Evaluation of CMIP5 Model Precipitation Using PERSIANN-CDR

PHU NGUYEN,[a,b] ANDREA THORSTENSEN,[a] SOROOSH SOROOSHIAN,[a] QIAN ZHU,[a,c] HOANG TRAN,[a] HAMED ASHOURI,[a] CHIYUAN MIAO,[d] KUOLIN HSU,[a] AND XIAOGANG GAO[a]

[a] *Center for Hydrometeorology and Remote Sensing, Department of Civil and Environmental Engineering, University of California, Irvine, Irvine, California*
[b] *Nong Lam University, Ho Chi Minh City, Vietnam*
[c] *Department of Hydraulic Engineering, Civil Engineering College, Zhejiang University, Hangzhou, China*
[d] *State Key Laboratory of Earth Surface Processes and Resource Ecology, College of Global Change and Earth System Science, Beijing Normal University, and Joint Center for Global Change Studies, Beijing, China*

(Manuscript received 25 August 2016, in final form 8 June 2017)

ABSTRACT

The purpose of this study is to use the PERSIANN–Climate Data Record (PERSIANN-CDR) dataset to evaluate the ability of 32 CMIP5 models in capturing the behavior of daily extreme precipitation estimates globally. The daily long-term historical global PERSIANN-CDR allows for a global investigation of eight precipitation indices that is unattainable with other datasets. Quantitative comparisons against CPC daily gauge; GPCP One-Degree Daily (GPCP1DD); and TRMM 3B42, version 7 (3B42V7), datasets show the credibility of PERSIANN-CDR to be used as the reference data for global evaluation of CMIP5 models. This work uniquely defines different study regions by partitioning global land areas into 25 groups based on continent and climate zone type. Results show that model performance in warm temperate and equatorial regions in capturing daily extreme precipitation behavior is largely mixed in terms of index RMSE and correlation, suggesting that these regions may benefit from weighted model averaging schemes or model selection as opposed to simple model averaging. The three driest climate regions (snow, polar, and arid) exhibit high correlations and low RMSE values when compared against PERSIANN-CDR estimates, with the exceptions of the cold regions showing an inability to capture the 95th and 99th percentile annual total precipitation characteristics. A comprehensive assessment of each model's performance in each continent–climate zone defined group is provided as a guide for both model developers to target regions and processes that are not yet fully captured in certain climate types, and for climate model output users to be able to select the models and/or the study areas that may best fit their applications of interest.

1. Introduction

Precipitation is one of the key elements for hydrological and climate studies. Atmosphere–ocean general circulation models (GCMs) are the most important tools for simulating the current state of the climate and projecting future changes for precipitation under different scenarios of greenhouse gas emissions (Jiang et al. 2012). With climate models being developed by different groups around the world, with different built-in assumptions, it comes as no surprise when there is a lack of correspondence between model projections of precipitation. Gaetani and Mohino (2013) assessed the capability of different models from phase 5 of the Coupled Model Intercomparison Project in predicting precipitation and found that predictive skills of CMIP5 precipitation simulations are highly model dependent. Therefore, it has been difficult to quantify what precipitation changes should be expected, given that disagreement among climate model projections is widespread (Neelin et al. 2006). A common remedy for such disagreement in model-produced precipitation is to simply employ the mean model, which considers all models equal. While this method has shown merit, particularly for global-scale studies evaluating the mean climate (Reichler and Kim 2008), the question remains as to whether this type of strategy can offer any insight into further understanding of the climate system, particularly in the scope of enhanced development of the climate models themselves to represent it. By identifying the individual climate models that perform well for a given region, one can begin to examine what aspects/assumptions those superior models have that less superior models do not. This notion loosely supports the development of a

Corresponding author: Phu Nguyen, ndphu@uci.edu

DOI: 10.1175/JHM-D-16-0201.1

model hierarchy for climate modeling suggested by Held (2005) by providing a basis for evaluating how much complexity may be needed to capture processes unique to various climate regions. Based on results shown here, a weighted mean might be a better approach for certain precipitation extremes. The weights can be based on the performance by the models for a particular index. Models that perform better get more weight than models that perform poorly.

With the release of the CMIP5 results (Taylor et al. 2012), researchers have conducted many studies. To evaluate the CMIP5 models' precipitation simulations, the reference data should be reliable. Rain gauge and radar observations are traditionally used as the reference datasets for evaluation. However, they are often sparse and inadequate to capture the spatial and temporal variability of precipitation systems (Miao et al. 2015). The current era of satellite-based precipitation estimates has made a great difference in detecting rainfall distribution and has been complementary to conventional rain and radar measurements. There are various high-resolution satellite precipitation products available, including the Tropical Rainfall Measuring Mission (TRMM) 3B42, version 7 (3B42V7; Huffman et al. 2007), and the Climate Prediction Center (CPC) morphing technique (CMORPH; Joyce et al. 2004). However, at least 30 years of historical weather data are required for long-term climatological studies (Burroughs 2003). It is essential to explore satellite-based precipitation products that can provide full global coverage at relatively high temporal and spatial resolutions suitable for daily precipitation extremes and long-term climate studies (e.g., Bosilovich et al. 2008; Peterson et al. 2012; Rossow et al. 2013; Wuebbles et al. 2014; Ashouri et al. 2016b). For example, the Global Precipitation Climatology Project (GPCP) data (1979–present; Adler et al. 2003), which merge the highest-quality satellite and gauge estimates, have been widely used as the reference data to evaluate the CMIP5 models for precipitation (Liu et al. 2014; Mehran et al. 2014).

Mehran et al. (2014) utilized the GPCP as the reference dataset to cross validate CMIP5 historical simulations of precipitation over land. However, the relatively coarse spatial and temporal resolution of GPCP precipitation datasets (2.5° latitude × 2.5° longitude; monthly) limits its application for higher-resolution studies, especially at a scale relevant to extreme events. Sillmann et al. (2013) compared global extreme precipitation indices from CMIP5 models with those from the Hadley Centre Global Climate Extremes Index 2 (HadEX2) observational precipitation dataset. They concluded that CMIP5 models were generally able to simulate precipitation extremes. Sillmann et al. (2013) also pointed out that the spatial coverage of the HadEX2 dataset (2.5° latitude × 3.75° longitude) is "far less than ideal." The results from this work were included in the IPCC Fifth Assessment Report (Flato et al. 2013; IPCC 2013).

Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks–Climate Data Record (PERSIANN-CDR) is a new retrospective satellite-based precipitation dataset intended for various applications requiring long-term, near-global coverage and high-resolution data including climate model evaluations and studies (Ashouri et al. 2015). PERSIANN-CDR provides the opportunity to evaluate the behavior of daily extreme precipitation patterns of CMIP5 models on a global scale and at a higher resolution over the past three decades. In this research, we aim to evaluate the capability of CMIP5 models in capturing the behavior of daily extreme precipitation estimates based on the daily PERSIANN-CDR dataset. Different from the other global evaluation works about CMIP5 models for precipitation, this study examines the performance of CMIP5 models over different continent–climate zone (CCZ) groups. This unique dataset not only enables users of climate model output to leverage the strengths of the top-performing models, but also allows climate model developers to get a more complete picture of how each model performs in a specific environment.

The remainder of this paper is organized as follows: the PERSIANN-CDR and CMIP5 models are briefly introduced in section 2; the methodology and results are presented in sections 3 and 4, respectively; and the conclusions and discussions are documented in section 5.

2. Datasets

PERSIANN-CDR is a new retrospective satellite-based precipitation dataset for long-term global climate studies (Ashouri et al. 2015). This dataset was developed under NOAA's Climate Data Record (CDR) initiative. PERSIANN-CDR provides near-global (60°S–60°N) daily precipitation estimates at 0.25° spatial resolution from 1 January 1983 to the present (http://chrsdata.eng.uci.edu). PERSIANN-CDR uses the existing PERSIANN algorithm (Hsu et al. 1997; Sorooshian et al. 2000) as its backbone model. In PERSIANN-CDR, in order to eliminate the need for passive microwave (PMW) observations, the nonlinear regression parameters of the neural network (NN) model are trained, using the National Centers for Environmental Prediction (NCEP) stage IV radar data, and kept fixed for the retrospective estimation of rainfall rates. The model then uses the archive of the

Gridded Satellite brightness temperature observations (GridSat-B1; Knapp 2008) to estimate surface rainfall rate at 0.25° × 0.25° pixels. To reduce the biases in the PERSIANN-estimated precipitation while preserving the spatial and temporal patterns in high resolution, 2.5° monthly GPCP precipitation data are utilized.

PERSIANN-CDR, even though a newly released dataset, is being rapidly adopted for use in different studies (e.g., Guo et al. 2015; Luchetti et al. 2016; Nguyen et al. 2017; Katiraie-Boroujerdy et al. 2017). Miao et al. (2015) evaluated PERSIANN-CDR's capability in reproducing the behavior of daily extreme precipitation events over China. When compared with the gauge-based East Asia (EA; Xie et al. 2007) dataset, PERSIANN-CDR showed similar spatial and temporal patterns of daily precipitation extremes, particularly in areas where the intensity and frequency of extreme precipitation were very high. Overall, PERSIANN-CDR showed slight underestimation of the magnitude of the daily extreme rainfall. Ashouri et al. (2016a) evaluated the efficacy of PERSIANN-CDR's precipitation and simulated streamflow over three test basins in the Distributed Hydrologic Model Intercomparison Project, phase 2 (DMIP2). The results showed close performance between PERSIANN-CDR and TRMM Multisatellite Precipitation Analysis (TMPA) when compared to stage IV radar data. Similar performances were observed when comparing the PERSIANN-CDR–derived and TMPA-derived streamflow with USGS streamflow observations. PERSIANN-CDR was then used to simulate historical streamflow starting from 1 January 1983 where only PERSIANN-CDR daily data are available. The results showed relatively high correlation coefficients (~0.67–0.73), low biases (~5%–12%), and high index of agreement criterion (~0.68–0.83) between PERSIANN-CDR–simulated daily streamflow and USGS daily observations during 1983–2012. Casse and Gosset (2015) used PERSIANN-CDR satellite rainfall estimates to analyze hydrological changes and flood increase in a watershed in Niger for the period of 1983–2013. The result of this study showed that PERSIANN-CDR depicts comparable annual rainfall amount similar to those of gauge-adjusted satellite rainfall estimates [TRMM 3B42V7; African Rainfall Estimation Algorithm, version 2 (RFE2)] and gauge product (CPC). Casse and Gosset (2015) also stated that "PERSIANN-CDR based hydrological simulation presents a realistic inter-annual variability, and detects flooded years, but not the exact flooded period day by day." In another study by Yang et al. (2016), PERSIANN-CDR and three other widely utilized satellite precipitation products were evaluated against rain gauge observations over a river basin in central-eastern China. TRMM 3B42V7 and PERSIANN-CDR showed reasonable performance, and the monthly gauge adjustment applied to these two products significantly reduced their systematic biases. Tan et al. (2015) used rain gauge data to evaluate the performance of different high-resolution precipitation data and ground-based precipitation products over Malaysia. The results show slight overestimation of observed precipitation by TRMM 3B42, real time (3B42RT); TRMM 3B42V7; and PERSIANN-CDR by 2%, 4.7%, and 2.1%, respectively. On the contrary, the study reports underestimation of observed precipitation by APHRODITE and CMORPH by 19.7% and 13.2%, respectively. Solmon et al. (2015) used PERSIANN-CDR for 2000–09 precipitation trend calculation over southern India. Given these studies, while there is still room for further evaluation of PERSIANN-CDR for different applications over different parts of the world, PERSIANN-CDR has shown its usefulness and acceptable accuracy for being used in different studies, including model testing. Therefore, in this study, the daily PERSIANN-CDR estimates are used as the reference data for evaluation of climate models' simulations during the overlapping period between PERSIANN-CDR and historical simulations of CMIP5 models (1983–2005). Though extreme precipitation may be expected to change in response to forced climate change (because of, for example, thermodynamical Clausius–Clapeyron scaling, dynamical changes, and microphysics changes), the year-to-year variability due to weather (noise in the CMIP5 historical experiment) is expected to be bigger on short time scales (i.e., 23 years in this study). For the United States, this was shown to be true in an ensemble of model experiments and based on different observational data estimates (van der Wiel et al. 2016).

PERSIANN-CDR 0.25° precipitation data are aggregated into 1° resolution using the bilinear interpolation method. To preserve the global rain total and minimize the effect from no-data pixels, the original PERSIANN-CDR images are regridded into 10 times smaller resolution (0.025°) before the bilinear interpolation process.

Regarding the regional classification method, it is worth noting that Köppen–Geiger classifications were used to define boundaries of study regions in this study (as further described in the following section), and these classifications are based on temperature and precipitation observations rather than geographical location. Therefore, although there are regions classified as "polar," they are not physically located at the poles where PERSIANN-CDR does not provide estimates.

TABLE 1. CMIP5 models used for analysis in this study. Horizontal resolution is given in longitude a (360°) and latitude b (180°).

No.	Model	Institute	Resolution ($a \times b$)
1	BNU-ESM	College of Global Change and Earth System Science, Beijing Normal University	128 × 64
2	CCSM4	National Center for Atmospheric Research	288 × 192
3	CESM1(BGC)	National Science Foundation, Department of Energy, National Center for Atmospheric Research	288 × 192
4	CESM1(CAM5)	National Science Foundation, Department of Energy, National Center for Atmospheric Research	288 × 192
5	CMCC-CESM	Centro Euro-Mediterraneo per i Cambiamenti Climatici	96 × 48
6	CMCC-CM	Centro Euro-Mediterraneo per i Cambiamenti Climatici	480 × 240
7	CMCC-CMS	Centro Euro-Mediterraneo per i Cambiamenti Climatici	192 × 96
8	CNRM-CM5	Centre National de Recherches Météorologiques/Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique	256 × 128
9	CSIRO Mk3.6.0	Commonwealth Scientific and Industrial Research Organisation in collaboration with the Queensland Climate Change Centre of Excellence	192 × 96
10	CanESM2	Canadian Centre for Climate Modelling and Analysis	128 × 64
11	EC-EARTH	EC-Earth Consortium	320 × 160
12	FGOALS-g2	State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences; and Center for Earth System Science, Tsinghua University	128 × 108
13	GFDL CM3	Geophysical Fluid Dynamics Laboratory	144 × 90
14	GFDL-ESM2G	Geophysical Fluid Dynamics Laboratory	144 × 90
15	GFDL-ESM2M	Geophysical Fluid Dynamics Laboratory	144 × 90
16	GISS-E2-H	NASA Goddard Institute for Space Studies	144 × 90
17	GISS-E2-R	NASA Goddard Institute for Space Studies	144 × 90
18	HadCM3	Met Office Hadley Centre	96 × 73
19	HadGEM2-ES	Met Office Hadley Centre	192 × 145
20	IPSL-CM5A-LR	L'Institut Pierre-Simon Laplace	96 × 96
21	IPSL-CM5A-MR	L'Institut Pierre-Simon Laplace	144 × 143
22	MIROC-ESM-CHEM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (University of Tokyo), and National Institute for Environmental Studies	128 × 64
23	MIROC-ESM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (University of Tokyo), and National Institute for Environmental Studies	128 × 64
24	MIROC4h	Atmosphere and Ocean Research Institute (University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	640 × 320
25	MIROC5	Atmosphere and Ocean Research Institute (University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	256 × 128
26	MPI-ESM-LR	Max Planck Institute for Meteorology	192 × 96
27	MPI-ESM-MR	Max Planck Institute for Meteorology	192 × 96
28	MPI-ESM-P	Max Planck Institute for Meteorology	192 × 96
29	MRI-CGCM3	Meteorological Research Institute	320 × 160
30	NorESM1-M	Norwegian Climate Centre	144 × 96
31	BCC_CSM1.1(m)	Beijing Climate Center, China Meteorological Administration	320 × 160
32	BCC_CSM1.1	Beijing Climate Center, China Meteorological Administration	128 × 64

Daily precipitation simulations by the 32 CMIP5 models, listed in Table 1, for the historical period 1983–2005 are processed for model evaluation. The precipitation simulations of the models in Table 1 are regridded into 1° × 1° global grids using the same bilinear interpolation method as for PERSIANN-CDR for comparison, with the focus on global land areas between 60°S and 60°N.

As previously discussed, various studies have utilized PERSIANN-CDR, for different applications, in which PERSIANN-CDR has shown its usefulness. In this study, before getting into the details of using PERSIANN-CDR for evaluating CMIP5 climate models on capturing the behavior of daily extremes, we compared PERSIANN-CDR performance against CPC daily gauge analysis, as well as GPCP One-Degree Daily (GPCP1DD; Huffman

et al. 2001) and TRMM 3B42V7 products, in reproducing daily extreme precipitation statistics.

For these comparisons, PERSIANN-CDR, CPC daily gauge, GPCP1DD, and TRMM 3B42V7 datasets are spatially upscaled into $1° \times 1°$ grids. It is noteworthy that although PERSIANN-CDR and GPCP1DD are bias corrected against similar monthly satellite–rain gauge analyses, their daily variability is different.

3. Methodology

a. Spatial divisions

Given that climate zones are primarily defined by temperature and precipitation and two of the most heavily evaluated variables of interest in climate projections are also temperature and precipitation, an assessment of global climate model performance naturally lends itself to the division of study regions by climate zone. This is because the performance of an individual model within a given CCZ group is likely to be consistently representative throughout that particular group. A Köppen–Geiger climate classification map (available at http://koeppen-geiger.vu-wien.ac.at/shifts.htm) was used to define boundaries of global climate zones. This map was derived using Climatic Research Unit Time Series, version 2.1 (CRU TS 2.1), for temperature and Global Precipitation Climatology Centre Full, version 2 (GPCC Full v2), for precipitation data from 1976 to 2000 (Rubel and Kottek 2010). The 31 climate zones were aggregated to 5 main climate zones: equatorial, arid, warm temperate, snow, and polar, with their spatial resolution at 1°.

These climate zones were further divided by continent to at least partially account for spatial proximity. The resulting combination of continent–climate zone defined areas yielded 25 study regions (Fig. 1, Table 2).

It is acknowledged that because of climate change, the boundaries of these climate zones are destined to change as well. In fact, this is the topic discussed by the climate zone map developers (Rubel and Kottek 2010). However, since the map derived from observations from 1976 to 2000 provides substantial overlap with the historical evaluation period used for evaluation, coupled with the generalization from 31 to 5 climate zones, the areas depicted in Fig. 1 were deemed stable enough for this analysis. It is also acknowledged that the CCZ groups have different sizes and some CCZ groups (e.g., CCZ 1, 2) are separated into several regions using the CCZ method. Although the focus of this study is limited to grouping similar climate types for climate model evaluation, it should be recognized that such differences in CCZ group size may have an impact on the comparison with PERSIANN-CDR.

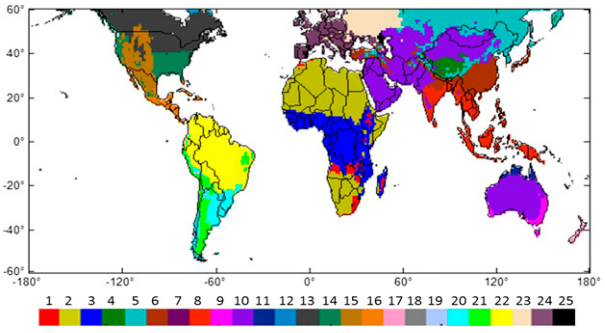

FIG. 1. The 25 CCZ groups.

b. Indices and statistical metrics

To demonstrate the capability of using PERSIANN-CDR as the "truth" to evaluate the performance of CMIP5 models, the performance of PERSIANN-CDR first should be ensured. Therefore, GPCP1DD is employed in this study to address this issue, particularly to demonstrate its capability in capturing the behavior of daily extreme precipitation with the statistical metrics used in this study. Eight daily precipitation extreme indices were computed for PERSIANN-CDR, GPCP1DD, TRMM 3B42V7, CPC daily gauge data, and 32 CMIP5 models at $1° \times 1°$ grid cells. Details of the indices are given in Table 3. Besides R10mmTOT, seven other daily extreme precipitation indices were selected from the recommended list by the joint Commission for Climatology (CCL)/Climate and Ocean: Variability, Predictability and Change (CLIVAR)/Joint Technical Commission for Oceanography and Marine Meteorology (JCOMM) Expert Team (ET) on Climate Change Detection and Indices (ETCCDI). The precipitation indices can be classified into four groups: total (R99pTOT, R95pTOT, R10mmTOT, and PRCPTOT), intensity [simple daily intensity index (SDII)], frequency (R10mm), and duration [consecutive wet days (CWD) and consecutive dry days (CDD)]. The indices are derived from daily data. For each year, the SDII is computed by taking the sum of precipitation in wet days (days with precipitation > 1 mm) and dividing that by the number of wet days.

To evaluate the performance of each of the CMIP5 models for each CCZ group based on PERSIANN-CDR as "observation," we calculated two commonly used comparison metrics: correlation (CORR) and relative root-mean-square error (RMSE). When comparing a daily extreme index of precipitation from

TABLE 2. Assigned numbers to continent and climate zone for each of 25 CCZ groups in Fig. 1. The numbers in parentheses show the area (10^3 km^2) of each CCZ group.

	Africa	Asia	Australia	North America	Oceania	South America	Europe
Polar	—	4 (1483)	—	12 (134)	—	19 (258)	—
Snow	—	5 (9536)	—	13 (7879)	—	—	23 (3000)
Warm temperate	1 (1561)	6 (5690)	9 (983)	14 (3525)	17 (289)	20 (3291)	24 (3965)
Arid	2 (17 144)	7 (14 530)	10 (6134)	15 (3838)	—	21 (2583)	25 (331)
Equatorial	3 (11 510)	8 (6338)	11 (674)	16 (1493)	18 (97)	22 (11 760)	—

each CMIP5 model with PERSIANN-CDR for a continent–climate zone group, CORR is calculated as follows:

$$\mathrm{CORR}=\frac{\sum_{i=1}^{n}(p_{\mathrm{CMIP5},i}-\overline{p}_{\mathrm{CMIP5}})\sum_{i=1}^{n}(p_{\mathrm{CDR},i}-\overline{p}_{\mathrm{CDR}})}{\sqrt{\sum_{i=1}^{n}(p_{\mathrm{CDR},i}-\overline{p}_{\mathrm{CDR}})^2}\sqrt{\sum_{i=1}^{n}(p_{\mathrm{CMIP5},i}-\overline{p}_{\mathrm{CMIP5}})^2}}. \quad (1)$$

For each CCZ group, we define relative RMSE of a daily extreme index of CMIP5 model k as follows:

$$\text{relative RMSE}_k=\frac{\mathrm{RMSE}_k-\mathrm{RMSE}_{\min}}{\mathrm{RMSE}_{\max}-\mathrm{RMSE}_{\min}}, \quad (2)$$

where $\mathrm{RMSE}_{\min}$ and $\mathrm{RMSE}_{\max}$ are the minimum and maximum RMSE values of the extreme index. RMSE is computed as follows:

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(p_{\mathrm{CMIP5},i}-p_{\mathrm{CDR},i})^2}, \quad (3)$$

where n is the total number of pixels in a given CCZ group; $p_{\mathrm{CMIP5},i}$ and $p_{\mathrm{CDR},i}$ are mean extreme indices of precipitation from the CMIP5 model and PERSIANN-CDR for pixel i; $\overline{p}_{\mathrm{CMIP5}}$ is the mean of the daily extreme index of precipitation from the CMIP5 model; and $\overline{p}_{\mathrm{CDR}}$ is the mean of the daily extreme index from PERSIANN-CDR for the CCZ group.

4. Results

As stated earlier, we first investigate the performance of PERSIANN-CDR against GPCP1DD. The resulting statistics are shown in Tables 4 and 5. As shown, there are generally high correlation coefficients and low RMSE between PERSIANN-CDR and GPCP1DD across all eight daily extreme precipitation indices. Most of the correlation coefficients are close to or larger than 0.9. For a few CCZs, this agreement is not significant. For instance, for CCZ 23 the correlation coefficient between the two products is low (except for SDII), which could be due to the snowy nature of precipitation in this CCZ. This evaluation over the period that GPCP1DD is available proves the accuracy of PERSIANN-CDR and its usefulness to be used for long-term validation of CMIP5 models over the past three decades.

Similarly, we investigated how PERSIANN-CDR compares against another high-resolution satellite-based precipitation product, TRMM 3B42V7, from 1998 to 2005. The derived correlation coefficient and RMSE values from this analysis are presented in Tables 6 and 7. In almost all the CCZs (except CCZ 17), correlation coefficients across all eight daily extreme indices are high (0.9 ± 0.1) between the two products.

At a more regional scale, we evaluated PERSIANN-CDR's performance against gridded gauge data analysis over the contiguous United States (CONUS), CPC daily gauge product, during the period of 1998–2005. Although the CPC daily gauge analysis is compromised

TABLE 3. Evaluation precipitation indices used in the analysis. A wet day has daily precipitation ≥ 1 mm.

Index	Definition	Unit
R99pTOT	Annual total precipitation when daily precipitation amount on a wet day is >99th percentile	mm
R95pTOT	Annual total precipitation when daily precipitation amount on a wet day is >95th percentile	mm
R10mmTOT	Annual total precipitation when daily precipitation amount is ≥10 mm	mm
PRCPTOT	Annual total precipitation in wet days	mm
SDII	Simple daily intensity index	mm day^{-1}
R10mm	Annual count of days when daily precipitation amount is ≥10 mm	days
CWD	Annual max number of consecutive days when daily precipitation amount is ≥1 mm	days
CDD	Annual max number of consecutive days when daily precipitation amount is <1 mm	days

TABLE 4. Correlation between extreme precipitation indices of PERSIANN-CDR and GPCP1DD from 1997 to 2005 [statistically significant correlations (at 0.05) are in boldface].

CCZ	CDD	CWD	PRCPTOT	R10mm	R10mmTOT	R95pTOT	R99pTOT	SDII
1	**0.957**	**0.906**	**0.979**	**0.949**	**0.976**	**0.963**	**0.932**	**0.828**
2	**0.959**	**0.930**	**0.998**	**0.977**	**0.967**	**0.986**	**0.968**	**0.901**
3	**0.908**	**0.865**	**0.986**	**0.940**	**0.927**	**0.916**	**0.908**	**0.860**
4	**0.809**	**0.943**	**0.993**	**0.980**	**0.968**	**0.960**	**0.893**	**0.940**
5	**0.868**	**0.584**	**0.762**	**0.905**	**0.922**	**0.770**	**0.739**	**0.840**
6	**0.965**	**0.896**	**0.990**	**0.975**	**0.981**	**0.956**	**0.874**	**0.879**
7	**0.958**	**0.832**	**0.986**	**0.923**	**0.949**	**0.962**	**0.927**	**0.880**
8	**0.983**	**0.852**	**0.991**	**0.983**	**0.979**	**0.916**	**0.891**	**0.859**
9	**0.941**	**0.894**	**0.990**	**0.983**	**0.981**	**0.969**	**0.927**	**0.878**
10	**0.960**	**0.921**	**0.997**	**0.977**	**0.987**	**0.981**	**0.949**	**0.905**
11	**0.946**	**0.782**	**0.979**	**0.970**	**0.974**	**0.739**	**0.667**	**0.486**
12	0.310	**0.567**	**0.616**	**0.758**	**0.772**	**0.443**	0.376	**0.995**
13	**0.718**	**0.505**	**0.813**	**0.899**	**0.888**	**0.805**	**0.764**	**0.954**
14	**0.965**	**0.908**	**0.991**	**0.981**	**0.982**	**0.935**	**0.878**	**0.863**
15	**0.940**	**0.866**	**0.984**	**0.927**	**0.920**	**0.903**	**0.835**	**0.773**
16	**0.959**	**0.918**	**0.988**	**0.982**	**0.983**	**0.944**	**0.891**	**0.928**
17	**0.862**	**0.580**	**0.750**	**0.867**	**0.522**	0.261	0.309	0.294
18	**0.990**	**0.953**	**0.999**	**0.997**	**0.994**	0.530	**0.812**	**0.745**
19	**0.987**	**0.881**	**0.976**	**0.949**	**0.946**	**0.888**	**0.758**	**0.891**
20	**0.927**	**0.868**	**0.989**	**0.978**	**0.983**	**0.960**	**0.916**	**0.894**
21	**0.956**	**0.867**	**0.972**	**0.902**	**0.913**	**0.926**	**0.885**	**0.877**
22	**0.973**	**0.859**	**0.991**	**0.984**	**0.979**	**0.903**	**0.769**	**0.783**
23	**0.474**	0.024	**0.165**	**0.391**	**0.427**	**0.232**	**0.230**	**0.823**
24	**0.917**	**0.318**	**0.782**	**0.819**	**0.789**	**0.622**	**0.572**	**0.681**
25	**0.885**	0.218	**0.976**	**0.912**	**0.834**	**0.731**	**0.577**	**0.609**

TABLE 5. RMSE between extreme precipitation indices of PERSIANN-CDR and GPCP1DD from 1997 to 2005 [statistically significant (at 0.05) are in boldface].

CCZ	CDD	CWD	PRCPTOT	R10mm	R10mmTOT	R95pTOT	R99pTOT	SDII
1	14.173	**18.605**	67.812	5.020	132.820	79.744	30.094	2.088
2	**26.498**	**4.632**	17.413	**3.196**	**91.105**	**63.003**	**31.064**	**2.957**
3	**15.377**	**21.758**	71.422	7.830	**239.192**	104.386	**35.823**	2.394
4	**31.349**	**5.808**	39.343	3.560	86.293	67.106	31.097	1.820
5	**14.849**	**3.762**	**172.052**	**4.607**	**78.384**	**68.472**	**23.609**	**1.256**
6	**7.630**	**11.406**	69.670	**5.533**	184.198	103.239	**37.891**	**3.666**
7	**22.222**	**2.681**	30.756	2.950	**62.062**	**51.629**	**26.750**	**2.437**
8	4.686	**36.015**	116.192	**9.100**	204.153	147.042	53.695	**2.816**
9	6.203	**2.739**	39.956	4.951	157.051	101.637	**40.148**	3.555
10	9.107	**4.423**	15.375	**3.632**	109.439	**90.266**	**46.306**	**4.312**
11	8.343	**17.139**	57.757	**6.105**	86.090	116.788	48.674	2.666
12	**14.338**	**4.833**	**360.785**	8.030	112.995	**135.925**	**44.338**	1.078
13	**5.725**	**3.879**	**194.261**	**8.175**	**141.499**	**65.599**	**21.946**	**1.305**
14	**7.170**	**8.181**	57.901	6.234	**237.578**	**124.881**	43.173	**4.171**
15	**10.008**	**4.285**	27.638	**4.798**	**102.022**	**71.968**	**32.653**	**2.457**
16	**5.534**	**30.784**	94.160	6.244	152.446	92.911	31.283	1.867
17	**3.700**	3.786	235.589	5.768	257.628	**146.961**	**53.695**	**3.193**
18	1.927	15.269	52.022	5.301	117.299	51.268	25.477	1.182
19	11.980	**14.975**	127.457	7.973	131.319	54.876	23.070	1.183
20	**6.511**	**7.846**	74.196	4.627	189.033	113.054	39.113	**5.009**
21	7.547	**7.794**	59.429	6.055	**137.011**	**85.827**	**34.868**	**2.587**
22	**5.119**	**25.985**	97.635	**5.391**	219.330	117.233	**41.268**	2.489
23	**5.524**	**2.806**	**268.368**	**7.428**	**116.992**	**92.013**	**34.249**	0.337
24	3.019	**4.167**	149.939	**7.128**	**144.267**	**68.580**	**27.503**	**1.802**
25	3.764	2.825	30.448	5.047	**115.954**	**74.566**	**32.520**	**2.337**

TABLE 6. Correlation between extreme precipitation indices of PERSIANN-CDR and TRMM 3B42V7 from 1998 to 2005 [statistically significant correlations (at 0.05) are in boldface].

CCZ	CDD	CWD	PRCPTOT	R10mm	R10mmTOT	R95pTOT	R99pTOT	SDII
1	**0.938**	**0.937**	**0.953**	**0.916**	**0.907**	**0.900**	**0.856**	**0.799**
2	**0.909**	**0.912**	**0.988**	**0.975**	**0.967**	**0.980**	**0.958**	**0.858**
3	**0.887**	**0.866**	**0.965**	**0.931**	**0.910**	**0.843**	**0.829**	**0.854**
4	**0.779**	**0.909**	**0.949**	**0.908**	**0.895**	**0.911**	**0.870**	**0.877**
5	**0.952**	**0.865**	**0.979**	**0.925**	**0.918**	**0.972**	**0.971**	**0.792**
6	**0.948**	**0.822**	**0.926**	**0.910**	**0.917**	**0.926**	**0.898**	**0.866**
7	**0.920**	**0.754**	**0.961**	**0.919**	**0.944**	**0.970**	**0.962**	**0.900**
8	**0.946**	**0.831**	**0.945**	**0.946**	**0.904**	**0.837**	**0.860**	**0.816**
9	**0.931**	**0.839**	**0.880**	**0.860**	**0.887**	**0.923**	**0.917**	**0.889**
10	**0.924**	**0.932**	**0.982**	**0.972**	**0.976**	**0.974**	**0.947**	**0.936**
11	**0.720**	**0.844**	**0.946**	**0.951**	**0.926**	**0.804**	**0.704**	**0.697**
12	0.108	0.368	**0.600**	**0.718**	**0.708**	**0.562**	**0.628**	**0.914**
13	**0.929**	**0.938**	**0.996**	**0.983**	**0.987**	**0.995**	**0.992**	**0.885**
14	**0.951**	**0.911**	**0.980**	**0.957**	**0.965**	**0.974**	**0.967**	**0.914**
15	**0.934**	**0.917**	**0.938**	**0.921**	**0.931**	**0.949**	**0.946**	**0.863**
16	**0.871**	**0.886**	**0.955**	**0.953**	**0.941**	**0.926**	**0.906**	**0.874**
17	**0.864**	0.352	**0.590**	**0.616**	**0.528**	**0.410**	0.289	0.329
18	**0.966**	**0.985**	**0.992**	**0.997**	**0.989**	0.283	0.232	**0.834**
19	**0.921**	**0.981**	**0.992**	**0.978**	**0.971**	**0.986**	**0.977**	**0.908**
20	**0.912**	**0.916**	**0.960**	**0.963**	**0.953**	**0.946**	**0.926**	**0.917**
21	**0.828**	**0.881**	**0.879**	**0.938**	**0.946**	**0.953**	**0.951**	**0.916**
22	**0.940**	**0.896**	**0.972**	**0.967**	**0.953**	**0.866**	**0.758**	**0.766**
23	**0.977**	**0.849**	**0.992**	**0.981**	**0.978**	**0.988**	**0.983**	**0.726**
24	**0.979**	**0.825**	**0.983**	**0.975**	**0.969**	**0.983**	**0.978**	**0.768**
25	**0.934**	**0.583**	**0.946**	**0.949**	**0.941**	**0.936**	**0.879**	**0.640**

TABLE 7. RMSE between extreme precipitation indices of PERSIANN-CDR and TRMM 3B42V7 from 1997 to 2005 [statistically significant (at 0.05) are in boldface].

CCZ	CDD	CWD	PRCPTOT	R10mm	R10mmTOT	R95pTOT	R99pTOT	SDII
1	17.557	**10.097**	94.493	5.565	144.009	80.877	37.133	**1.415**
2	**31.939**	**3.846**	**42.587**	**1.992**	**51.242**	**43.399**	**27.337**	**1.956**
3	16.455	**16.356**	112.949	**6.119**	**191.821**	**134.582**	**58.711**	1.592
4	**28.094**	**4.383**	**135.741**	**6.344**	**107.192**	**50.621**	**17.496**	**0.801**
5	**7.511**	**2.991**	**62.551**	3.133	**74.335**	**46.338**	**20.020**	**1.457**
6	9.498	**8.088**	**208.958**	7.926	**233.574**	**117.525**	**49.392**	**2.297**
7	25.438	**2.818**	53.651	**2.416**	45.478	**32.674**	**20.799**	**1.642**
8	**10.595**	**29.126**	282.726	**16.254**	323.276	**200.697**	**87.798**	**1.990**
9	6.123	1.837	122.385	**5.038**	128.579	**84.933**	**36.341**	**2.058**
10	**13.622**	**3.126**	38.473	**2.439**	65.344	**57.916**	**35.106**	2.172
11	32.797	**12.219**	102.356	**9.779**	73.479	**132.076**	**73.497**	**1.789**
12	**1.298**	**0.546**	**41.028**	1.001	15.515	**16.090**	**5.526**	1.606
13	**3.499**	**2.513**	45.751	**3.097**	**114.973**	**80.478**	**34.064**	**2.189**
14	7.694	**5.074**	86.541	4.017	**176.936**	**135.092**	**58.362**	**3.264**
15	9.818	**2.887**	63.118	2.894	**61.215**	**44.101**	**22.984**	**1.285**
16	11.141	**22.492**	192.695	9.036	218.144	**128.578**	**54.079**	1.378
17	6.377	**3.694**	534.036	20.795	400.006	**144.288**	**55.470**	**1.878**
18	6.573	15.831	239.607	7.709	326.291	**242.751**	99.768	**1.864**
19	16.849	**10.587**	121.956	**5.068**	61.914	**16.505**	10.041	0.516
20	**10.213**	**5.494**	141.409	**5.149**	192.865	**135.482**	**59.510**	**3.497**
21	**32.656**	**6.314**	160.393	3.174	**81.856**	**57.136**	**27.627**	**1.637**
22	**12.535**	**18.063**	170.495	**8.626**	203.686	**124.303**	**51.748**	**1.537**
23	**1.744**	2.034	32.273	**1.520**	**34.667**	**21.804**	**9.838**	0.661
24	**6.301**	**3.189**	93.822	2.832	**87.494**	**58.374**	**26.895**	**1.504**
25	**8.169**	**3.277**	89.222	2.249	51.405	34.529	20.471	1.218

TABLE 8. Correlation between extreme precipitation indices of PERSIANN-CDR and CPC daily gauge data over the CONUS from 1998 to 2005 [statistically significant correlations (at 0.05) are in boldface].

CCZ	CDD	CWD	PRCPTOT	R10mm	R10mmTOT	R95pTOT	R99pTOT	SDII
13	**0.983**	**0.964**	**0.990**	**0.988**	**0.988**	**0.994**	**0.993**	**0.954**
14	**0.983**	**0.962**	**0.988**	**0.981**	**0.981**	**0.990**	**0.988**	**0.905**
15	**0.977**	**0.944**	**0.953**	**0.922**	**0.923**	**0.974**	**0.973**	**0.886**

by sparse station networks, quality of the analysis over grid boxes where there is at least one reporting gauge inside a 0.25° grid box is reliable. However, it is also noteworthy that using interpolation techniques to combine point measurements into a gridded product can introduce uncertainties in the final gridded product, especially at remote and ungauged regions where a sufficient number of gauge stations, if any, is not available. As presented in Tables 8 and 9, there exist significant agreement between PERSIANN-CDR and CPC daily gauge data with low RMSEs and high correlation coefficient values ranging from 0.886 for SDII to 0.994 for R95pTOT.

The result of the above analyses, tests, and examinations proves that PERSIANN-CDR can be a reliable data source to serve as the reference dataset for evaluating climate model's historical simulations, for a long-term period and at a global scale.

To better understand why the evaluation of different climate models' simulations against a global set of observations is important, we computed the mean precipitation total (PRCPTOT) for PERSIANN-CDR and for the 32 CMIP5 models (Fig. 2) for 1983–2005. One can find the differences in spatial patterns of the annual total precipitation in wet days from different CMIP5 models and also from PERSIANN-CDR. Figure 2 indicates different CMIP5 models perform quite differently from each other. For example, in Africa, mean PRCPTOT of BNU-ESM is much larger than that of EC-EARTH. This means the evaluation of CMIP5 models is of great importance.

For each precipitation index within each CCZ group, a comparison between the individual 32 CMIP5 models is made to that of the corresponding PERSIANN-CDR–derived precipitation index. For a given index with this strategy, each pixel within the CCZ group of focus will have a sample of that index (mean of indices from 23 years of data under consideration) for each model. In other words, a CCZ group composed of 100 pixels will have 100 samples of a given index for each model. For each model, these samples are plotted against those obtained from PERSIANN-CDR estimates, and RMSE and correlation are calculated from the linear fit of these samples. Figure 3 illustrates a sample of the method we used to compare each daily extreme index for each CMIP5 model against PERSIANN-CDR. A correlation of 0.683 and an RMSE of 352.923 mm between PRCPTOT from PERSIANN-CDR and precipitation from the CNRM-CM5 model were calculated for CCZ 3 (Africa equatorial) in this particular example. Relative RMSE was calculated based on Eq. (3).

Following the analyses illustrated in Fig. 3, correlation and relative RMSE were calculated for each precipitation index, for each model, in each of the 25 CCZ groups against PERSIANN-CDR as the baseline. These statistical results are collected by climate zone type and summarized in portrait diagrams (Figs. 4–8; note that Figs. 5–8 are described in greater detail below). Statistics for equatorial CCZ groups (Fig. 4) are largely mixed for correlation and mostly moderate (around 0.5) for relative RMSE. One standout for this climate zone is northern South America (CCZ 22), which touts low relative RMSE and high correlation for nearly all models for the R10mmTOT index. Similar performance for CCZ 22 was observed for the SDII index as well. R95pTOT seems to be doing just as poorly or worse for all models and CCZs (except 3 and 8) based on Fig. 4. Among the eight metrics, the correlation coefficient for R99pTOT index is the worst across almost all the CMIP5 climate models and all equatorial CCZ groups except Asia equatorial (CCZ 8). The correlation coefficients of R10mmTOT index vary across equatorial

TABLE 9. RMSE between extreme precipitation indices of PERSIANN-CDR and CPC daily gauge data from 1983 to 2005 [statistically significant (at 0.05) are in boldface].

CCZ	CDD	CWD	PRCPTOT	R10mm	R10mmTOT	R95pTOT	R99pTOT	SDII
13	**3.546**	0.958	**59.324**	**1.433**	24.313	14.166	**6.198**	0.778
14	**6.359**	1.510	90.964	3.433	79.875	47.327	**20.106**	**1.248**
15	**9.602**	1.479	68.639	**2.238**	**41.414**	**29.603**	**13.702**	**0.895**

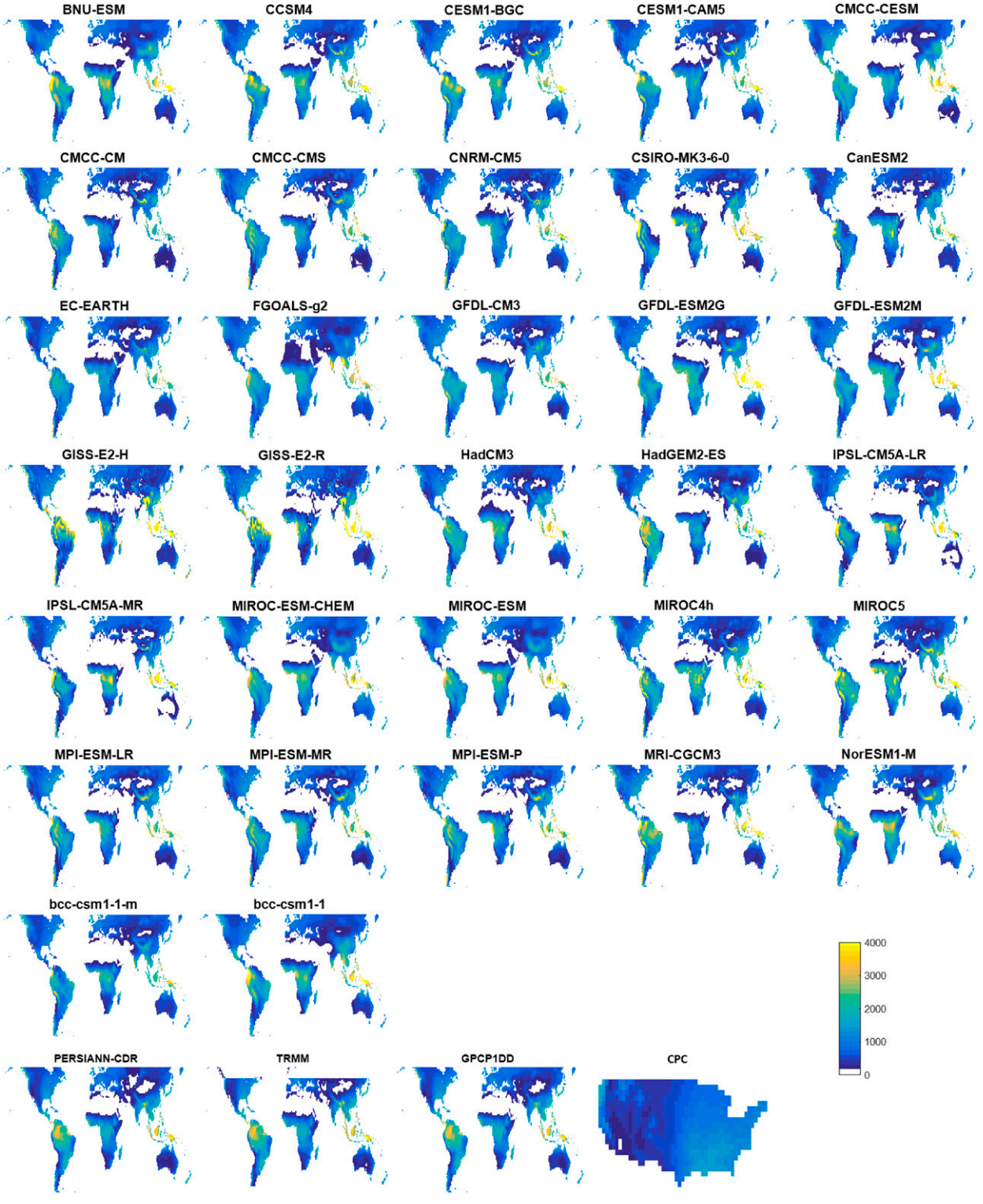

FIG. 2. Mean PRCPTOT (mm) for PERSIANN-CDR and 32 CMIP5 models from 1983 to 2005.

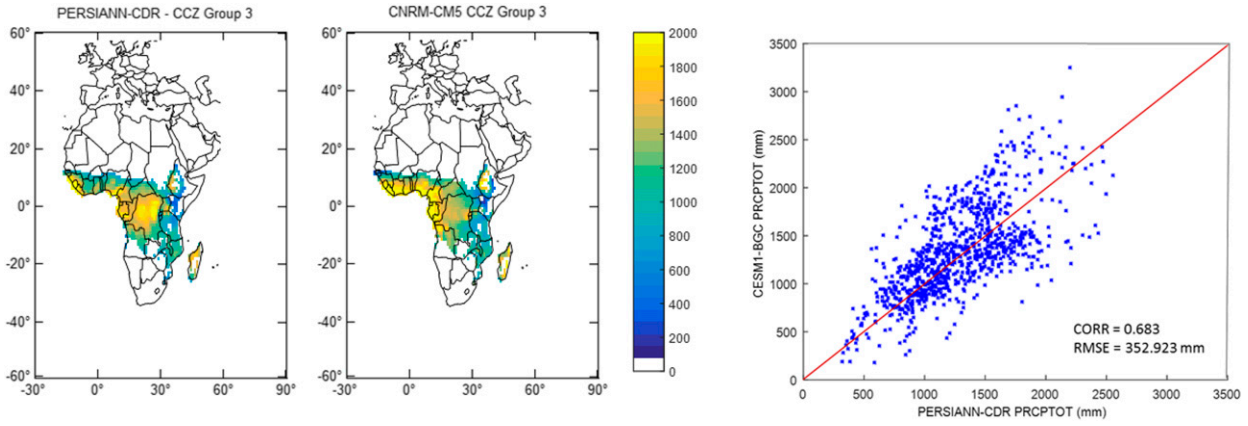

FIG. 3. (right) Correlation between (left) PERSIANN-CDR PRCPTOT (mm) and (center) CNRM-CM5 PRCPTOT (mm) for CCZ 3.

CCZ groups from lower than −0.5 (CCZ 16) to larger than 0.9 (CCZ 20), and so does R95pTOT index.

Arid CCZ groups (Fig. 5) have high correlations for most models and across all indices, particularly PRCPTOT and CDD. The correlations between daily extreme indices (e.g., SDII and CWD indices) from CMIP5 models 10 (CanESM2), 16 (GISS-E2-H), and 17 (GISS-E2-R) and that from PERSIANN-CDR are relatively worse than the other models. In terms of the RMSE, markedly poor performance in most models appears in PRCPTOT for CCZ 15 (North America arid), especially CMIP5 models 5, 6, 7, 20, and 21. Similar poor performance in most models for SDII, CWD, and R10mmTOT appears in CCZ 25 (Europe arid).

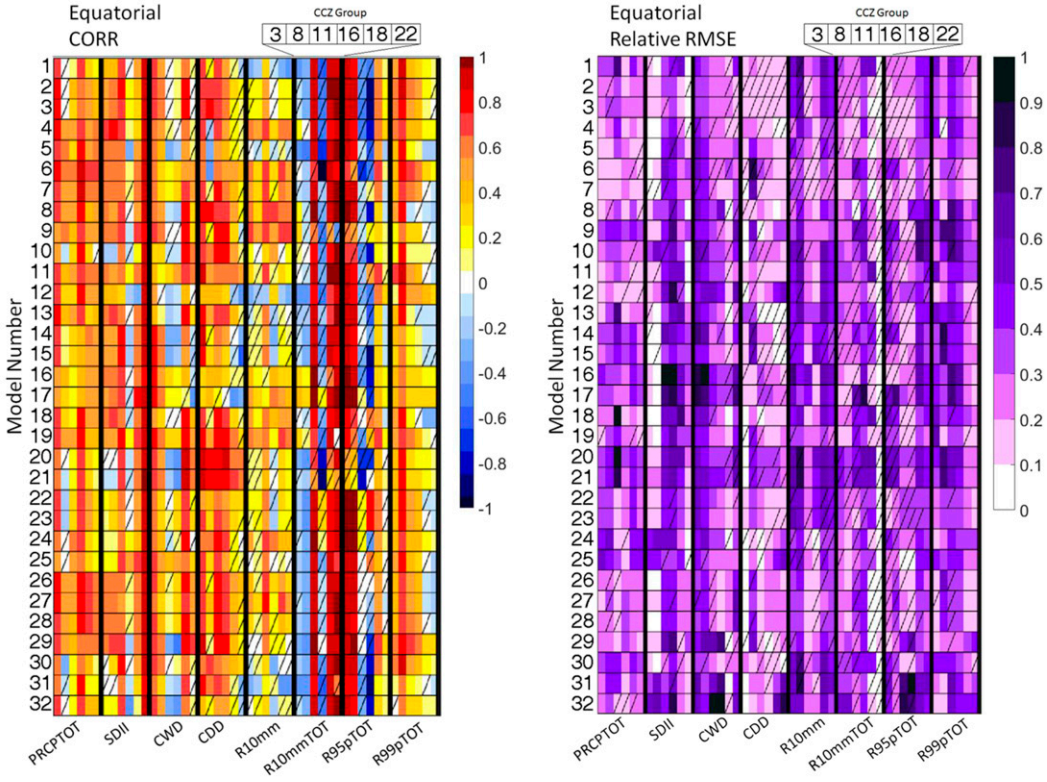

FIG. 4. (left) Correlation and (right) relative RMSE for precipitation indices in equatorial CCZ groups [statistically insignificant correlations (at 0.05) in hatched boxes].

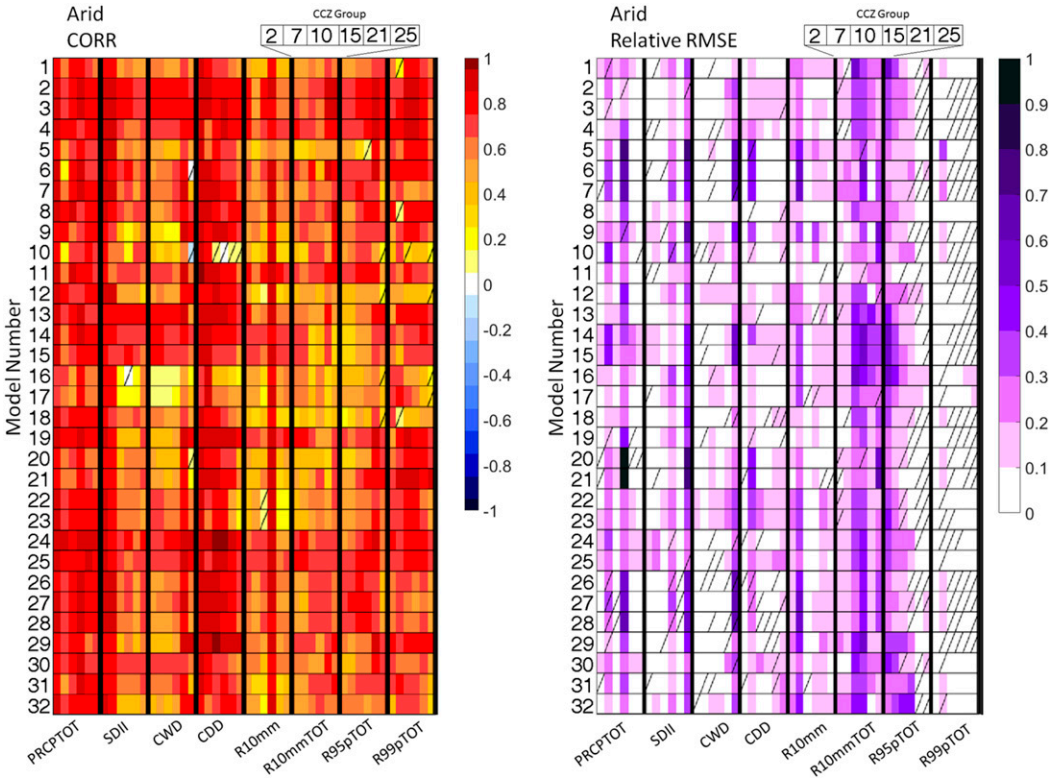

FIG. 5. As in Fig. 4, but for arid CCZ groups.

Among the eight indices, CWD and CDD perform better than the others in terms of correlation efficient, in which R10mm shows the worst performance. In terms of relative RMSE, R99pTOT behaves the best, while R10mmTOT performs the worst in terms of RMSE.

Warm temperate CCZ groups (Fig. 6) have fairly high (~0.5–0.9) correlation and low relative RMSE (~0.2, except for models 16 and 17). The results show that almost all the CMIP5 climate models depict high correlation coefficients across all the daily extreme precipitation indices, with the exception of R10mmTOT few models [e.g., model 12 (FGOALS-g2)]. With respect to the error between climate models and PERSIANN-CDR, CMIP5 models 16 and 17 exhibit relatively worse performance than the other models. These two models have comparably high values over CCZ 6, 9, and 25 in SDII. Similar to arid CCZ groups, CDD performs better than the others in terms of correlation efficient and R99pTOT behaves the best in terms of RMSE.

Cold regions (snow and polar in Figs. 7 and 8, respectively) generally show low relative RMSE compared to the other three climate zone varieties. The main exception occurs over CCZ 23 (Europe snow) for R10mmTOT index, where almost all CMIP5 models show poor performance. With respect to the correlation coefficients between different models and PERSIANN-CDR precipitation, generally high values are observed for intensity (SDII), frequency (R10mm), and duration (CWD, CDD) indices. With respect to the daily extreme total indices over the snow CCZ groups, almost all CMIP5 models depict poor performances in R95pTOT and R99pTOT with low correlation coefficients (~0.2). RMSE is low for almost all models for R10mm in the polar groups according to Fig. 8. These climate zones show relatively high correlations for most models in all indices except R95pTOT and R99pTOT.

Regarding the performance of CMIP5 models in different climate zones, it can be seen that the performance of arid zones is the best in terms of correlation coefficient. However, snow zones exhibit outstanding performance in terms of relative root-mean-square error. This may be because, like the arid zones, both models and PERSIANN-CDR agree on small precipitation amounts overall for snow zones (Fig. 2). With less precipitation to capture, it becomes easier to obtain a lower RMSE. Unlike arid zones, the snow regions have an even greater advantage in that the little precipitation that is received is less likely to be in the form of convective events with intensities that may be more difficult to capture. This gives the snow zones a

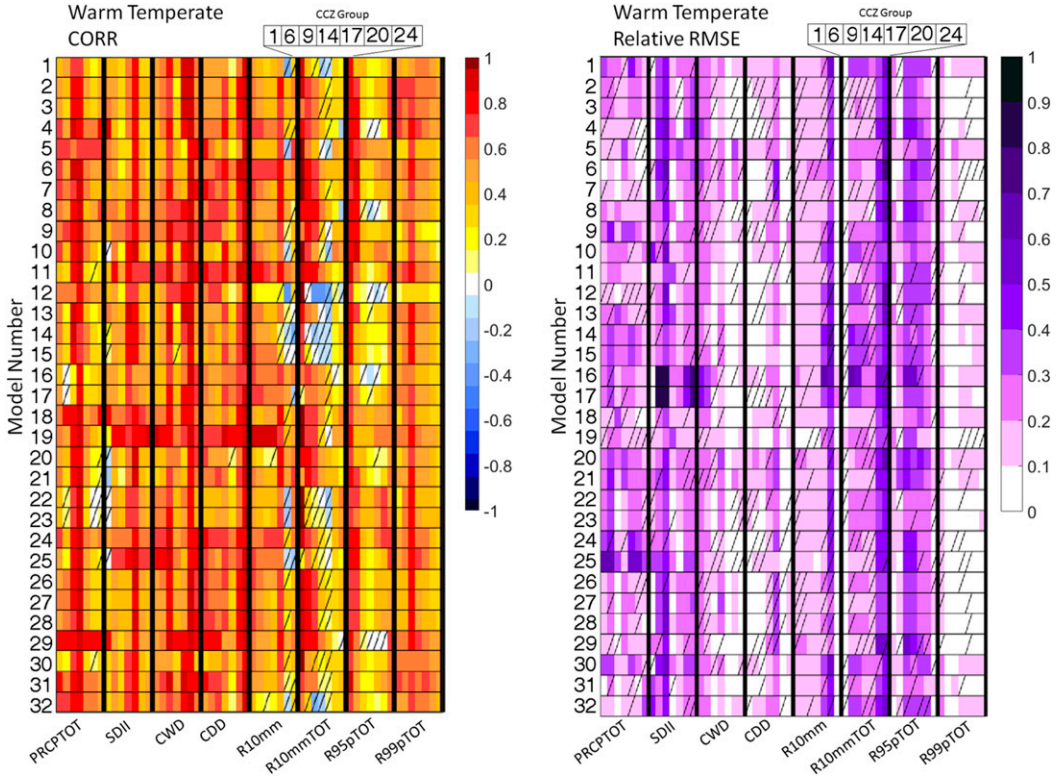

FIG. 6. As in Fig. 4, but for warm temperate CCZ groups.

slight edge over the arid regions in terms of RMSE and is reflected in Figs. 5 and 7.

Because of the choice of dividing study regions into CCZ groups, the analysis incorporates areas of varying sizes (see Table 2). One of the largest CCZ groups (22) was divided into uniform 6° × 6° areas (roughly the size of the smallest CCZ groups) to provide a quantitative sample of the uncertainty resulting from this division choice. After discarding subareas smaller than the smallest CCZ group (18), which appear along the irregular boundaries of CCZ 22, the analysis was performed on 33 subareas (Figs. 9, 10). Some of the indices (especially PRCPTOT, R10mm, and R99pTOT) are only slightly affected by a large spatial extent difference, as the statistics for the 33 subareas cluster tightly around the complete CCZ 22 CORR and relative RMSE values for most CMIP5 models. The models examined exhibit a degradation in performance in SDII CORR and in both CORR and relative RMSE for R10mmTOT when transitioning from a large analysis area to small areas. Other indices exhibit mixed levels of representation of CCZ 22 by the 33 subareas. R95pTOT, for example, shows that the subarea populations characterize the large CCZ 22 CORR quite well, but consistently outperform the large area in terms of RMSE. However, whether improving, degrading, or not changing performance of CORR and relative RMSE, the intermodel performance is broadly preserved as the spatial scale is refined.

5. Summary and discussion

This study provides a first overview of the performance of all CMIP5 models in simulating climate daily extreme indices in comparison to PERSIANN-CDR precipitation estimates with high resolution on daily and nearly global scales. The set of 8 indices is calculated for 32 CMIP5 models over 25 CCZ groups.

We use PERSIANN-CDR as the reference data for CMIP5 model evaluation on the assumption that it is accurate enough for the study. It is acknowledged that satellite-based precipitation estimates are subject to uncertainties, for example, systematic errors. However, the assumption is based on the fact that PERSIANN-CDR has been tested for its performance for different applications against ground-truth observational data and also in comparison to other satellite-based products, as explained in section 2. The use of multiple precipitation estimates does not reduce the uncertainty; rather, it shows that there is uncertainty in all estimates and tries to quantify that uncertainty. The quantitative

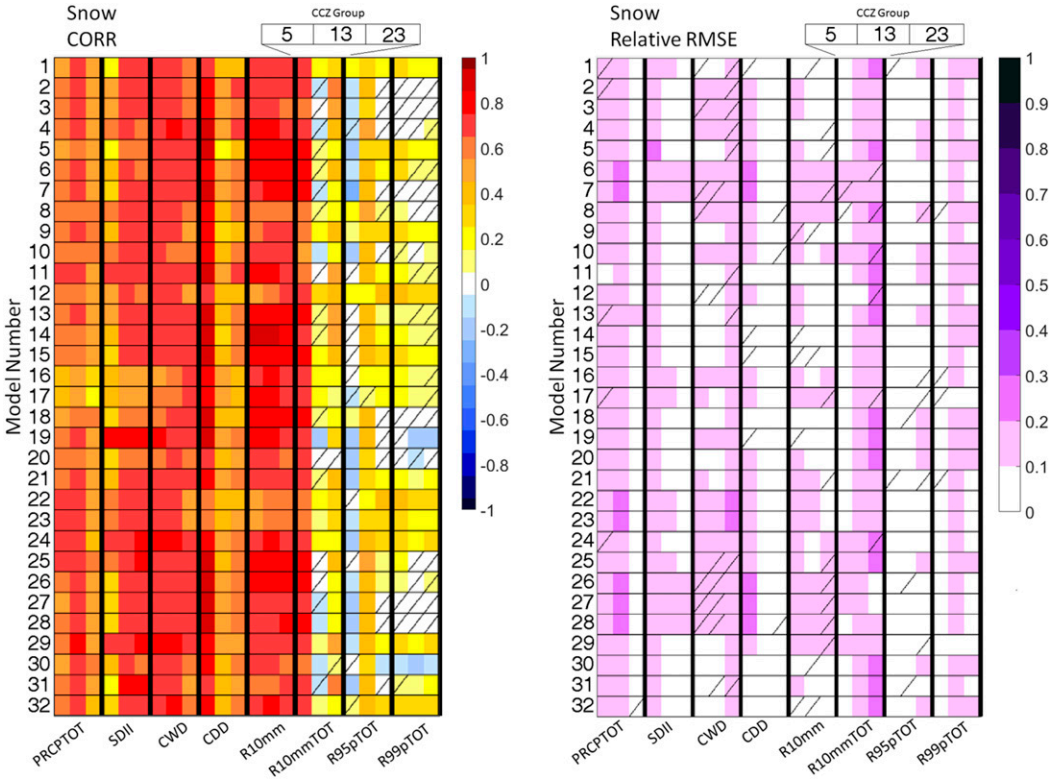

FIG. 7. As in Fig. 4, but for snow CCZ groups.

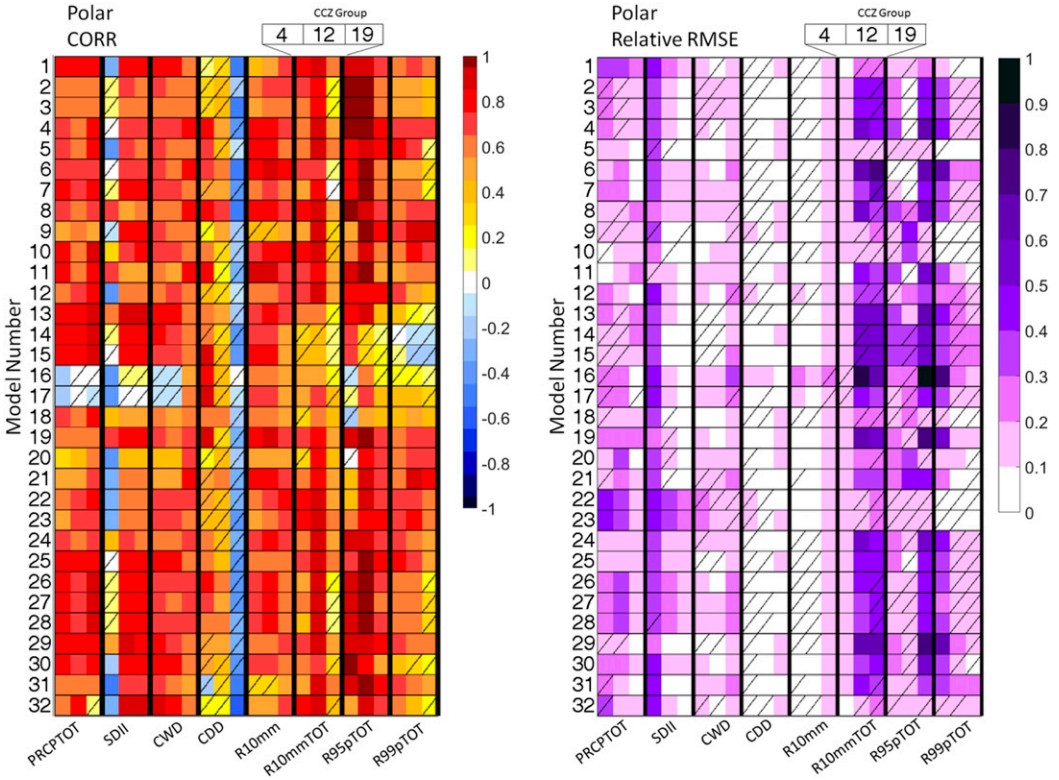

FIG. 8. As in Fig. 4, but for polar CCZ groups.

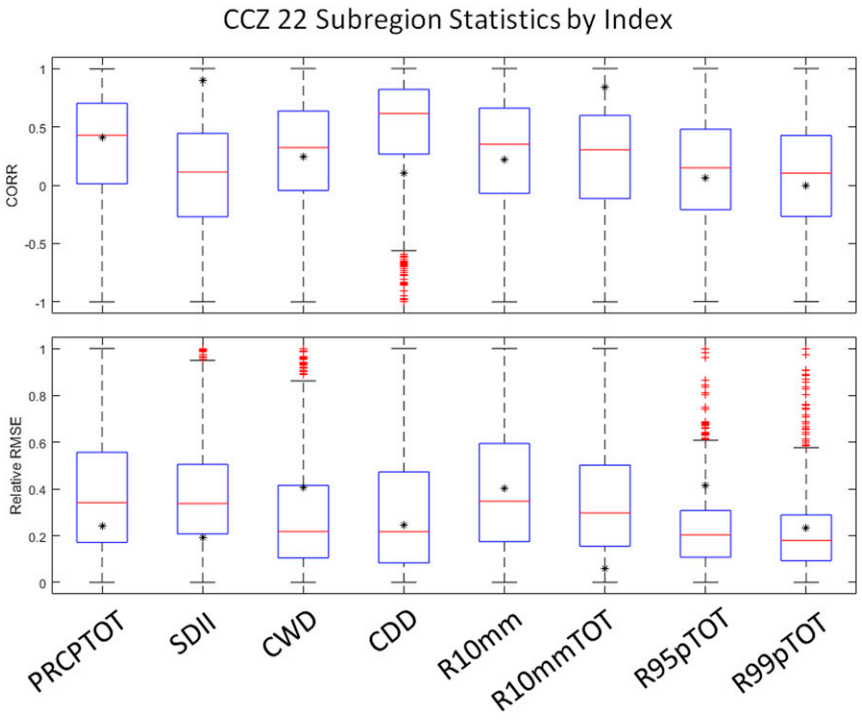

FIG. 9. Box-and-whisker plots of (top) CORR and (bottom) relative RMSE values for subregions in CCZ 22. Each plot is a grouping of all CORR or relative RMSE values for all models and all 33 subareas in CCZ 22. Red lines indicate means of subarea samples, red crosses indicate outliers, and black stars indicate corresponding statistics (averaged over the 33 models) of the entire CCZ 22 region.

comparisons and analysis between PERSIANN-CDR and GPCP1DD, TRMM 3B42V7, and U.S. CPC daily gauge data products proved that PERSIANN-CDR is effective as a reference dataset used in evaluating the performance of CMIP5 models. PERSIANN-CDR has its priority for climate studies with its long period (from 1983 to near present), near-global coverage, and high resolution. It is a useful dataset for addressing various key climatological and hydrological research questions that require longer and finer-resolution data than previously available.

This study examines performance of climate models based on their subdivision into climate groups and continents. These CCZ groups are a natural division unit for evaluation of climate models since the CCZ groups are defined by climate-scale precipitation and temperature and the primary outputs for the climate models are also precipitation and temperature. Such a grouping is useful not only for users of climate model output who would like to leverage the strengths of the top-performing models, but also for climate model developers. Through this categorization, developers get a more complete picture of how each model performs in a specific environment, which may then be related to how well a given model is able to capture certain processes unique to that environment. Having insight on the strengths and weaknesses of a given model will allow developers to focus their efforts where they are most needed. For example, it is obvious that the GISS models perform poorly in polar regions and many others perform the worst in equatorial regions.

The performance of percentile indices (R95pTOT and R99pTOT) indicates that all CMIP5 models exhibit a deficit in reproducing observed precipitation characteristics in daily extremes over snow and polar CCZ groups. The inconsistency originates from the fact that the PERSIANN-CDR algorithm is designed to estimate liquid precipitation while these regions are mostly snowy regions. It indicates that CMIP5 models' precipitation estimates capture the tendency (CORR) and amount (RMSE) of PERSIANN-CDR observations over arid CCZ groups better than the other CCZ groups. However, there is an exception, which is the snow CCZ group, over which CWD and R10mm have better correlation than the other CCZ groups. This indicates CMIP5 models have the capability to capture the frequency and duration of wet days. The persistence of these errors in all 32 evaluated CMIP5 simulations indicates the models have a

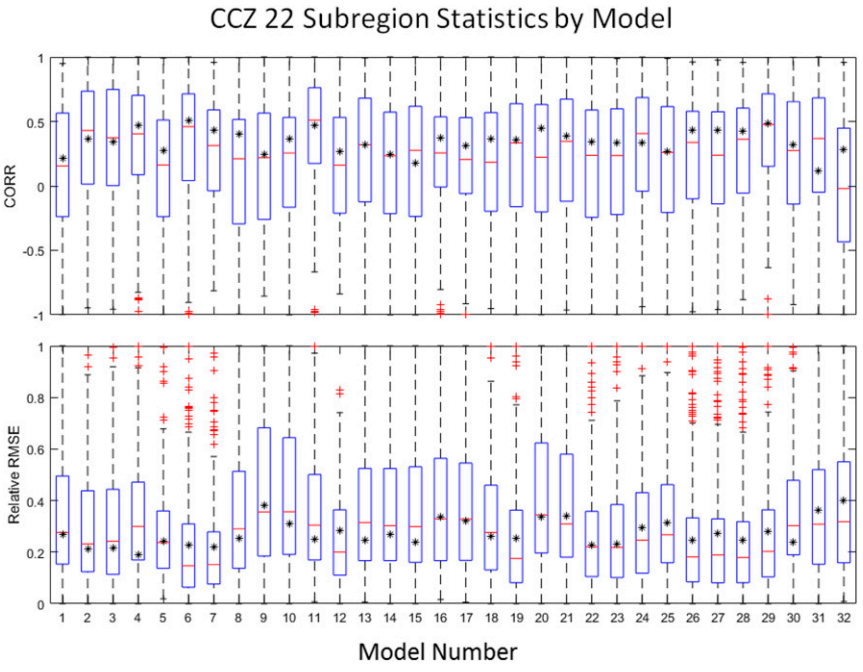

FIG. 10. Box-and-whisker plots of (top) CORR and (bottom) relative RMSE values for subregions in CCZ 22. Each plot is a grouping of all CORR or relative RMSE values for all indices and all 33 subareas in CCZ 22. Red lines indicate means of subarea samples, red crosses indicate outliers, and black stars indicate corresponding statistics (averaged over the eight indices) of the entire CCZ 22 region.

general deficiency in simulating the daily extreme upper tail of the precipitation distribution.

The inconsistent correlation performance across various models and precipitation indices within the equatorial and, to a slightly lesser extent, warm temperate CCZ groups is not entirely surprising given the more complex precipitation patterns in these areas. Similarly, these two categories of climate zones tend to show higher RMSE values for most indices compared to the other three climate zone categories. This is somewhat of an artifact of normalizing all RMSE values globally. Since zero precipitation days are effectively removed from some indices (R99pTOT, R95pTOT, R10mmTOT, and PRCPTOT) and dominate others (SDII, R10mm, CWD, and CDD) in the comparably drier arid, polar, and snow areas, these regions have the advantage of capturing the more simplistic dry tendencies. On this note, the result that the arid regions had the highest correlation and lowest RMSE out of the climate zone types is expected.

A sample investigation into the role of analyzing different-sized regions reveals that larger areas may have an overall advantage in terms of SDII CORR and R10mmTOT CORR and relative RMSE, and that PRCPTOT and R99pTOT CORR and relative RMSE and R95pTOT CORR are unaffected. Comparisons between models reveal that a rank in model performance according to the metrics used in this study is largely preserved through the spatial-scale change.

As shown in the initial part of the analysis, PERSIANN-CDR in general depicts high correlation coefficients and low RMSE across all the daily extreme indices and climate zones (with few exceptions), when compared against three precipitation data products, GPCP1DD, TRMM 3B42V7, and CPC daily gauge data. Therefore, one can conclude that PERSIANN-CDR validates well against gauge-based and other high-resolution satellite-based precipitation products in reproducing the daily extreme precipitation indices and can be used for model evaluation. Despite this, and as is the case with all the satellite-based products, we acknowledge that the uncertainties in PERSIANN-CDR can also contribute to results we are deriving; however, given the favorable results from the aforementioned evaluation of PERSIANN-CDR, it does not seem to be significant.

This study provides insightful suggestions for CMIP5 model selection for hydrometeorological and climate studies. For instance, one can use the correlation and/or RMSE criteria in Fig. 6 to select a number of models among the 32 CMIP5 models for southeastern China (CCZ 6) based on one or some of the eight precipitation indices. If the selection is based on the SDII with the

highest correlation, HadGEM2-ES, EC-EARTH, and MIROC5 are the three best choices, respectively. Such model selection can be adapted to any region globally depending on the phenomenon of interest and can be combined with user-specific criteria.

As some previous works (Kim et al. 2012; Fang and Li 2016) suggested, the performance of the CMIP5 models in simulating precipitation can be improved with some model averaging techniques. However, it is unreasonable to utilize all the CMIP5 models in some techniques (e.g., Bayesian model averaging). This study can provide good suggestions for hydrometeorology researchers to select those CMIP5 models that perform well over the specific region to do the multimodel ensemble simulations. Besides, this study informs the model developers of the strengths and weakness of their models. At the same time, the study provides insights on which models should be chosen when concentrating on the daily extremes of precipitation and which models should be selected when emphasizing the total amount of precipitation.

Acknowledgments. This research was partially supported by the Cooperative Institute for Climate and Satellites (CICS) program (NOAA Prime Award NA14NES4320003, Subaward 2014-2913-03) for OHD-NWS student fellowship, the Army Research Office (Award W911NF-11-1-0422), National Science Foundation (NSF Award 1331915), and Department of Energy (DOE Prime Award DE-IA0000018).

REFERENCES

Adler, R. F., and Coauthors, 2003: The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeor.*, **4**, 1147–1167, doi:10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2.

Ashouri, H., K. Hsu, S. Sorooshian, D.K. Braithwaite, K.R. Knapp, L.D. Cecil, B.R. Nelson, and O.P. Prat, 2015: PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bull. Amer. Meteor. Soc.*, **96**, 69–83, doi:10.1175/BAMS-D-13-00068.1.

——, P. Nguyen, A. Thorstensen, K. Hsu, S. Sorooshian, and D. Braithwaite, 2016a: Assessing the efficacy of high-resolution satellite-based PERSIANN-CDR precipitation product in simulating streamflow. *J. Hydrometeor.*, **17**, 2061–2076, doi:10.1175/JHM-D-15-0192.1.

——, S. Sorooshian, K. Hsu, M. G. Bosilovich, J. Lee, and M. F. Wehner, 2016b: Evaluation of NASA's MERRA precipitation product in reproducing the observed trend and distribution of extreme precipitation events in the United States. *J. Hydrometeor.*, **17**, 693–711, doi:10.1175/JHM-D-15-0097.1.

Bosilovich, M. G., J. Chen, F. R. Robertson, and R. F. Adler, 2008: Evaluation of global precipitation in reanalyses. *J. Appl. Meteor. Climatol.*, **47**, 2279–2299, doi:10.1175/2008JAMC1921.1.

Burroughs, W., 2003: *Climate: Into the 21st Century*. Cambridge University Press, 240 pp.

Casse, C., and M. Gosset, 2015: Analysis of hydrological changes and flood increase in Niamey based on the PERSIANN-CDR satellite rainfall estimate and hydrological simulations over the 1983–2013 period. *Proc. IAHS*, **370**, 117–123, doi:10.5194/piahs-370-117-2015.

Fang, M., and X. Li, 2016: Application of Bayesian model averaging in the reconstruction of past climate change using PMIP3/CMIP5 multi-model ensemble simulations. *J. Climate*, **29**, 175–189, doi:10.1175/JCLI-D-14-00752.1.

Flato, G., and Coauthors, 2013: Evaluation of climate models. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 741–866.

Gaetani, M., and E. Mohino, 2013: Decadal prediction of the Sahelian precipitation in CMIP5 simulations. *J. Climate*, **26**, 7708–7719, doi:10.1175/JCLI-D-12-00635.1.

Guo, H., S. Chen, A. Bao, J. Hu, A. S. Gebregiorgis, X. Xue, and X. Zhang, 2015: Inter-comparison of high-resolution satellite precipitation products over central Asia. *Remote Sens.*, **7**, 7181–7211, doi:10.3390/rs70607181.

Held, I. M., 2005: The gap between simulation and understanding in climate modeling. *Bull. Amer. Meteor. Soc.*, **86**, 1609–1614, doi:10.1175/BAMS-86-11-1609.

Hsu, K., X. Gao, S. Sorooshian, and H. V. Gupta, 1997: Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks. *J. Appl. Meteor. Climatol.*, **36**, 1176–1190, doi:10.1175/1520-0450(1997)036<1176:PEFRSI>2.0.CO;2.

Huffman, G. J., 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. Hydrometeor.*, **2**, 36–50, doi: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. Hydrometeor.*, **8**, 38–55, doi:10.1175/JHM560.1.

IPCC, 2013: *Climate Change 2013: The Physical Science Basis*. Cambridge University Press, 1535 pp., doi:10.1017/CBO9781107415324.

Jiang, Z., J. Song, L. Li, W. Chen, Z. Wang, and J. Wang, 2012: Extreme climate events in China: IPCC-AR4 model evaluation and projection. *Climatic Change*, **110**, 385–401, doi:10.1007/s10584-011-0090-0.

Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeor.*, **5**, 487–503, doi:10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2.

Katiraie-Boroujerdy, P.-S., H. Ashouri, K. Hsu, and S. Sorooshian, 2017: Trends of precipitation extreme indices over a subtropical semi-arid area using PERSIANN-CDR. *Theor. Appl. Climatol.*, doi:10.1007/s00704-016-1884-9, in press.

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.

Knapp, K. R., 2008: Scientific data stewardship of International Satellite Cloud Climatology Project B1 global geostationary observations. *J. Appl. Remote Sens.*, **2**, 023548, doi:10.1117/1.3043461.

Liu, Z. Y., and Coauthors, 2014: Chinese cave records and the East Asia summer monsoon. *Quat. Sci. Rev.*, **83**, 115–128, doi:10.1016/j.quascirev.2013.10.021.

Luchetti, N. T., J. R. Sutton, E. E. Wright, M. C. Kruk, and J. J. Marra, 2016: When El Niño rages: How satellite data can help water-stressed islands. *Bull. Amer. Meteor. Soc.*, **97**, 2249–2255, doi:10.1175/BAMS-D-15-00219.1.

Mehran, A., A. AghaKouchak, and T. J. Phillips, 2014: Evaluation of CMIP5 continental precipitation simulations relative to satellite-based gauge-adjusted observations. *J. Geophys. Res. Atmos.*, **119**, 1695–1707, doi:10.1002/2013JD021152.

Miao, C., H. Ashouri, K. Hsu, S. Sorooshian, and Q. Duan, 2015: Evaluation of the PERSIANN-CDR daily rainfall estimates in capturing the behavior of extreme precipitation events over China. *J. Hydrometeor.*, **16**, 1387–1396, doi:10.1175/JHM-D-14-0174.1.

Neelin, J. D., M. Münnich, H. Su, J. E. Meyerson, and C. E. Holloway, 2006: Tropical drying trends in global warming models and observations. *Proc. Natl. Acad. Sci. USA*, **103**, 6110–6115, doi:10.1073/pnas.0601798103.

Nguyen, P., and Coauthors, 2017: Exploring trends through ''RainSphere'': Research data transformed into public knowledge. *Bull. Amer. Meteor. Soc.*, **98**, 653–658, doi:10.1175/BAMS-D-16-0036.1.

Peterson, T. C., P. A. Stott, and S. Herring, Eds., 2012: Explaining extreme events of 2011 from a climate perspective. *Bull. Amer. Meteor. Soc.*, **93**, 1041–1067, doi:10.1175/BAMS-D-12-00021.1.

Reichler, T., and J. Kim, 2008: How well do coupled models simulate today's clime? *Bull. Amer. Meteor. Soc.*, **89**, 303–311, doi:10.1175/BAMS-89-3-303.

Rossow, W. B., A. Mekonnen, C. Pearl, and W. Goncalves, 2013: Tropical precipitation extremes. *J. Climate*, **26**, 1457–1466, doi:10.1175/JCLI-D-11-00725.1.

Rubel, F., and M. Kottek, 2010: Observed and projected climate shifts 1901–2100 depicted by world maps of the Köppen–Geiger climate classification. *Meteor. Z.*, **19**, 135–141, doi:10.1127/0941-2948/2010/0430.

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. Atmos.*, **118**, 1716–1733, doi:10.1002/jgrd.50203.

Solmon, F., V. S. Nair, and M. Mallet, 2015: Increasing Arabian dust activity and the Indian summer monsoon. *Atmos. Chem. Phys.*, **15**, 8051–8064, doi:10.5194/acp-15-8051-2015.

Sorooshian, S., K. Hsu, X. Gao, H. V. Gupta, B. Imam, and D. Braithwaite, 2000: Evaluation of PERSIANN system satellite–based estimates of tropical rainfall. *Bull. Amer. Meteor. Soc.*, **81**, 2035–2046, doi:10.1175/1520-0477(2000)081<2035:EOPSSE>2.3.CO;2.

Tan, M. L., A. L. Ibrahim, Z. Duan, A. P. Cracknell, and V. Chaplot, 2015: Evaluation of six high-resolution satellite and ground-based precipitation products over Malaysia. *Remote Sens.*, **7**, 1504–1528, doi:10.3390/rs70201504.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, doi:10.1175/BAMS-D-11-00094.1.

van der Wiel, K., and Coauthors, 2016: The resolution dependence of contiguous U.S. precipitation extremes in response to CO_2 forcing. *J. Climate*, **29**, 7991–8012, doi:10.1175/JCLI-D-16-0307.1.

Wuebbles, D. J., and Coauthors, 2014: CMIP5 climate model analyses: Climate extremes in the United States. *Bull. Amer. Meteor. Soc.*, **95**, 571–583, doi:10.1175/BAMS-D-12-00172.1.

Xie, P. P., A. Yatagai, M. Y. Chen, T. Hayasaka, Y. Fukushima, C. M. Liu, and S. Yang, 2007: A gauge-based analysis of daily precipitation over East Asia. *J. Hydrometeor.*, **8**, 607–626, doi:10.1175/JHM583.1.

Yang, X., B. Yong, H. Yong, S. Chen, and X. Zhang, 2016: Error analysis of multi-satellite precipitation estimates with an independent raingauge observation network over a medium-sized humid basin. *Hydrol. Sci. J.*, **61**, 1813–1830, doi:10.1080/02626667.2015.1040020.