of September precipitation in these regions (Knight and Davis 2007). It is

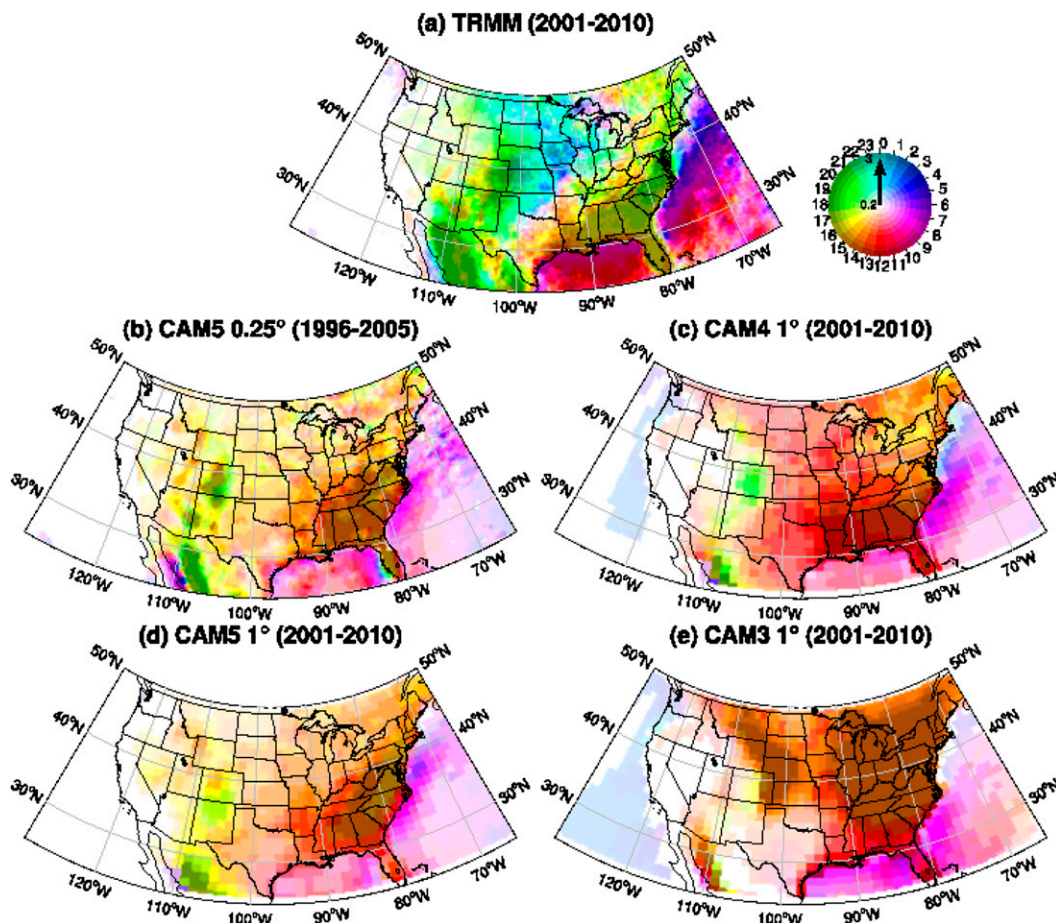


FIG. 9. Local time of first harmonic of JA diurnal precipitation (color hue) in the (a) TRMM, (b) CAM5 at 0.25°, (c) CAM4 at 1°, (d) CAM5 at 1°, and (e) CAM3 at 1°. Color intensity represents the mean daily precipitation (mm day^{-1}) beginning at 0.2 mm day^{-1} . The observation period is 2001–10, except the CAM5 at 0.25°, where it is 1996–2005.

well established that GCMs at standard resolutions have severe issues with the production of tropical cyclones and, when they are produced, they are generally much weaker than in observations. It is possible therefore that these errors are associated with errors in the production of tropical cyclones in the model.

To put the errors in extreme precipitation in perspective, we compare climatological annual extreme precipitation in the NAAP, GPCP, CFSR, CAM4 (not shown), CAM5, and CCSM4 (Fig. 10). We see that errors in the CFSR are generally the smallest: for the absolute value of the percent difference, they are within 0%–20% in central and eastern Canada/United States, much greater (30%–60%) in the U.S. Rocky Mountains, and in some areas in the Canadian Rockies and around Hudson Bay up to 80% (Figs. 10c,d). The errors in the GPCP are typically higher (0%–40%) in most of the United States (Figs. 10a,b). The GPCP has a distinct change in the errors at 40°N, where the input data for the

GPCP change from the TMP1 (south) to TOVS (north). South of 40°N, the GPCP, particularly in the southwest, there are larger extremes with errors ranging from 10% up to 100%, while north of 40°N the extremes are generally smaller (up to 40%), except along the West Coast, which approach 60%. The creators of the GPCP have noted large errors when the GPCP is used at the daily frequency provided and have suggested that the errors decrease when it is temporally averaged (Huffman et al. 2001).

The GCMs have some large errors in the Rocky Mountains and West Coast (60%–100%), similar in pattern to the CFSR (Fig. 10). A central difference between the CFSR and the models is in the southeastern and central United States, where there are errors between 20% and 60%, which as discussed previously are likely due to issues with the representation of convection. Errors in extreme precipitation in the CAM4 (not shown) are very similar to the CAM5 and CCSM4.

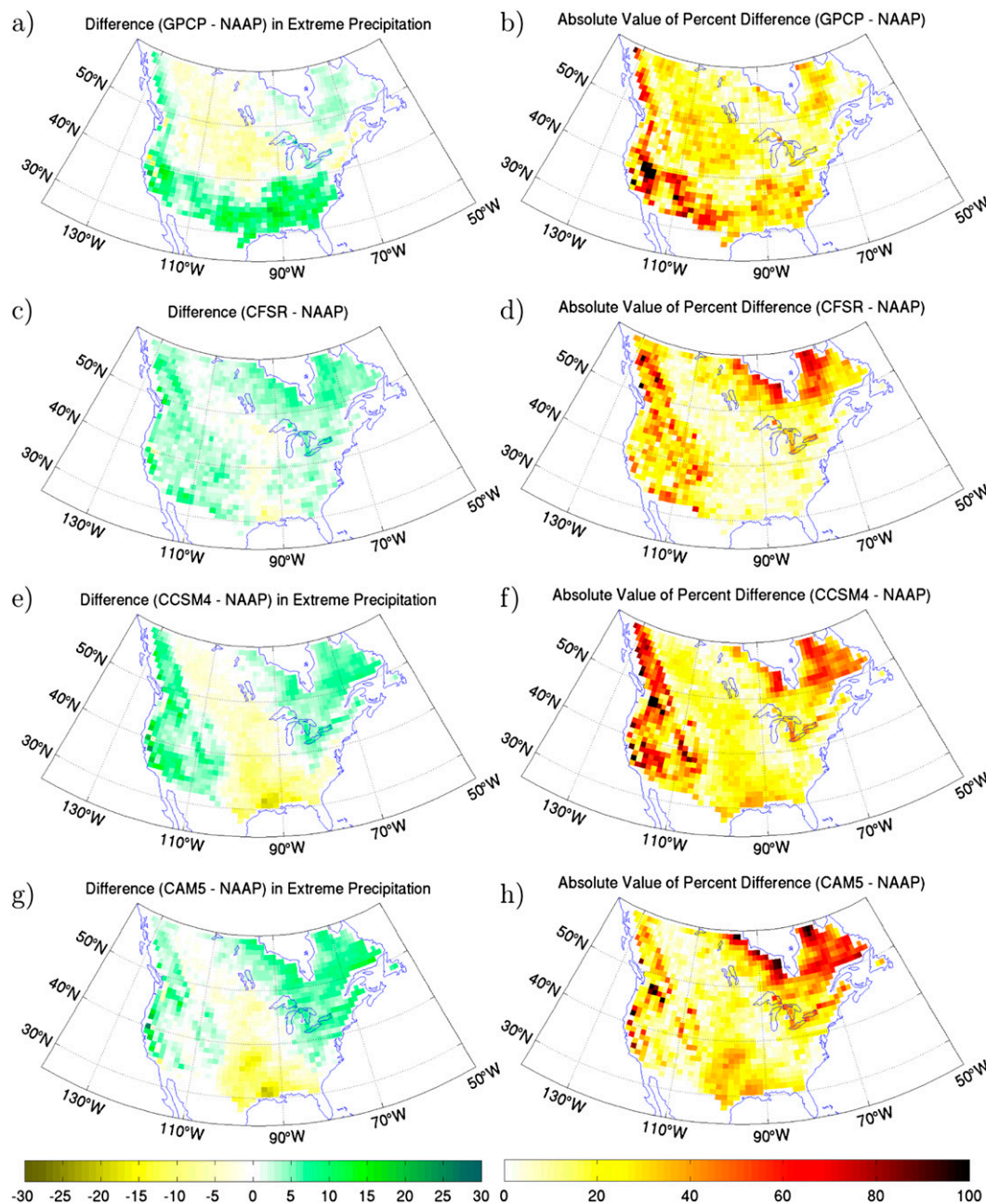


FIG. 10. (left) Anomalies and (right) absolute value of percent anomalies, relative to the NAAP, for the (a),(b) GPCP; (c),(d) CFSR; (e),(f) CCSM4; and (g),(h) CAM5 climatological (1979–2005) annual extreme precipitation (mm day^{-1}). Anomalies are computed, for example, as $(\text{GPCP} - \text{NAAP})$, and absolute value of percent anomalies are computed, for example, as $[(\text{GPCP} - \text{NAAP})/\text{NAAP}] \times 100$.

Although the locations of the errors are different, the magnitude of errors in extreme precipitation in the GPCP and the models are on the same order, which implies that the models are within observational error (Fig. 10). However, given the known errors in the GPCP at daily resolution and the highly visible differences in the error field at the location of the data

transition in the GPCP previously discussed, these results imply that caution should be exerted when using the GPCP to evaluate extremes. It should be noted that errors in the CFSR typically have a similar pattern to those in the GCMs; as such, using the CFSR for validation could result in an underestimation of errors in GCMs.

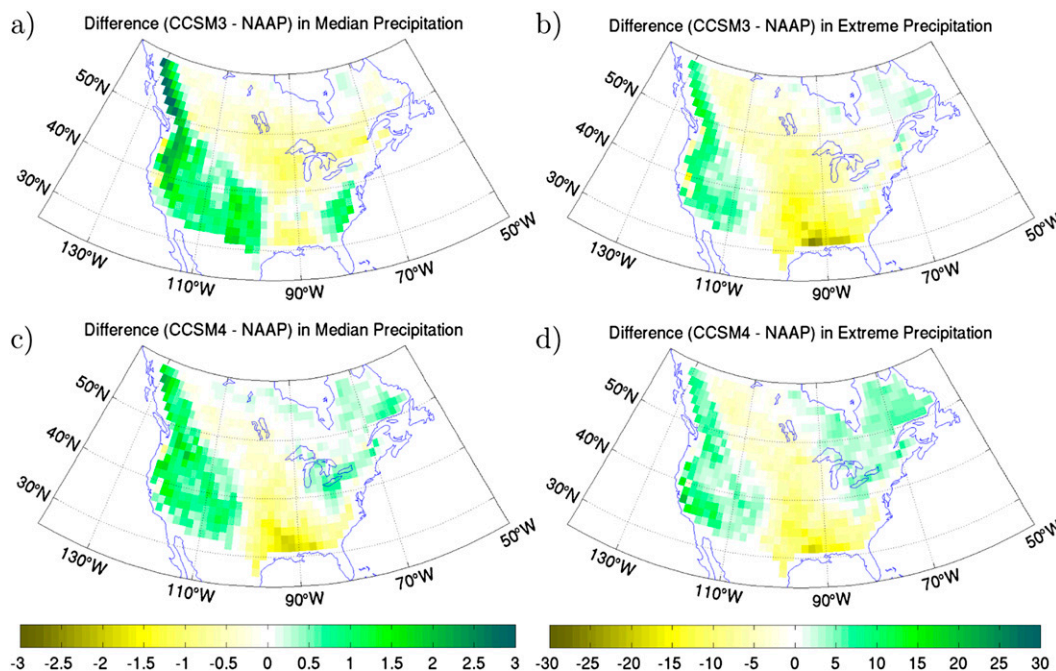


FIG. 11. Anomalies relative to the NAAP, for the (a),(b) CCSM3 and (c),(d) CCSM4 of climatological (1979–99) annual (a),(c) median and (b),(d) extreme precipitation (mm day^{-1}). All data are remapped onto the CCSM3 resolution and anomalies are computed, for example, as CCSM3 minus NAAP.

In addition to examining errors that exist in the current version of the model, it is important to determine what improvements if any there are over the previous version of the model. Gent et al. (2011) show that the distribution of precipitation is much better represented in the CCSM4 than the CCSM3 in the tropics, even when the CCSM4 is run at lower resolution; however, it has not yet been shown how the distributions compare in the midlatitudes. Using the SCRIP remapping procedure once again, we remap both the NAAP and the CCSM4 data onto the lower-resolution CCSM3 T85 spectral grid (approximately 1.4° latitude–longitude) over the 1979–99 period and examine similar metrics as above. Comparison between the CCSM3 and CCSM4 median precipitation errors with respect to the NAAP reveal that there is a notable reduction in excess precipitation over the northern West Coast but over much of the rest of the domain there is little to no improvement or in some cases a small increase in error in the representation of the median precipitation in the CCSM4 (Fig. 11). For the extreme precipitation, we do see modest improvements in the southeastern United States as well as the West Coast (Fig. 11).

5. Conclusions

Although it is well known that precipitation observations and reanalysis can have some large errors, the

traditional assumption has been that these errors are small relative to those in the precipitation produced by a GCM. As the representation of precipitation in GCMs improves, however, we find that errors in model precipitation are approaching these observational errors and so they must be carefully considered. Results from Gervais et al. (2014) related station density to errors in gridded station measurements and suggest the potential for large errors in gridded precipitation station analysis when station density is low. Applying this relationship between errors and station density found in Gervais et al. (2014) to the NAAP dataset, we find the potential for large errors exists in the mountainous regions of the western United States, as well as in the majority of Canada. A general bias in the models toward negatively skewed precipitation in northern Canada with respect to the NAAP was interpreted as an error in the NAAP since intercomparison with the other reference sources all had the same bias. We also see large changes in the GPCP at the border of its data change. Changes in the skewness and extreme precipitation across this border can be of the same magnitude as errors in the models. This implies that the GPCP for some metrics, the extremes in particular, can be as biased as the models and thus should be used with caution. The errors in the CFSR generally have the same pattern as the GCMs and so when validating against this data you are likely to see fewer errors than actually exist. These results support

the viewpoint that in many cases the model error is approaching observational error and thus utilizing several reference products to constrain these model errors is important.

CAM4, CAM5, and CCSM4 all had very similar biases in their representation of the distribution of precipitation. In general, the well-known tendency of positive skewness in climate models is present in these models but only east of the Rocky Mountains. This coincides with regions that experience a larger portion of convective precipitation. Regions within and to the west of the Rockies generally have a more negatively skewed distribution than the NAAP, indicative of a shift toward higher daily precipitation rates than that found in the observations. Errors in skewness are higher in the east than the west. Several locations have distributions that are very close to or fall between the various observations and reanalysis, which is very promising.

Examining the extreme precipitation for different bi-monthly periods implicates different processes as sources of bias. In particular, there is a large underestimation of extreme precipitation in the Gulf Coast region in the winter and spring months. Heavy precipitation is also underestimated in the spring and summer from the Gulf Coast through the Great Plains. These types of errors were seen in previous versions of the model and were attributed in those cases to issues with the convective parameterization and an unrealistic representation of the buildup of CAPE. An analysis of the phase of diurnal precipitation for various versions of the CAM model reveal that these issues with diurnal timing are ongoing in the newer versions of the model and are not remedied by increasing the resolution from 1° to 0.25° . A final notable issue with model's extreme precipitation is the lack of heavy precipitation in the SO period from the Gulf Coast inland and up the eastern seaboard. Although this is likely not the only source of errors in this region, this error is attributed in part to a lack of extreme precipitation from tropical systems as they are not well resolved in the model.

When errors in CCSM3 and CCSM4 were compared by Gent et al. (2011) in the tropical regions they found large improvements in the representation of the distribution of precipitation. However, when comparisons are made in North America using the metrics in this study, we find the gains in the representation of the CCSM4 to be minimal. There are many potential reasons of which one contributing factor may be the continued issue with convective parameterization in the model.

When examining output of future climate model runs, we can use these results to inform us of which processes we can and cannot expect the CCSM model to adequately produce. This study focused on the United

States and southern Canada; however, this area contains a wide range of precipitation climatologies that may be useful in understanding GCM prediction of precipitation in other midlatitude regions. Although the CCSM models are doing relatively well in producing an adequate distribution in many locations, they do have some difficulty in producing extremes with higher magnitudes. Our confidence in their abilities to represent future changes in these extremes is therefore low, but for the production of larger-scale precipitation processes it is higher.

Acknowledgments. Support for this work is provided in part by the Fonds de Recherche en Science du Climat (FRSCO) from the Ouranos Consortium and the Natural Sciences and Engineering Council (NSERC) Discovery Foundation. M.G. is grateful to Québec-Océan for financial support during the course of this work. B.T. is supported by the Canadian Sea Ice and Snow Evolution (CanSISE) Network, which is funded by the NSERC Climate Change and Atmospheric Research program. We thank the NOAA/Climate Prediction Center (CPC) for making available the UPD dataset; the National Land and Water Information Service, Agriculture and Agri-Food, Canada, for the GCDC dataset, with special thanks to Dan McKenney and Pia Papadopol; the NASA Goddard Space Flight Center's Laboratory for the development of the GPCP dataset as part of the GEWEX Global Precipitation Climatology Project; the NCAR Computational Information Systems Laboratory for access to the CFSR data; and NCAR for access to CCSM data through the Earth System grid, as well as the CESM polar working group, particularly Dr. David Bailey and Dr. Marika Holland. We thank Dr. Mike Pritchard, Dr. Mitch Moncrieff, Dr. Camille Li, and Dr. Graeme Stephens for useful discussions. The NAAP dataset is available from the corresponding author upon permission of the data providers of the UPD and the GCDC datasets.

REFERENCES

- Adam, J. C., and D. P. Lettenmaier, 2003: Adjustment of global gridded precipitation for systematic bias. *J. Geophys. Res.*, **108**, 4257, doi:10.1029/2002JD002499.
- Anderson, T. W., 1962: On the distribution of the two-sample Cramer-von Mises criterion. *Ann. Math. Stat.*, **33**, 1148–1159, doi:10.1214/aoms/1177704477.
- , and D. A. Darling, 1952: Asymptotic theory of certain “goodness of fit” criteria based on stochastic processes. *Ann. Math. Stat.*, **23**, 193–212, doi:10.1214/aoms/1177729437.
- Bacmeister, J. T., M. F. Wehner, R. B. Neale, A. Gettelman, C. Hannay, P. H. Lauritzen, J. M. Caron, and J. E. Truesdale, 2014: Exploratory high-resolution climate simulations using the Community Atmosphere Model (CAM). *J. Climate*, **27**, 3073–3099, doi:10.1175/JCLI-D-13-00387.1.

- Chen, C.-T., and T. Knutson, 2008: On the verification and comparison of extreme rainfall indices from climate models. *J. Climate*, **21**, 1605–1621, doi:[10.1175/2007JCLI1494.1](https://doi.org/10.1175/2007JCLI1494.1).
- Chen, M., R. Dickinson, Z. Xubin, and A. Hahmann, 1996: Comparison of precipitation observed over the continental United States to that simulated by a climate model. *J. Climate*, **9**, 2233–2249, doi:[10.1175/1520-0442\(1996\)009<2233:COPOOT>2.0.CO;2](https://doi.org/10.1175/1520-0442(1996)009<2233:COPOOT>2.0.CO;2).
- Cherry, J. E., L.-B. Tremblay, M. Stieglitz, G. Gong, and S. J. Déry, 2007: Development of the Pan-Arctic snowfall reconstruction: New land-based solid precipitation estimates for 1940–99. *J. Hydrometeorol.*, **8**, 1243–1263, doi:[10.1175/2007JHM765.1](https://doi.org/10.1175/2007JHM765.1).
- Collins, W., and Coauthors, 2006: The Community Climate System Model version 3 (CCSM3). *J. Climate*, **19**, 2122–2143, doi:[10.1175/JCLI3761.1](https://doi.org/10.1175/JCLI3761.1).
- Dai, A., and K. E. Trenberth, 2004: The diurnal cycle and its depiction in the Community Climate System Model. *J. Climate*, **17**, 930–951, doi:[10.1175/1520-0442\(2004\)017<0930:TDCAID>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<0930:TDCAID>2.0.CO;2).
- , F. Giorgi, and K. E. Trenberth, 1999: Observed and model-simulated diurnal cycles of precipitation over the contiguous United States. *J. Geophys. Res.*, **104**, 6377–6402, doi:[10.1029/98JD02720](https://doi.org/10.1029/98JD02720).
- Daly, C., W. P. Gibson, G. H. Taylor, G. L. Johnson, and P. Pasteris, 2002: A knowledge-based approach to the statistical mapping of climate. *Climate Res.*, **22**, 99–113, doi:[10.3354/cr022099](https://doi.org/10.3354/cr022099).
- DeMott, C. A., D. A. Randall, and M. Khairoutdinov, 2007: Convective precipitation variability as a tool for general circulation model analysis. *J. Climate*, **20**, 91–112, doi:[10.1175/JCLI3991.1](https://doi.org/10.1175/JCLI3991.1).
- Dirmeyer, P. A., and Coauthors, 2012: Simulating the diurnal cycle of rainfall in global climate models: Resolution versus parameterization. *Climate Dyn.*, **39** (1–2), 399–418, doi:[10.1007/s00382-011-1127-9](https://doi.org/10.1007/s00382-011-1127-9).
- Emori, S., A. Hasegawa, R. Suzuki, and K. Dairaku, 2005: Validation, parameterization dependence, and future projection of daily precipitation simulated with a high-resolution atmospheric GCM. *Geophys. Res. Lett.*, **32**, L06708, doi:[10.1029/2004GL022306](https://doi.org/10.1029/2004GL022306).
- Fritsch, J. M., R. J. Kane, and C. R. Chelius, 1986: The contribution of mesoscale convective weather system to the warm-season precipitation in the United States. *J. Climate Appl. Meteor.*, **25**, 1333–1345, doi:[10.1175/1520-0450\(1986\)025<1333:TCOMCW>2.0.CO;2](https://doi.org/10.1175/1520-0450(1986)025<1333:TCOMCW>2.0.CO;2).
- Gandin, L., 1965: *Objective Analysis of Meteorological Fields*. Israel Program for Scientific Translations, 242 pp.
- Gent, P. R., S. G. Yeager, R. B. Neale, S. Levis, and D. A. Bailey, 2010: Improvements in a half degree atmosphere/land version of the CCSM. *Climate Dyn.*, **34**, 819–833, doi:[10.1007/s00382-009-0614-8](https://doi.org/10.1007/s00382-009-0614-8).
- , and Coauthors, 2011: The Community Climate System Model version 4. *J. Climate*, **24**, 4973–4991, doi:[10.1175/2011JCLI4083.1](https://doi.org/10.1175/2011JCLI4083.1).
- Gervais, M., B. Tremblay, J. Gyakum, and E. Atallah, 2014: Representing extremes in a daily gridded precipitation analysis over the United States: Impacts of station density, resolution, and gridding methods. *J. Climate*, **27**, 5201–5218, doi:[10.1175/JCLI-D-13-00319.1](https://doi.org/10.1175/JCLI-D-13-00319.1).
- Goodison, B., P. Louis, and D. Yang, 1998: WMO solid precipitation measurement intercomparison. World Meteorological Association Instruments and Observing Systems Rep. 67, 318 pp.
- Gornall, J., R. Betts, E. Burke, R. Clark, J. Camp, K. Willett, and A. Wiltshire, 2010: Implications of climate change for agricultural productivity in the early twenty-first century. *Philos. Trans. Roy. Soc. London*, **365B**, 2973–2989, doi:[10.1098/rstb.2010.0158](https://doi.org/10.1098/rstb.2010.0158).
- Huffman, G. J., and Coauthors, 1997: The Global Precipitation Climatology Project (GPCP) combined precipitation dataset. *Bull. Amer. Meteor. Soc.*, **78**, 5–20, doi:[10.1175/1520-0477\(1997\)078<0005:TGPCPG>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<0005:TGPCPG>2.0.CO;2).
- , R. F. Adler, M. M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind, 2001: Global precipitation at one-degree daily resolution from multisatellite observations. *J. Hydrometeorol.*, **2**, 36–50, doi:[10.1175/1525-7541\(2001\)002<0036:GPAODD>2.0.CO;2](https://doi.org/10.1175/1525-7541(2001)002<0036:GPAODD>2.0.CO;2).
- , and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.*, **8**, 38–55, doi:[10.1175/JHM560.1](https://doi.org/10.1175/JHM560.1).
- Hutchinson, M. F., 1995: Interpolating mean rainfall using thin plate smoothing splines. *Int. J. Geogr. Info. Syst.*, **9**, 385–403, doi:[10.1080/02693799508902045](https://doi.org/10.1080/02693799508902045).
- , D. W. McKenney, K. Lawrence, J. H. Pedlar, R. F. Hopkinson, E. Milewska, and P. Papadopol, 2009: Development and testing of Canada-wide interpolated spatial models of daily minimum-maximum temperature and precipitation for 1961–2003. *J. Appl. Meteor. Climatol.*, **48**, 725–741, doi:[10.1175/2008JAMC1979.1](https://doi.org/10.1175/2008JAMC1979.1).
- Iorio, J., P. Duffy, B. Govindasamy, S. Thompson, M. Khairoutdinov, and D. Randall, 2004: Effects of model resolution and subgrid-scale physics on the simulation of precipitation in the continental United States. *Climate Dyn.*, **23** (3–4), 243–258, doi:[10.1007/s00382-004-0440-y](https://doi.org/10.1007/s00382-004-0440-y).
- Jones, P., 1999: First- and second-order conservative remapping schemes for grids in spherical coordinates. *Mon. Wea. Rev.*, **127**, 2204–2210, doi:[10.1175/1520-0493\(1999\)127<2204:FASOCR>2.0.CO;2](https://doi.org/10.1175/1520-0493(1999)127<2204:FASOCR>2.0.CO;2).
- Knight, D. B., and R. E. Davis, 2007: Climatology of tropical cyclone rainfall in the southeastern United States. *Phys. Geogr.*, **28**, 126–147, doi:[10.2747/0272-3646.28.2.126](https://doi.org/10.2747/0272-3646.28.2.126).
- Kursinski, A. L., and X. Zeng, 2006: Areal estimation of intensity and frequency of summertime precipitation over a midlatitude region. *Geophys. Res. Lett.*, **33**, L22401, doi:[10.1029/2006GL027393](https://doi.org/10.1029/2006GL027393).
- Massey, F., 1951: The Kolmogorov-Smirnov test for goodness of fit. *J. Amer. Stat. Assoc.*, **46**, 68–78, doi:[10.1080/01621459.1951.10500769](https://doi.org/10.1080/01621459.1951.10500769).
- Neale, R. B., C.-C. Chen, P. H. Lauritzen, D. L. Williamson, A. J. Conley, A. K. Smith, M. Mills, and H. Morrison, 2010: Description of the NCAR Community Atmosphere Model (CAM 5.0). National Center for Atmospheric Research Tech. Rep., 283 pp.
- , J. Richter, S. Park, P. H. Lauritzen, S. J. Vavrus, P. J. Rasch, and M. Zhang, 2013: The mean climate of the Community Atmosphere Model (CAM4) in forced SST and fully coupled experiments. *J. Climate*, **26**, 5150–5168, doi:[10.1175/JCLI-D-12-00236.1](https://doi.org/10.1175/JCLI-D-12-00236.1).
- Pritchard, M. S., M. W. Moncrieff, and R. C. J. Somerville, 2011: Orographic propagating precipitation systems over the United States in a global climate model with embedded explicit convection. *J. Atmos. Sci.*, **68**, 1821–1840, doi:[10.1175/2011JAS3699.1](https://doi.org/10.1175/2011JAS3699.1).
- Rosenzweig, C., F. N. Tubiello, R. Goldberg, E. Mills, and J. Bloomfield, 2002: Increased crop damage in the US from excess precipitation under climate change. *Global Environ. Change*, **12**, 197–202, doi:[10.1016/S0959-3780\(02\)00008-0](https://doi.org/10.1016/S0959-3780(02)00008-0).
- Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057, doi:[10.1175/2010BAMS3001.1](https://doi.org/10.1175/2010BAMS3001.1).

- 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*, **26**, 7902–7911, doi:[10.1175/JCLI-D-12-00850.1](https://doi.org/10.1175/JCLI-D-12-00850.1).
- Shepard, D., 1968: A two-dimensional interpolation function for irregularly-spaced data. *Proc. 23rd ACM National Conf.*, New York, NY, Association for Computing Machinery, 517–524, doi:[10.1145/800186.810616](https://doi.org/10.1145/800186.810616).
- Shiu, C.-J., S. C. Liu, C. Fu, A. Dai, and Y. Sun, 2012: How much do precipitation extremes change in a warming climate? *Geophys. Res. Lett.*, **39**, L17707, doi:[10.1029/2012GL052762](https://doi.org/10.1029/2012GL052762).
- Sillmann, J., V. V. Kharin, X. Zhang, F. W. Zwiers, and D. Bronaugh, 2013: Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate. *J. Geophys. Res.*, **118**, 1716–1733, doi:[10.1002/jgrd.50203](https://doi.org/10.1002/jgrd.50203).
- Stephens, M. A., 1970: Use of the Kolmogorov-Smirnov, Cramer-Von Mises and related statistics without extensive tables. *J. Roy. Stat. Soc.*, **32**, 115–122.
- Sun, X., and A. P. Barros, 2010: An evaluation of the statistics of rainfall extremes in rain gauge observations, and satellite-based and reanalysis products using universal multifractals. *J. Hydrometeor.*, **11**, 388–404, doi:[10.1175/2009JHM1142.1](https://doi.org/10.1175/2009JHM1142.1).
- Sun, Y., S. Solomon, A. Dai, and R. Portmann, 2006: How often does it rain? *J. Climate*, **19**, 916–934, doi:[10.1175/JCLI3672.1](https://doi.org/10.1175/JCLI3672.1).
- Susskind, J., P. Piraino, L. Rokke, L. Iredell, and A. Mehta, 1997: Characteristics of the TOVS Pathfinder Path A dataset. *Bull. Amer. Meteor. Soc.*, **78**, 1449–1472, doi:[10.1175/1520-0477\(1997\)078<1449:COTTPP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<1449:COTTPP>2.0.CO;2).
- Xie, P., M. Chen, S. Yang, A. Yatagai, T. Hayasaka, Y. Fukushima, and C. Liu, 2007: A gauge-based analysis of daily precipitation over East Asia. *J. Hydrometeor.*, **8**, 607–626, doi:[10.1175/JHM583.1](https://doi.org/10.1175/JHM583.1).
- Zhang, G. J., and N. A. McFarlane, 1995: Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian Climate Centre general circulation model. *Atmos.–Ocean*, **33**, 407–446, doi:[10.1080/07055900.1995.9649539](https://doi.org/10.1080/07055900.1995.9649539).