An Analysis of Precipitation Variability, Persistence, and Observational Data Uncertainty in the Western United States

KRISTEN J. GUIRGUIS AND RONI AVISSAR

Department of Civil and Environmental Engineering, Duke University, Durham, North Carolina

(Manuscript received 28 August 2007, in final form 16 March 2008)

ABSTRACT

This paper presents an intercomparison of precipitation observations for the western United States. Using nine datasets, the authors provide a comparative climatology and season- and location-specific evaluations of precipitation uncertainty for the western United States and for five subregions that have distinct precipitation climates. All data are shown to represent the general climate features but with high bias among datasets. Interannual variability is similar among datasets with respect to the timing of precipitation excesses and deficits, but important differences occur in the spatial distribution of specific anomalous events. Dataset distribution differences, as represented by their cumulative density functions (CDFs), are statistically significant for 80% of data combinations stratified by subregion and season. The CDFs of anomaly fields are more similar but uncertainty remains, as data differences are significant for 40% of dataset comparisons. Observational uncertainty is low for persistence studies because the data are found to be similar with respect to (i) grid cell estimates of a characteristic persistence time scale and (ii) distributions of anomaly length scales. Spatially, the greatest uncertainty in magnitude differences occurs along the Rocky Mountains in winter, spring, and fall, and along the California coastline in summer. In linear (phase) association, the greatest differences occur in northern Mexico during all seasons; along the Rocky Mountains in winter, spring, and fall; and in California, Nevada, and the intermountain region in summer. Overall, data similarity is lowest in summer as a result of a reduction in phase association and an increase in amplitude differences.

1. Introduction

The western United States is a region marked by limited water resources and a fast-growing population, making it sensitive to variations in the water cycle. Interannual climate variability can alter the amount of water stored as snow in the Western Cordillera, which affects river flow and, consequently, the region's water resources. The frequency and occurrence of extreme precipitation events can be highly variable, and reliable forecasts of these events remain elusive (e.g., Ralph et al. 2005). This is partly a result of weakness in model parameterizations in their ability to approximate complex land-atmosphere dynamics, particularly over complex terrain. However, uncertainty in observations is also a contributing factor. Model evaluation and diagnostics require the use of systematic and high-quality observations, which are logistically difficult to obtain in

DOI: 10.1175/2008JHM972.1

mountainous regions. There are several sources of precipitation data available for the western United States, including estimates from rain gauges, ground radar, satellite, and reanalysis. Each data product contains error that is space and time variant, and it is difficult to know a priori which data product is most reliable and best suited for model evaluation. If a high level of observational data uncertainty exists, then it is possible that the choice of dataset for model evaluation could affect conclusions regarding model skill.

Data reliability depends on issues such as latitude, topography, and seasonality. For example, sampling error associated with rain gauge data can become large over mountains where gauge coverage is sparse. Some data products have aimed to solve this problem by applying statistical methods to rain gauge data to correct for the affect of orography (Daly et al. 1994). However, error associated with these products could be large for those areas where gauge density is particularly low, and error could be introduced from poorly fit regression parameters. Satellite data are available over mountain regions, but precipitation estimates become less reliable poleward of approximately 40° where geostationary satellite measurements are not available and where

Corresponding author address: Roni Avissar, Department of Civil and Environmental Engineering, Edmund T. Pratt School of Engineering, Duke University, P.O. Box 90287, Durham, NC 27708-0287.

E-mail: avissar@duke.edu

precipitation estimates rely on polar-orbiting satellites, which have a poor temporal sampling rate. Additionally, microwave scattering algorithms are problematic over snow- or ice-covered surfaces (Xie and Arkin 1997), which limits their use during cold seasons. Merged precipitation products use information from various sources, including rain gauges and satellites, as a means of rectifying latitudinal variations in sampling error and correcting for bias present in the individual data sources (Huffman et al. 1997; Xie and Arkin 1997; Adler et al. 2003). However, sampling error remains an issue over mountainous regions because these data products use rain gauge data as a major data component. Global and regional reanalysis data are also available for the western United States. However, these precipitation estimates are heavily influenced by model parameterizations (Kalnay et al. 1996; Kistler et al. 2001; Kanamitsu et al. 2002; Mesinger et al. 2006), which do not perform well for all locations and weather regimes.

Precipitation data intercomparison studies aim to provide a measure of the current state of observational data quality and uncertainty by comparing different datasets or precipitation estimation algorithms. Xie and Arkin (1995) compared IR and microwave satellite precipitation estimates with rain gauge data and found general agreement for warm seasons over the tropical Pacific but poorer correspondence over land areas during cold seasons. Similar findings were reported by Ebert et al. (1996), who compared three precipitation estimation algorithms against rain gauge data and found good agreement in the tropical western Pacific and over Japan in summer but poor agreement over Europe in winter. Costa and Foley (1998) compared six precipitation datasets for the Amazon Basin and found general agreement among gauge-based datasets of long-term average climatology but noted important differences with interannual variability and significant bias in comparisons with reanalysis. Janowiak et al. (1998) compared a merged product (Huffman et al. 1997) with reanalysis and found strong large-scale similarity but noted poor agreement for some regional features. Gruber et al. (2000) compared two merged satellite-gauge products (Huffman et al. 1997; Xie and Arkin 1997) and found strong spatial and temporal correlation but also noted significant differences, which they attribute to differences in the use of atoll rain gauge data and aerodynamic gauge corrections. Gottschalck et al. (2005) compared several daily and subdaily datasets for the continental United States and found that, when used to force a land surface hydrology model, observational data differences produced large differences in some prognostic hydrology fields.

This paper contributes to these efforts by providing

an intercomparison for the western United States, using nine state-of-the-art precipitation datasets commonly used in hydrometeorological research. Specifically, we consider the Global Precipitation Climatology Project Combined Precipitation Dataset, version 2 (GPCP); Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP); Global Precipitation Climatology Center monitoring product (GPCC); CPC retrospective United States and Mexico daily precipitation analysis (USMex); Parameter-elevation Regressions on Independent Slopes Model (PRISM); National Centers for Environmental Prediction–Department of Energy Reanalysis 2 (NCEP2); North American Regional Reanalysis (NARR); Variable Infiltration Capacity (VIC) Retrospective Land Surface Dataset (VIC); and the Global Meteorological Forcing Dataset for land surface modeling (GMFD). These data have spatial resolutions ranging from 1/8° to 2.5° but are rescaled to a common grid for this analysis to allow for direct comparison. We focus on data uncertainty, as it pertains to moderately long-term climatology studies on the regional-tocontinental scale. Therefore, we consider monthly 2.5° precipitation fields over a 15-yr period. The research presented here differs from other intercomparison studies; this research includes the bulk of observational precipitation datasets available for moderately longterm studies in the western United States. Additionally, we provide location-specific assessments of uncertainty, using the results of a companion paper (Guirguis and Avissar 2008, hereafter GA08) to focus on specific regions within the domain of the western United States that have been found to have distinct precipitation climates. The precipitation datasets, the rescaling method, and the notation are discussed in section 2. In section 3, we briefly consider the effects of spatial rescaling. The results are presented in sections 4 and 5. Section 4 focuses on general climate features including seasonality, interannual variability, and persistence. Section 5 provides a quantitative assessment of observational data uncertainty. A summary and conclusions are presented in section 6.

2. Data and methods

The time domain for our analysis is January 1986– July 2000, which is the period of maximum overlap between datasets. The results are stratified by season throughout. For this analysis, winter is December– February (DJF), spring is March–May (MAM), summer is June–August (JJA), and fall is September– November (SON). Results are also stratified by location. Specifically, we use the results of GA08, which identifies five unique precipitation climates in the western United States using principal component analysis.



FIG. 1. Composite regionalization for the western United States based on GA08.

In GA08, regionalization results are presented for nine precipitation datasets. In this paper, a composite of those regionalization results is used to subdivide the domain for spatial averaging and discussion. The five unique precipitation regions are referred to as the Pacific Northwest, West Coast, Southwest, Northern Plains, and Colorado Plateau regions. They are mapped in Fig. 1.

a. Datasets

The datasets employed in this study and their spatial/ temporal characteristics are shown in Table 1, along with relevant resources for data documentation. This section provides a short data description, primarily to address issues of dataset independence to highlight those datasets sharing some common sources. A more detailed description of each dataset can be found in GA08.

The GPCC and USMex are rain gauge datasets, of which the GPCC is available globally and the USMex is available over the continental United States and Mexico. These two datasets are likely to have some overlapping stations. However, the GPCC contains only a fraction of the information represented by the USMex (7500 rain gauges globally for the GPCC versus 13 000-15 000 in the United States for the USMex (Higgins et al. 2000; Fuchs et al. 2007). The PRISM and VIC datasets are rain gauge products that use orographic adjustment to account for the affect of elevated terrain on precipitation. The development of the PRISM uses linear regression between gauge measurements and elevation (Daly et al. 1994). The VIC uses an adjustment factor, calculated as the ratio of monthly precipitation from the PRISM to that of rain gauge data (Maurer et al. 2002). The GPCP and CMAP datasets are merged satellite-gauge products. Over land, they use many of the same data sources, including the GPCC, as their rain gauge component, but they differ in their merging methodology (Adler et al. 2003; Xie and Arkin 1997). The NCEP2 and NARR are reanalysis products, which use a data assimilation system to merge observations

TABLE 1. Main characteristics of precipitation datasets used in this study.

Dataset	Spatial resolution	Temporal domain	Data source
GPCC ^a	$2.5^{\circ} \times 2.5^{\circ}$	From 1986	Rain gauge
GPCP ^b	$2.5^{\circ} \times 2.5^{\circ}$	From 1979	Rain gauge, satellite
CMAP ^c	$2.5^{\circ} \times 2.5^{\circ}$	From 1979	Rain gauge, satellite
NCEP2 ^d	$210 \text{ km} \times 210 \text{ km}$	From 1979	Reanalysis
GMFD ^e	$1^{\circ} \times 1^{\circ}$	1948–2000	Reanalysis, rain gauge, satellite
USMex ^f	$1^{\circ} \times 1^{\circ}$	From 1948	Rain gauge
NARR ^g	$32 \text{ km} \times 32 \text{ km}$	From 1979	Reanalysis
VIC ^h	$1/8^{\circ} \times 1/8^{\circ}$	1950–2000	Rain gauge with orographic adjustment
PRISM ⁱ	$4 \text{ km} \times 4 \text{ km}$	From 1890	Rain gauge with orographic adjustment

^a Global Precipitation Climatology Center Monitoring Product (Rudolf and Schneider 2005; Fuchs et al. 2007).

- ^b Global Precipitation Climatology Project Combined Precipitation Dataset, version 2 (Adler et al. 2003).
- ^c CPC Merged Analysis of Precipitation (Xie and Arkin 1997).
- ^d NCEP-DOE Reanalysis 2 (Kanamitsu et al. 2002).
- ^e Global Meteorological Forcing Dataset for land surface modeling (Sheffield et al. 2004).
- ^f CPC retrospective United States and Mexico daily precipitation analysis (Higgins et al. 2000).
- ^g North American Regional Reanalysis (Mesinger et al. 2006).
- ^h VIC Retrospective Land Surface Dataset (Maurer et al. 2002).
- ⁱ Parameter-elevation Regressions on Independent Slopes Model (Daly et al. 1994).

from many sources. Precipitation is a prognostic field in both reanalysis products and is heavily dependent on model parameterizations (Kanamitsu et al. 2002; Mesinger et al. 2006). The GMFD uses data from the NCEP reanalysis as its primary input and applies rain gauge and satellite data to correct for known errors and to downscale the data to a higher spatial and temporal resolution (Sheffield et al. 2006).

b. Rescaling

For this analysis, all datasets are rescaled to a common $2.5^{\circ} \times 2.5^{\circ}$ grid that runs from the Pacific coast to the eastern Rocky Mountains and from northern Mexico to just north of the U.S.–Canadian border (28.75° – 48.75° N and 128.75° – 103.75° W). Box averaging was used to upscale those datasets having spatial resolutions smaller than $2.5^{\circ} \times 2.5^{\circ}$. For datasets having an original horizontal grid spacing of $1^{\circ} \times 1^{\circ}$ or smaller, the upscaling was done iteratively, where at each iteration the horizontal scale was increased by 100% until the desired 2.5° scale was achieved. The iterative method was implemented to prevent the propagation of undefined values occurring over the Pacific Ocean for rain gauge data and to prevent the unrealistic spatial propagation of small-scale, intense precipitation events to the coarser scale. For example, upscaling directly from $1/8^{\circ}$ to 2.5° might cause an isolated precipitation event to be disproportionately represented over a large land area because the mean is sensitive to outliers, whereas upscaling iteratively seemed to eliminate this problem while still maintaining the general features of the original high-resolution data. Finally, bilinear interpolation was employed to adjust the 2.5° data so that all datasets have collocated grid cell centers.

c. Notation

Precipitation anomaly fields are calculated by subtracting the long-term climatic monthly mean from each monthly observation as

$$x'_{m}(i,k) = x_{m}(i,k) - \bar{x}_{m}(i),$$
 (1)

where the prime represents an anomaly field, $x_m(i, k)$ represents monthly precipitation at position *i* and time *k* occurring in month *m*, and $\overline{x}_m(i)$ is the long-term climatology for month *m* at position *i*. Angle brackets denote spatial averages as

$$\langle x \rangle = \frac{1}{p} \sum_{i=1}^{p} x, \qquad (2)$$

where p is the number of 2.5° grid cells in the region being averaged over, which ranges from 10 for the Colorado Plateau region to 18 for the Southwest region. Overbars represent temporal averages as

$$\overline{x} = \frac{1}{t} \sum_{k=1}^{t} x, \qquad (3)$$

where t is the temporal sample size. The standard deviation is represented by S_x and σ_x for spatial and temporal spread, respectively. The correlation coefficient for data pairs x and y is represented by r_{xy} and ρ_{xy} for correlation in the space and time domain, respectively.

It is convenient for some analyses to compare each dataset against a common reference dataset to identify data differences. For this, we use an ensemble dataset as the reference dataset, which is taken as the mean over all datasets as

$$E(i,k) = \frac{1}{N} \sum_{n=1}^{N} x_n(i,k),$$
(4)

where E(i, k) is the ensemble at position *i* and time *k*, $x_n(i, k)$ is the *n*th ensemble member at position *i* and time *k*, and N = 9 is the number of ensemble members.

The PRISM dataset does not include data for northern Mexico. Therefore, for the Southwest region, any analyses involving the PRISM data use only those grid cells located in the United States.

3. Effects of rescaling

Figure 2a shows the spatial distribution of long-term mean precipitation for each dataset at its original spatial resolution and rescaled to a common $2.5^{\circ} \times 2.5^{\circ}$ grid. Considering first the original data in Fig. 2a (left column), all datasets, regardless of spatial scale, represent the general features of the precipitation climate. The Pacific Northwest and northern California are shown to receive much more precipitation than the rest of the domain, and southern California, western Arizona, and northern Mexico are comparatively dry. All datasets represent the enhancement of precipitation along the windward side of the coastal ranges and Cascades in northern California, Oregon, and Washington where Pacific airstreams encounter orographic uplift. The higher-resolution data having spatial scales of 1° or smaller (NARR, USMex, GMFD, VIC, and PRISM) additionally represent orographic precipitation associated with the Sierra Nevada and coastal ranges in central and southern California. These datasets also represent the zone of precipitation enhancement that occurs on the windward side of the Rocky Mountains, which is observed as a roughly linear zone of high precipitation, running from the northwest part of the domain over Idaho to the southeastern part of the domain over Colorado. The very high-resolution data having spatial scales smaller than 1° (NARR, VIC, and PRISM) are further able to show the rain shadow effect that occurs over the San Joaquin Valley, which lies between the coastal ranges and the Sierra Nevada in central California. The effect of rescaling the high- and very highresolution data to a $2.5^{\circ} \times 2.5^{\circ}$ grid (Fig. 2a, right column) is to lose much detail of precipitation spatial variability. However, some important details are transferred to the larger grid. The NARR, USMex, GMFD, VIC, and PRISM still show the enhancement of precipitation over the Rocky Mountains, which is not as clearly represented by other datasets. This orographic signal is muted according to the GMFD as a result of wetter conditions in the northern plains.

4. Dataset intercomparison of climatology

a. Seasonality

Long-term seasonal mean precipitation is given in Fig. 3. Here, all datasets capture the general seasonal features of the western United States. During winter, all data show the eastward wet–dry precipitation gradient associated with orographic uplift of Pacific airstreams along the western coast and the subsequent rain shadow effect to the east. During spring, wetter conditions are observed over the Colorado Plateau and Northern Plains regions, which are associated with the

SON

GPCC

GPCP

СМАР

NCEP2

GMFD

USMe

NARR

PRISM

10

15

5

JJA





mm/month

mm/month

FIG. 2. (a) Spatial distribution of long-term mean precipitation for the datasets at (left) their original resolution and (right) rescaled to $2.5^{\circ} \times 2.5^{\circ}$. (b) Bias as measured against the ensemble reference dataset for winter, spring, summer, and fall.



FIG. 3. Long-term seasonal precipitation for (left to right) winter, spring, summer, and fall.

advection of Gulf of Mexico moisture into the deep continental interior, whereas drier conditions are seen along the northern coastal regions. During summer, the West Coast and Pacific Northwest regions are shown to be at the peak of their dry season, whereas precipitation increases occur in the Southwest region, Northern Plains, and eastern Colorado, which are associated with the North American monsoon (NAM) and Great Plains low-level jet. During fall, the domain transitions from the warm- to cold-season precipitation regime and, therefore, fall exhibits many of the same characteristics as winter but to a lesser degree.

Figure 2b gives the seasonal bias β for each dataset *n*

calculated against the ensemble reference dataset *E*. Specifically, Fig. 2b shows

$$\beta_n(i) = \frac{1}{t} \sum_{k=1}^{t} [x_n(i,k) - E(i,k)].$$
 (5)

The overall mean precipitation $\langle \bar{x}_n \rangle$ is given in Table 2, and the mean bias $\langle \beta_n \rangle$ is given in Table 3. In Figs. 2b and 3 and Tables 2 and 3, the GPCP, CMAP, and GPCC data show drier conditions relative to other datasets for all seasons and regions. Conversely, the GMFD, VIC, and PRISM generally show wetter conditions. Spatially, dataset differences are greatest along the western coast and along the path of the Rocky



FIG. 3. (Continued)

Mountains. This may be related to the original spatial scale of the datasets, owing to the ability of the higher resolution data to represent orographic precipitation enhancement over the Sierra Nevada in California and over the western flanks of the Rocky Mountains (cf. Fig. 2a). However, sampling error associated with poor gauge coverage by the GPCC dataset (also used in the CMAP and GPCP) may also be a factor. The orographic adjustment employed in the development of the VIC and PRISM would contribute to the higher precipitation estimates by these datasets over these parts. Also notable in Figs. 2b and 3 are differences between the GPCP and CMAP over eastern Oregon, Nevada, and the northern plains, especially in winter, where larger precipitation amounts are observed for the GPCP compared to the CMAP. This is likely a result of the use of systematic gauge corrections in the development of the GPCP data, whereas the CMAP gauge data are uncorrected, and most systematic errors occur as undercatch (Fuchs et al. 2007). The GMFD data and especially the NCEP2 data show wetter conditions over the Northern Plains in spring and summer relative to the other data, which may be related to convection parameterizations used in the derivation of NCEP2 precipitation fields. The GMFD data show wetter conditions over the Colorado Plateau region and the state of Nevada in winter than is observed by any other dataset. During summer, the GMFD, VIC, USMex, and especially NCEP2 data show a stronger monsoon signal in the Southwest region, whereas the NAM signal is

VOLUME	9
--------	---

		GPCC	GPCP	CMAP	NCEP2	GMFD	USMex	NARR	VIC	PRISM
Domain	DJF	34.1	39.2	34.0	43.7	54.6	45.5	46.3	57.3	58.8
	MAM	31.2	35.3	31.0	46.3	44.4	37.1	38.8	45.3	46.4
	JJA	28.9	31.7	28.4	45.8	40.2	34.1	32.2	38.9	35.5
	SON	27.3	29.7	26.8	34.9	38.7	32.9	33.8	40.6	40.6
Pacific Northwest	DJF	78.5	87.0	78.4	99.2	115.3	101.0	115.5	134.1	134.8
	MAM	54.9	58.6	54.6	79.3	75.9	65.9	77.4	87.0	84.9
	JJA	23.9	26.3	23.8	42.9	37.2	26.7	28.7	36.6	34.4
	SON	53.1	56.0	52.1	65.5	75.2	64.3	73.4	86.5	84.8
West Coast	DJF	43.9	48.1	44.2	50.5	70.8	62.2	53.8	74.4	71.4
	MAM	24.0	28.4	24.0	33.1	39.4	33.7	29.5	38.7	37.9
	JJA	6.9	8.3	6.7	11.3	12.6	8.7	6.7	9.5	9.7
	SON	15.0	17.4	14.9	18.4	25.5	19.8	17.0	22.9	22.8
Southwest	DJF	15.6	17.0	15.4	23.0	23.6	21.0	17.0	22.1	22.2
	MAM	13.3	15.5	12.9	23.9	16.7	15.5	12.0	16.9	19.8
	JJA	39.7	42.6	38.6	56.7	55.1	51.0	44.4	56.2	48.3
	SON	21.5	22.5	21.1	32.2	27.3	25.8	22.0	28.4	27.4
Northern Plains	DJF	12.9	18.3	12.7	19.7	25.8	16.6	18.3	22.9	22.8
	MAM	36.3	42.4	36.0	51.8	52.2	38.7	42.5	46.4	45.0
	JJA	45.3	51.2	45.1	73.8	55.9	48.1	47.9	53.7	52.0
	SON	23.5	27.1	23.2	28.9	32.0	26.0	28.1	31.6	31.5
Colorado Plateau	DJF	17.0	24.3	16.5	22.2	37.4	23.6	23.5	29.6	28.9
	MAM	30.7	35.7	30.9	47.7	43.0	35.2	37.2	41.4	39.9
	JJA	24.8	25.9	24.2	39.7	34.7	30.7	29.0	31.9	32.0
	SON	21.4	23.9	20.9	25.8	32.8	26.8	27.3	31.0	30.3

TABLE 2. Long-term mean precipitation (mm month $^{-1}$).

much more subtle in the CMAP and GPCC data (Fig. 3). For the NCEP2, the spatial extent of the monsoon signal is restricted westward, as observed by comparatively dry conditions over the western part of northern

Mexico and extending northward into Arizona and Utah (Fig. 2b). Another important difference observed in Figs. 2 and 3 is that, whereas (Fig. 3) the NARR data generally show a similar spatial pattern of precipitation

		GPCC	GPCP	CMAP	NCEP2	GMFD	USMex	NARR	VIC	PRISM
Domain	DJF	-11.3	-6.2	-11.4	-1.7	9.2	0.1	0.8	11.9	11.7
	MAM	-8.0	-3.9	-8.2	7.1	5.2	-2.1	-0.3	6.1	5.3
	JJA	-6.5	-3.6	-6.9	10.4	4.9	-1.3	-3.2	3.6	1.3
	SON	-6.4	-3.9	-6.8	1.3	5.1	-0.7	0.2	7.0	6.3
Pacific Northwest	DJF	-24.3	-15.8	-24.4	-3.6	12.5	-1.8	12.8	31.3	32.0
	MAM	-14.8	-11.1	-15.1	9.6	6.2	-3.8	7.6	17.3	15.2
	JJA	-7.3	-5.0	-7.4	11.6	6.0	-4.6	-2.5	5.3	3.2
	SON	-13.5	-10.6	-14.5	-1.1	8.5	-2.3	6.7	19.9	18.2
West Coast	DJF	-13.4	-9.1	-13.1	-6.8	13.5	4.9	-3.5	17.1	14.1
	MAM	-8.1	-3.8	-8.1	0.9	7.2	1.5	-2.6	6.6	5.7
	JJA	-2.4	-1.0	-2.6	2.1	3.3	-0.6	-2.6	0.2	0.5
	SON	-4.6	-2.2	-4.6	-1.2	5.9	0.2	-2.6	3.3	3.2
Southwest	DJF	-4.1	-2.7	-4.3	3.3	3.9	1.3	-2.7	2.4	2.0
	MAM	-2.6	-0.4	-3.0	8.0	0.8	-0.4	-3.9	1.0	1.0
	JJA	-8.5	-5.5	-9.5	8.5	7.0	2.9	-3.7	8.1	2.4
	SON	-3.7	-2.7	-4.0	7.1	2.2	0.6	-3.2	3.2	1.7
Northern Plains	DJF	-6.3	-0.9	-6.5	0.5	6.6	-2.5	-0.9	3.7	3.6
	MAM	-7.8	-1.7	-8.1	7.6	8.0	-5.4	-1.6	2.3	0.9
	JJA	-7.5	-1.6	-7.7	21.0	3.1	-4.7	-4.9	0.9	-0.8
	SON	-4.5	-0.9	-4.8	0.9	3.9	-2.0	0.0	3.6	3.5
Colorado Plateau	DJF	-8.6	-1.4	-9.1	-3.5	11.7	-2.1	-2.1	4.0	3.2
	MAM	-7.4	-2.4	-7.2	9.5	4.9	-3.0	-1.0	3.3	1.8
	JJA	-5.8	-4.7	-6.5	9.1	4.0	0.0	-1.6	1.2	1.4
	SON	-5.5	-3.0	-6.0	-1.1	5.9	0.0	0.4	4.1	3.5

TABLE 3. Same as Table 2 but for the mean bias as measured against the reference dataset (mm month $^{-1}$).



FIG. 4. (a) Ensemble mean, (b) standard deviation, and (c) coefficient of variation.

as do other high-resolution data, much drier conditions are seen along the California coast during winter and to a lesser extent in spring and fall (Fig. 2b). This is also observed for the NCEP2, suggesting that model parameterizations used in the data assimilation process may be an issue here for both reanalysis products.

To estimate overall bias among datasets and its spatial distribution, we consider the spread among ensemble members (datasets) and the corresponding coefficient of variation (as in Yang and Arritt 2002). The spread is represented by the ensemble standard deviation as

$$S_{E}(i,k) = \sqrt{\frac{1}{N} \sum_{n=1}^{N} \left[x_{n}(i,k) - E(i,k) \right]^{2}}, \qquad (6)$$

and the corresponding coefficient of variation is given by

$$CV(i, k) = \sigma_E(i, k)/E(i, k).$$
(7)

Figure 4 gives temporal averages of the ensemble mean, spread, and coefficient of variation. The ensemble spread (Fig. 4b) is shown to be large for the Pacific Northwest during spring, fall, and particularly winter; the Rocky Mountains in winter, spring and summer; coastal California during winter; and the Great Plains and Southwest in summer. In Fig. 4c, the coefficient of variation is large for dry regions where the precipitation frequency distributions are heavily skewed toward zero, and the mean precipitation [denominator in (7)] is small. The spread among datasets for California reaches up to 55% of the mean during winter, 90% during spring and fall, and up to 137% for some coastal areas during summer. This implies that data uncertainty for the West Coast region can be larger than the mean precipitation received there. This is similarly observed for the Southwest region but to a lesser degree. Here, the spread can reach 93% of the mean during spring and 75% during winter, summer, and fall. The spread for the Pacific Northwest, whereas generally larger in magnitude (Fig. 4b), represents a smaller ratio of the mean (Fig. 4c), as here the coefficient of variation ranges from 15%-70% depending on season and location. Smaller spreads are observed for the Northern Plains and Colorado Plateau regions. For the Northern Plains, the ensemble standard deviation ranges from 5 to 19 mm month $^{-1}$, which corresponds to a coefficient of variation of 22%-58%. The spread for the Colorado Plateau region ranges from 5 to 21 mm month⁻¹, which represents 21%–51% of the mean.

Empirical probability densities of precipitation values are shown in Fig. 5, stratified by season and region. Each data vector represented in Fig. 5 is of length $t \times p$,



FIG. 5. Empirical cumulative density functions for (a) winter, (b) spring, (c) summer, and (d) fall.

where *t* is the temporal sample size (t = 42-45 depending on season), and *p* is the number of spatial grid cells (p = 10-18 depending on region). We apply the twosample Kolmogorov–Smirnov (K–S) test to compare the distributions of precipitation between each pair of datasets. The null hypothesis for the K–S test is that data vectors **x** and **y** are drawn from the same continuous distribution, against the alternative that they are from different distributions. The test statistic is $\max |F_x(j) - F_y(j)|$, where F_x and F_y are the empirical cumulative density functions (CDFs) from sample vectors **x** and **y**. The null hypothesis is rejected if the test statistic exceeds a critical value.

Figure 6 gives the results of the K–S test. The shaded parts indicate that differences between distributions are significant at the 95% level. Here, the null hypothesis that **x** and **y** are drawn from the same population is rejected more often than not, implying a significant degree of data uncertainty. Specifically, the null is rejected 579 times out of $5 \times 4\sum_{n=1}^{N-1} n = 720$ comparisons, where 5 is the number of regions, 4 is the number of seasons, and *n* is the number of datasets. Figure 7 gives the results of the K–S test for anomaly fields in which the climatology has been removed, as in (1). Here, the null hypothesis is rejected less frequently, implying that seasonal anomaly distributions are more similar. However, uncertainty remains an issue, as significant differences are observed for 39% of the dataset comparisons (283 out of 720). In Fig. 7, the greatest similarity is observed during summer and fall for the Southwest, Northern Plains, and Colorado Plateau regions and in spring for the Northern Plains.

b. Interannual variability

Figure 8a shows spatially averaged annual anomaly fields in which December–November is used as the averaging period to maintain seasonal congruence (e.g., calendar year 1990 represents the period December 1989–November 1990). Also shown in Fig. 8a is the phase and strength of El Niño–Southern Oscillation (ENSO) according to the monthly multivariate ENSO index. There is strong agreement between data sources with respect to interannual peaks and troughs. All data show a Pacific Northwest dry period between 1986 and 1994, after which time the region entered a wet phase that extended through 1999. Extremes occurring in the



FIG. 6. Results of the two-sample K–S test for (a) winter, (b) spring, (c) summer, and (d) fall. Shading indicates that the difference between the distributions of datasets x and y is significant at the 95% level.

Pacific Northwest are the negative anomalies in 1987 and 1992 coinciding with warm phases of ENSO and the positive anomalies during 1995–97, which occurred during a cold ENSO period. The West Coast region shows a similar pattern of dry conditions during the first part of the record and wetter conditions later. Strong positive anomalies are observed for the West Coast region in 1993, 1995, and 1998, all of which coincide with the warm phase of ENSO. The Southwest region shows an opposite anomaly pattern as compared to the Pacific Northwest and West Coast, with wet conditions dominating prior to 1994 and drier conditions observed thereafter. Precipitation extremes for the Southwest include the relatively dry 1989 and the wet 1992, which occur during cold and warm phases of ENSO, respectively. The Northern Plains and Colorado Plateau regions demonstrate some evidence of a linear positive trend, with regularly spaced cycles of peaks and troughs that do not deviate dramatically from normal. The observed peaks and troughs appear to correspond fairly consistently with warm and cold phases of ENSO, respectively.

The data series for seasonal anomalies are shown in Fig. 9. Strong similarity is observed among data for all regions and seasons with respect to interannual highs and lows. The NCEP2 data are shown to exaggerate some anomalies relative to the other datasets, especially in summer. The difference between data in their estimation of $\langle x' \rangle$ in Fig. 9 (measured on the y axis as the difference between two time series) ranges from 2.3 to 35.8 mm month⁻¹ for the Pacific Northwest, 1.0–25.7 mm month⁻¹ for the West Coast, 2.8–57.6 mm month⁻¹ for the Southwest, $1.8-58.0 \text{ mm month}^{-1}$ for the Northern Plains, and 0.9–39.1 mm month⁻¹ for the Colorado Plateau region. For the Southwest, Northern Plains, and Colorado Plateau regions, the maximum data difference drops to 26.4, 10.5, 14.5 mm month $^{-1}$, respectively, if the NCEP2 data are not considered. Figure 10



FIG. 7. As in Fig. 6 but for precipitation anomaly fields x' and y'.

shows the seasonal time series in which the climatology has not been removed. Here, the difference between data in their estimation of $\langle x \rangle$ ranges from 10.8 to 92.3 mm month⁻¹ for the Pacific Northwest, 2.7–55.6 mm month⁻¹ for the West Coast, 4.0–73.4 mm month⁻¹ for the Southwest, 6.5–83.7 mm month⁻¹ for the Northern Plains and 6.0–50.1 mm month⁻¹ for the Colorado Plateau region. If the NCEP2 data are not considered, the upper limit is reduced to 34.1, 21.5, and 29.2 mm month⁻¹ for the Southwest, Northern Plains and Colorado Plateau regions, respectively.

The winters of 1996/97 and 1997/98 are known for being particularly eventful for the western United States, and we consider them in more detail. The 1996/97 winter season brought heavy and extensive flooding to the Pacific Northwest, Nevada, and California. The Sierra Nevada region was especially affected, and December precipitation in parts of Idaho was recorded at more than 300% of normal (Lott et al. 1997). The 1997/98 winter season saw record-breaking warm and wet conditions across the United States. The western United States was particularly hard hit during February, which brought 4 weeks of near-continuous storm activity to California; and parts of the Southwest and Northern Plains regions were also affected (Ross et al. 1998). In Fig. 9, the 1996/97 winter season is wetter for the Pacific Northwest, Northern Plains, and Colorado Plateau, and the 1997/98 winter season is wetter for the West Coast and Southwest. Also in Fig. 9, the spread among data is relatively large for the 1996/97 winter season (6.1–33.0 mm month⁻¹), whereas a smaller spread is observed for the 1997/98 winter season (2.4–18.0 mm month⁻¹).

The spatial distributions of the 1996/97 and 1997/98 winter anomalies according to each dataset are shown in Fig. 8b. All datasets show the 1996/97 winter season as having widespread wet conditions over the Pacific Northwest, western Montana, Nevada, and California. However, there are some important discrepancies between data sources. For example, eastern Montana, Wyoming, and Colorado are shown to be mildly dry according to the GPCP, CMAP, and GPCC data, whereas the other datasets suggest wet conditions.



% Difference from normal

FIG. 8. (a) Annual precipitation anomaly fields and the phase and strength of ENSO according to the monthly multivariate ENSO index. The warm (cold) phases of ENSO are represented by red and upward (blue and downward) triangular markers. Markers are centered at the midpoint of a warm or cold period and their size is weighted by the magnitude of the ENSO anomaly. (b) Spatial distribution of precipitation anomalies for the winters of 1996/97 and 1997/98. The anomalies were calculated by subtracting the long-term DJF mean.



FIG. 9. Spatially averaged seasonal precipitation anomaly fields. The vertical dashed lines in the DJF time series correspond to the anomaly fields shown in Fig. 8b.

There is also disagreement over the severity of the anomaly over northern Nevada and Idaho during 1996/97. The NARR, USMex, and VIC data suggest dramatically wet conditions over Idaho as compared to the other data, and the GPCC, GPCP, and CMAP data show drier conditions over northern Nevada. For 1997/98, all the datasets show anomalously wet conditions over California. However, the magnitude of the anomaly is highly varied. The CMAP and GPCC data show extremely widespread and wet conditions covering the entire state of California and western Nevada, whereas the PRISM, VIC, USMex, and NARR data show intensely wet conditions only over the southern and coastal parts of California.

c. Persistence

For our analysis of persistence, we calculate the onemonth lag autocorrelation coefficient and estimate the characteristic time scale of persistence. The autocorrelation is calculated as

$$\alpha(\tau) = \sum_{k=1}^{t-\tau} (x'_k - \bar{x})(x'_{k+\tau} - \bar{x})/\sigma_{x'}^2, \tag{8}$$

where τ is the lag in months, x' is the precipitation anomaly calculated as in (1), t is the number of months in the time series, and $\sigma_{x'}$ is the standard deviation of x'. The data are detrended prior to the analysis to remove any linear tendency in the data series.

The characteristic time scale of persistence is estimated using a count of consecutive, like-signed monthly anomalies as in Liu and Avissar (1999). A run of length *l* represents the period after an anomaly is observed in which x' maintains the same sign. Beginning with the first anomaly $x'_{k=k_1}$, we count the number of consecutive months in which x' maintains the same sign, and this value is denoted l_1 . The next run (which is



FIG. 10. As in Fig. 9 but without removing the climatology.

opposite in sign) begins with $x'_{k=k_2}$, and the length of the second run is denoted l_2 . This process repeats until the end of the time series; however, the first and last runs are omitted to avoid truncation at the end points. An anomaly that maintains its sign for only one month demonstrates no persistence and, therefore, l = 0. The characteristic time scale of persistence is estimated as the average length scale according to

$$L = \frac{1}{\eta} \sum_{k=1}^{n} l_k, \tag{9}$$

where η is the number of persistent runs in the time series. We also consider the probability distribution of l_k for each dataset and region using the Kolmogorov– Smirnov test to determine the degree of similarity between datasets in their representation of persistence.

The one-month lag autocorrelation coefficients are shown in Fig. 11a. The coefficients range from -0.08 to 0.41 overall and are shown to vary geographically and

by dataset. Grid cells with statistically significant correlations are shown in Fig. 11b. The correlations are statistically significant at the 95% level (critical value of 0.148 for 173 degrees of freedom) only in the intermountain region, northern Mexico, and southern California. The VIC, USMex, and NARR data give larger coefficients over northern Mexico relative to the other data. The NCEP2 data show larger coefficients for grid cells spanning Idaho, western Montana, and Wyoming, which may be related to moisture recycling efficiency in the spectral model. Figure 11c shows the estimated characteristic persistence time scale. The time scale Lranges from 0.68 to 2.2 months overall, with most of the domain showing persistence on the order of 1–1.5 months. Longer time scales of 1.5-2 months occur for parts of California and northern Mexico. The USMex and NCEP2 data show longer time scales for northern Mexico, and the PRISM, VIC, and USMex data demonstrate longer time scales over a larger portion of California.



FIG. 11. (a) One-month lag autocorrelation coefficient, (b) correlations significant at the 95% level shown as shaded, and (c) characteristic persistence time scale in months.



FIG. 12. The distribution of persistence length scales for each precipitation region. There is overlap of outliers such that each plus sign (+) represents more than one length scale occurrence.

Figure 12 shows the distribution of persistence length scales for each precipitation region. Here, the ensemble reference dataset is used as the precipitation data, from which we calculate l_1, l_2, \ldots, l_n for each grid cell, and these persistence length scale time series are then aggregated over each precipitation region to give the distributions shown in Fig. 12. The median length scale for the Pacific Northwest, Southwest, and Colorado Plateau regions is one month, and the median length scale is zero for the West Coast and Northern Plains. The spread as represented by the interquartile range (IQR) is two months, in which the 25th percentile is zero, and the 75th percentile is two months for all regions. Outliers are shown as those larger than $1.5 \times IQR$, or $l_k >$ 5 months. The majority of outliers correspond to negative anomalies (not shown), suggesting that droughts are more likely to persist for long periods than for periods of precipitation excess. The proportion of positive-to-negative anomalies exceeding five months is (displayed as positive:negative) 12:10, 12:26, 23:53, 6:21, and 6:17 for the Pacific Northwest, West Coast, Southwest, Northern Plains, and Colorado Plateau regions, respectively.

The degree of similarity between datasets in their representation of persistence is determined by compiling length scale distributions as described above for each of the nine datasets. Then we apply the twosample K–S test to compare the length scale distributions among datasets. The null hypothesis that two samples comprised of l_k from two different precipitation datasets are drawn from the same distribution is rejected only 4 times out of $5 \times \sum_{n=1}^{N-1} n = 180$ comparisons, where 5 is the number of regions and N is the number of datasets. Statistically significant differences occur for comparisons of the GPCP data against the USMex, NARR, and VIC data for the West Coast region, and between the PRISM and NCEP2 data for the Northern Plains region. For the West Coast, there are slight differences between the cumulative densities of l_k for the GPCC, GPCP, and CMAP data as compared to the USMex, NARR, VIC, and PRISM data. Specifically, the GPCC, GPCP, and CMAP data show higher densities for l_k in the range of 0–2 months and lower densities for l_k in the range of 2–4 months relative to the USMex, NARR, VIC, and PRISM data; this difference is large enough to achieve statistical significance for comparisons between the GPCP data with the USMex, NARR, and VIC data. For the Northern Plains, the NCEP2 data show higher densities for l_k in the range of 5–10 months and lower densities for l_k in the range of 0-2 months; this difference is large enough to reach significance for the comparison between the NCEP2 and PRISM data.

5. Observational data uncertainty

a. Spatial distribution of uncertainty

Figure 13 is used to assess the spatial distribution of observational data uncertainty. For each grid cell, the correlation coefficient and normalized root-meansquare error (NRMSE) are calculated, in which the NRMSE normalization factor is the standard deviation. These quantities are calculated according to

$$\rho_{xy}(i) = \sum_{k=1}^{t} (x - \overline{x})(y - \overline{y})/\sigma_x \sigma_y, \text{ and}$$
(10)
$$\text{NRMSE}_{xy}(i) = \frac{1}{\sigma} \left\{ \frac{1}{t} \sum_{k=1}^{t} [x(i,k) - y(i,k)]^2 \right\}^{1/2},$$
(11)

respectively, where σ in (11) is the average of σ_x and σ_y . The quantities (10) and (11) are calculated for each dataset pair, and the average of these quantities over the 36 dataset combinations is taken as an estimate of data uncertainty for each grid cell. Specifically, in Fig. 13 we show

$$\rho(i) = \frac{1}{N} \sum_{n=1}^{N} \rho_n(i), \text{ and }$$
(12)

$$NRMSE(i) = \frac{1}{N} \sum_{n=1}^{N} NRMSE_n(i), \qquad (13)$$

where N = 36 is the number of dataset combinations.



FIG. 13. (a) Average correlation coefficient and (b) NRMSE between datasets.

The correlation coefficient provides a measure of the phase association (ρ), or the phase error $(1 - \rho)$ between datasets. In Fig. 13a, the average dataset correlation is in the range 0.34-0.98 for the domain as a whole over all seasons. The correlation is lowest over the Rocky Mountains in winter, spring, and fall; over northern Mexico for all seasons; and over southern California and the intermountain region in summer. Table 4 gives the range of values of $\rho(i)$ for each season and region. The correlation across seasons is in the range of 0.78-0.98, 0.34-0.97, 0.43-0.95, 0.63-0.93, and 0.63-0.93 for the Pacific Northwest, West Coast, Southwest, Northern Plains, and Colorado Plateau regions, respectively. The lowest correlations occur in spring for the Pacific Northwest, in summer for the West Coast, in fall for the Southwest, and in winter for the Northern Plains and Colorado Plateau regions.

The NRMSE provides a measure of amplitude differences between datasets. In Fig. 13b, the average NRMSE is in the range of 0.28–1.62 for the domain as a whole over all seasons. The NRMSE is generally high over the Rocky Mountains, with the largest values observed in these parts in winter and the smallest values observed in summer. Large amplitude differences are also observed in coastal California in summer and in northern Mexico during all seasons. Values of NRMSE(i) across seasons are in the range of 0.30–1.37, 0.28-1.37, 0.37-1.11, 0.44-1.36, and 0.47-1.62 for the Pacific Northwest, West Coast, Southwest, Northern Plains, and Colorado Plateau regions, respectively. In Table 4, the largest errors are observed in winter for the Pacific Northwest, Northern Plains, and Colorado Plateau regions, whereas the largest errors are observed in summer for the West Coast and Southwest regions.

b. Decomposition of the mean squared error

To give an overall indication of the degree of agreement among datasets, we use the mean square error (MSE) and apply the decomposition of Murphy (1988). This methodology is commonly used in quantitative forecast verification to compare forecasted and observed fields to assess forecast skill. Here, we use the method to compare observations from one dataset

TABLE 4. Range of values by season and precipitation region for the correlation coefficient (ρ) and RMSE presented in Fig. 13.

		ρ	RMSE
Domain	DJF	0.59-0.96	0.41-1.62
	MAM	0.50-0.97	0.43-1.13
	JJA	0.34-0.94	0.50-1.37
	SON	0.43-0.98	0.28-1.11
Pacific Northwest	DJF	0.85-0.96	0.43-1.37
	MAM	0.78-0.97	0.47-1.13
	JJA	0.83-0.94	0.55-0.88
	SON	0.83-0.98	0.30-0.97
West Coast	DJF	0.75-0.96	0.42-0.87
	MAM	0.83-0.97	0.44-0.86
	JJA	0.34-0.93	0.55-1.37
	SON	0.53-0.97	0.28-1.07
Southwest	DJF	0.59-0.95	0.41-0.97
	MAM	0.50-0.95	0.43-1.06
	JJA	0.61-0.89	0.54 - 1.07
	SON	0.43-0.94	0.37-1.11
Northern Plains	DJF	0.63-0.89	0.81-1.36
	MAM	0.82-0.93	0.46 - 1.07
	JJA	0.83-0.93	0.50-0.71
	SON	0.81-0.92	0.44 - 1.00
Colorado Plateau	DJF	0.63-0.93	0.56-1.62
	MAM	0.77-0.88	0.53-0.90
	JJA	0.74-0.90	0.62-0.95
	SON	0.80-0.91	0.47-0.99



FIG. 14. NMSE, the terms of its decomposition, and the ratio of standard deviations for the domain of the western United States. The datasets listed on the y axis (x axis) represent x' (y') in Eq. (16).

against observations of another dataset. The MSE operating on anomaly fields is represented as

$$MSE_{x'y'} = \frac{1}{m} \sum_{i=1}^{m} (x'_i - y'_i)^2, \qquad (14)$$

where x' and y' are calculated as in (1) for two sets of observations (datasets), and m is the number of paired observations for a given 2D field (map). The MSE can be decomposed according to Murphy (1988) as

$$MSE_{x'y'} = (\langle x' \rangle - \langle y' \rangle)^2 + S_{x'}^2 + S_{y'}^2 - 2S_{x'}S_{y'}r_{x'y'},$$
(15)

where the angle brackets represent spatial averages, $S_{x'}$ and $S_{y'}$ represent the spatial standard deviations of x'and y', and $r'_{x'y'}$ represents the spatial anomaly correlation coefficient. Equation (15) can be manipulated algebraically by dividing by $S_{y'}^2$ and completing the square to give the normalized MSE (NMSE), represented as

$$\text{NMSE}_{x'y'} = \left(\underbrace{\frac{\langle x' \rangle - \langle y' \rangle}{S_{y'}}}_{A^2}\right)^2 + \left(\underbrace{r_{x'y'} - \frac{S_{x'}}{S_{y'}}}_{B^2}\right)^2 + \underbrace{(1 - r_{x'y'}^2)}_{C^2},$$
(16)

where A^2 is a nondimensional measure of the unconditional bias, B^2 is a measure of the conditional bias, C^2 is the phase error, and NMSE is the normalized MSE, where NMSE = $MSE/S_{y'}^2$. For analyses of fields (maps), the terms in (16) can be physically interpreted as representing the contribution to NMSE as a result of phase error (C^2), amplitude differences (B^2), and error as a result of map mean differences (A^2 ; based on Livezey et al. 1995).

Figure 14 provides the NMSE (averaged spatially and temporally) and the terms of its decomposition for all data combinations, and Table 5 gives the median value taken over all data pairs. Also shown in Fig. 14 is the term $S_{x'}/S_{y'}$ to describe how standard deviation differences might affect the NMSE and terms A^2 and B^2 .

An examination of Fig. 14 and Table 5 shows that greater similarity between datasets is observed in winter, spring, and fall as compared to summer, with median NMSE values of 0.42, 0.42, 0.41, and 0.70, respectively. This summer reduction in similarity is a result of a reduction in phase association and an increase in amplitude differences (Fig. 14 and Table 5). The clustering of low NMSE values in the southwest and northeast quadrants of Fig. 14 highlights two groups of "similar" data: Group 1 are GPCC, GPCP, and CMAP data, and Group 2 are USMex, NARR, VIC, and PRISM. These datasets generally show strong phase association and low bias within groups, and notably lesser phase association and higher bias between groups. The GPCC, GPCP, and CMAP data are shown to be very similar,

		NMSE	A^2	B^2	C^2
Domain	DJF	0.42	0.05	0.08	0.27
	MAM	0.42	0.05	0.07	0.29
	JJA	0.70	0.06	0.15	0.45
	SON	0.41	0.04	0.08	0.29
Pacific Northwest	DJF	0.71	0.22	0.20	0.27
	MAM	0.73	0.18	0.16	0.32
	JJA	0.85	0.21	0.18	0.41
	SON	0.66	0.18	0.16	0.29
West Coast	DJF	0.85	0.25	0.26	0.34
	MAM	1.09	0.31	0.34	0.39
	JJA	1.22	0.22	0.42	0.49
	SON	0.87	0.22	0.22	0.38
Southwest	DJF	1.19	0.38	0.33	0.50
	MAM	0.85	0.17	0.24	0.41
	JJA	1.53	0.43	0.43	0.60
	SON	1.04	0.22	0.34	0.51
Northern Plains	DJF	1.46	0.48	0.46	0.53
	MAM	0.86	0.20	0.24	0.40
	JJA	0.71	0.14	0.16	0.41
	SON	0.95	0.26	0.30	0.43
Colorado Plateau	DJF	1.45	0.65	0.27	0.46
	MAM	1.43	0.62	0.32	0.49
	JJA	1.23	0.35	0.37	0.48
	SON	1.03	0.43	0.20	0.40

TABLE 5. Median value of the NMSE and its decomposition $(A^2, B^2, \text{ and } C^2)$ taken over all data pairs.

with NMSE less than 0.22 for all seasons. The conditional and unconditional bias between the GPCC. GPCP, and CMAP data is small (A^2 , $B^2 < 0.03$) and phase association is strong $(1 - C^2 = r_{x'y'}^2 = 0.82 - C^2$ 0.97). Strong similarity is also observed between the VIC and PRISM data, with NMSE less than 0.13 in winter, spring, and fall, and NMSE = 0.22 in summer. The VIC and PRISM data demonstrate very little conditional or unconditional bias $(A^2, B^2 < 0.04)$ and strong phase association $(r_{x'y'}^2 = 0.82 - 0.92)$. The USMex and NARR data show relatively strong similarity between each other and with the VIC and PRISM data during winter, spring and fall, with NMSE ranging from 0.12 to 0.41, squared anomaly correlation coefficients ranging from 0.80 to 0.91, and conditional (unconditional) bias ranging from 0.01 to 0.21 (0.01-0.07). In summer, comparisons of the USMex and NARR data against the VIC data show a reduced similarity (NMSE = 0.58-0.73), which is primarily a result of an increase in phase error ($C^2 = 0.38$). The PRISM and VIC data demonstrate a minimum similarity with Group 1 during winter (NMSE = 0.38-1.20) as a result of increased conditional ($B^2 = 0.06-0.69$) and unconditional ($A^2 = 0.06-0.23$) bias. The NCEP2 data are shown to be generally dissimilar to the other data during all seasons except winter, and it is especially dis similar in summer. In Fig. 14, the NCEP2 data are shown to exhibit lower phase association and higher conditional and unconditional bias than is observed for comparisons between other datasets. The GMFD data generally show a greater similarity with Group 2 than with Group 1 (mean NMSE = 0.47 and 0.58, respectively).

Table 6 provides seasonal averages of the spatial standard deviation for each dataset, which affects the NMSE through the term $S_{x'}/S_{y'}$. Standard deviation differences between datasets are shown in Fig. 14 to contribute to the bias terms A^2 and B^2 and to the NMSE in winter, spring, and fall. During summer, the term $S_{x'}/S_{y'}$ is near one for all data pairs except the NCEP2, which has a much larger spread than the other data (also in Table 3). Aside from the NCEP2, the standard deviation differences are greatest between Group 1 and Group 2 datasets, where for Group 1 $S_{x'} = 16.4-23.3$ and for Group 2 $S_{x'} = 15.8-36.7 \text{ mm day}^{-1}$. In Table 6, the standard deviations of the GMFD and NCEP2 data generally occur in the midrange of these two groups except during summer, when the NCEP2 spread is large, as discussed above.

Figure 15 and Table 5 give the median NMSE and decomposition terms for each season and region. The NMSE is shown to be largest in the Southwest region in summer, and the largest in the Northern Plains and Colorado Plateau regions in winter (median NMSE = 1.53, 1.46, and 1.45, respectively). The summer reduction in dataset similarity in the Southwest is a result of increases of all error components. For the Northern Plains in winter, the increase in NMSE is a result of increases in all error terms, but conditional and unconditional bias rise more sharply in winter than phase error. For the Colorado Plateau in winter, the higher NMSE is a result of an increase in unconditional bias. For the Pacific Northwest region, the largest NMSE is observed in summer, which is primarily a result of a reduction in phase association because the bias terms change very little across seasons. For the West Coast region, the NMSE peaks in summer, which is a result of increases in conditional bias and phase error.

6. Summary and conclusions

A comparison of precipitation between nine data products revealed two groups of similar datasets. The GPCP, CMAP, and GPCC (Group 1) were shown to behave similarly as were the USMex, NARR, VIC, and PRISM (Group 2). In general, Group 1 depicts drier conditions relative to other data during all seasons and for most of the domain. In previous studies, the satellite components of the GPCP and CMAP data have been shown to produce biased estimates of precipitation in the midlatitudes (e.g., Adler et al. 2001), which could help explain these findings. However, over land the GPCP and CMAP are strongly influenced by gauge

TABLE 6. Seasonal	l averages of	the spatial	standard	deviation S_v	, for each dataset.
	0			v	

		GPCC	GPCP	CMAP	NCEP2	GMFD	USMex	NARR	VIC	PRISM
Domain	DJF	23.3	22.90	23.2	27.9	26.4	28.8	29.0	36.7	36.1
	MAM	17.9	17.6	17.8	24.9	18.9	19.4	21.2	23.2	23.1
	JJA	16.8	16.4	16.6	30.8	16.9	15.8	16.3	16.9	16.2
	SON	18.7	18.3	18.4	22.3	19.9	19.1	21.1	24.0	24.4
Pacific Northwest	DJF	29.7	28.6	29.3	31.8	27.4	33.5	34.5	43.4	41.9
	MAM	19.3	18.1	19.5	24.1	17.1	20.4	22.1	25.4	24.3
	JJA	10.2	10.4	10.2	19.3	10.4	8.7	11.2	11.7	11.1
	SON	22.3	21.4	22.2	23.1	21.0	22.3	24.4	29.2	29.0
West Coast	DJF	22.8	20.5	23.1	25.2	24.3	28.3	25.8	39.6	36.4
	MAM	12.8	12.6	12.7	16.4	13.7	15.0	15.4	19.3	18.2
	JJA	5.5	5.6	5.2	13.0	6.5	5.4	6.3	5.7	5.6
	SON	11.2	10.3	11.2	13.3	12.1	11.5	11.9	14.8	15.0
Southwest	DJF	8.9	7.5	8.9	11.9	9.8	10.0	10.4	11.9	11.2
	MAM	9.2	8.2	9.0	15.8	8.9	8.7	9.5	10.2	10.4
	JJA	18.2	15.6	18.1	30.6	16.8	16.0	16.3	16.9	14.9
	SON	12.6	11.0	12.2	15.5	11.2	10.1	11.3	11.8	12.2
Northern Plains	DJF	5.3	6.1	5.3	6.1	7.9	6.7	9.1	8.7	8.7
	MAM	12.5	12.9	12.5	15.1	13.9	10.5	13.2	12.6	12.6
	JJA	14.4	15.4	14.3	22.1	12.7	12.0	12.7	13.6	14.1
	SON	9.2	9.4	9.2	10.0	9.8	9.0	11.5	11.5	12.3
Colorado Plateau	DJF	6.3	7.6	6.1	7.3	9.9	6.8	9.1	8.2	8.0
	MAM	9.4	9.2	9.1	14.7	9.8	8.8	11.4	10.1	9.7
	JJA	9.2	8.2	8.8	17.4	9.4	8.0	9.7	8.8	9.1
	SON	7.6	7.5	7.5	9.8	9.1	7.4	9.5	8.6	8.5

observations. This study has shown a strong similarity between the GPCP, CMAP and the GPCC rain gauge datasets, and any dry bias is similarly expressed by all three datasets. Therefore, it is more likely that sampling issues related to sparse gauge coverage over the western United States by the GPCC rain gauge dataset (which is also used in the GPCP and CMAP) explains the drier conditions observed by Group 1 relative to the other datasets.

The largest differences within Group 1 occur as higher precipitation reported by the GPCP. This is likely a result of gauge corrections used in the development of the GPCP data, which are not used in the CMAP and GPCC. This is consistent with other studies involving the GPCP and CMAP over land. For example, Gruber et al. (2000) noted these differences over northern hemispheric land areas including central North America, which they also attribute to the GPCP gauge corrections. The largest data differences within Group 1 were observed in the Pacific Northwest and Colorado Plateau regions in winter and in the northern plains during all seasons (Table 2 and Fig. 2b).

Group 2 data generally show wetter conditions relative to other datasets, particularly in the Pacific Northwest in winter, spring, and fall and in the West Coast in winter (Table 2 and Fig. 2b). Within this group, the VIC



FIG. 15. Median NMSE and decomposition terms for each season.

and PRISM exhibit the wettest conditions among the group as a result of the orographic adjustment employed in their development. These two datasets were found to be very similar, which is not unexpected because the rain gauge data used in the VIC is adjusted using data from the PRISM. The greatest differences among Group 2 datasets in long-term mean precipitation occur in the Pacific Northwest in winter, spring, and fall and in the West Coast in winter (Table 2). However, when error components are considered, the largest differences are observed in summer for comparisons of the USMex and NARR against the VIC as result of a summertime increase in phase errors. With respect to data distributions (as represented by empirical CDFs), comparisons of Group 2 datasets were shown to be significantly different for most regions and seasons except for comparisons between the VIC and PRISM. However, removing the climatology generally eliminates any distribution differences between Group 2 datasets (Figs. 6 and 7).

The NCEP2 precipitation was generally observed in the midrange of the two groups during winter and fall; however, during spring and summer, the NCEP2 depicts wetter conditions than Group 2 (Table 2 and Fig. 2b). Previous studies have shown the NCEP2 to overestimate precipitation in the central United States during spring and summer (Higgins et al. 1996) and in the central-western United States during the North American monsoon (Janowiak et al. 1998). This has been attributed to inadequacies in the treatment of the Great Plains low-level jet by the spectral model, which results in the transport of excess moisture into the central United States (Higgins et al. 1996; Mo and Higgins 1996). This is consistent with our results; the NCEP2 excesses are observed to occur predominantly in the Northern Plains and Colorado Plateau regions in spring and summer and in the Southwest region where the NAM dominates during summer and fall (the NAM season is typically late summer and into September). Our results also show higher precipitation values for the NCEP2 relative to other data for the Pacific Northwest in spring and summer. In general, the pattern of precipitation bias associated with the NCEP2 over Washington and Oregon (Fig. 2b) is consistent with Widmann and Bretherton (2000), who compared the NCEP reanalysis to gauge-based observations adjusted with PRISM climatology. Their study showed that the NCEP data (relative to gauge observations) underestimated precipitation west of the Cascades, while overestimating it to the east as a result of coarse representation of topography by the spectral model.

The GMFD was shown to be more similar to Group 2 than Group 1. However, it shows a high level of phase

error and bias when compared to both groups. The GMFD shows the wettest conditions of long-term mean precipitation relative to other datasets for all seasons and regions except the Pacific Northwest during winter, spring and fall. Although the GMFD uses the NCEP reanalysis in its development, this study shows the GMFD to be dissimilar to the NCEP2.

The results of this analysis underscore the high level of uncertainty in precipitation observations for the western United States. Additionally, the uncertainty demonstrates important space/time dependencies, which makes it difficult to quantify data error outright. Uncertainty in observations should be considered when using precipitation data for numerical model evaluation. It is possible that the choice of the observational dataset selected for a particular study could affect conclusions regarding model skill.

One limitation of this analysis is that it is not possible to separate the effects of the rescaling from the effects of other data differences on the results. This is recognized as an important issue, especially because the most similar datasets (GPCP, CMAP, and GPCC and PRISM and VIC) share common data sources as well as a common spatial scale. The similarities and differences between datasets are likely a result of some combination of scale and precipitation estimation methodology. However, the contribution of each dataset to the results is difficult to quantify individually. Another limitation is our inability to draw any conclusions about superiority among datasets. This will likely remain a problem because the ability to quantify error requires that we make comparisons against the truth, and no dataset can be assumed to represent the truth free from error.

Acknowledgments. This research was funded by the National Oceanic and Atmospheric Administration (NOAA) under Grants NA04OAR4310078 and NA050AR4310014. This work was additionally supported by NASA Headquarters under the Earth System Science Fellowship Grant NNG04GQ60H. The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA or NASA. The authors wish to thank all the data providers for use of their datasets. The GPCC dataset was provided by the Global Precipitation Climatology Center, Deutscher Wetterdienst (DWD), Germany (available online at http://gpcc.dwd.de). The GPCP, CMAP, and NCEP2 data were provided by the NOAA Climate Diagnostics Center (available online at http://www.cdc.noaa.gov/ cdc/data.gpcp.html, http://www.cdc.noaa.gov/cdc/ data.cmap.html, and http://www.cdc.noaa.gov/cdc/ data.reanalysis2.html, respectively). The USMex dataset was provided by the NOAA Climate Prediction Center (available online at ftp.cpc.ncep.noaa.gov/ precip/wd52ws/us-mex). NARR data were obtained from the NOAA National Climate Data Center (available online at http://nomads.ncdc.noaa.gov). The GMFD dataset was provided by the Land Surface Hydrology Research Group at Princeton University (available online at http://hydrology.princeton.edu/data. php). PRISM data were provided by Oregon State University (available online at http://www.prism.oregonstate. edu). The VIC dataset was supplied by the Land Surface Hydrology Research Group at the University of Washington (available online at ftp.hydro.washington.edu/ pub/HYDRO/data/VIC_retrospective/monthly).

REFERENCES

- Adler, R. F., C. Kidd, G. Petty, M. Morissey, and H. M. Goodman, 2001: Intercomparison of global precipitation products: The third Precipitation Intercomparison Project (PIP-3). *Bull. Amer. Meteor. Soc.*, 82, 1377–1396.
- —, and Coauthors, 2003: The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–Present). J. Hydrometeor., 4, 1147–1167.
- Costa, M. H., and J. A. Foley, 1998: A comparison of precipitation datasets for the Amazon basin. *Geophys. Res. Lett.*, 25, 155– 158.
- Daly, C., R. P. Neilson, and D. L. Phillips, 1994: A statisticaltopographic model for mapping climatological precipitation over mountainous terrain. J. Appl. Meteor., 33, 140–158.
- Ebert, E. E., M. J. Manton, P. A. Arkin, R. J. Allam, G. E. Holpin, and A. Gruber, 1996: Results from the GPCP Algorithm Intercomparison Programme. *Bull. Amer. Meteor. Soc.*, 77, 2875–2887.
- Fuchs, T., U. Schneider, and B. Rudolf, 2007: Global precipitation analysis products of the GPCC. Global Precipitation Climatology Centre, Deutscher Wetterdienst, 12 pp.
- Gottschalck, J., J. Meng, M. Rodell, and P. Houser, 2005: Analysis of multiple precipitation products and preliminary assessment of their impact on global land data assimilation system land surface states. J. Hydrometeor., 6, 573–598.
- Gruber, A., X. J. Su, M. Kanamitsu, and J. Schemm, 2000: The comparison of two merged rain gauge-satellite precipitation datasets. *Bull. Amer. Meteor. Soc.*, **81**, 2631–2644.
- Guirguis, K. J., and R. Avissar, 2008: A precipitation climatology and dataset intercomparison for the western United States. J. Hydrometeor., 9, 825–841.
- Higgins, R. W., K. C. Mo, and S. D. Schubert, 1996: The moisture budget of the central United States in spring as evaluated in the NCEP/NCAR and the NASA/DAO reanalyses. *Mon. Wea. Rev.*, **124**, 939–963.
- —, A. Leetmaa, Y. Xue, and A. Barnston, 2000: Dominant factors influencing the seasonal predictability of U.S. precipitation and surface air temperature. J. Climate, 13, 3994–4017.
- Huffman, G. J., and Coauthors, 1997: The Global Precipitation Climatology Project (GPCP) combined precipitation dataset. *Bull. Amer. Meteor. Soc.*, **78**, 5–20.
- Janowiak, J. E., A. Gruber, C. R. Kondragunta, R. E. Livezey, and G. J. Huffman, 1998: A comparison of the NCEP–NCAR reanalysis precipitation and the GPCP rain gauge–satellite combined dataset with observational error considerations. J. *Climate*, **11**, 2960–2979.

- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. Bull. Amer. Meteor. Soc., 77, 437–471.
- Kanamitsu, M., W. Ebisuzaki, J. Woolen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP-DEO AMIP-II reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, 83, 1631–1643.
- Kistler, R., and Coauthors, 2001: The NCEP–NCAR 50-Year Reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.*, 82, 247–267.
- Liu, Y., and A. Avissar, 1999: A study of persistence in the land– atmosphere system with a fourth-order analytical model. J. Climate, 12, 2154–2168.
- Livezey, R. E., J. D. Hoopingarner, and J. Huang, 1995: Verification of official monthly mean 700-HPA height forecasts: An update. *Wea. Forecasting*, **10**, 512–527.
- Lott, N., G. Ross, and M. Sittel, 1997: The winter of '96-'97 west coast flooding. NCDC Tech. Rep. 97-01, National Climate Data Center, 22 pp.
- Maurer, E., A. Wood, J. Adam, D. Lettenmaier, and B. Nijssen, 2002: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States. J. Climate, 15, 3237–3251.
- Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. Bull. Amer. Meteor. Soc., 87, 343–360.
- Mo, K. C., and R. W. Higgins, 1996: Large-scale atmospheric moisture transport as evaluated in the NCEP/NCAR and the NASA/DAO reanalyses. J. Climate, 9, 1531–1545.
- Murphy, A. H., 1988: Skill scores based on the mean square error and their relationships to the correlation coefficient. *Mon. Wea. Rev.*, **116**, 2417–2424.
- Ralph, F. M., and Coauthors, 2005: Improving short-term (0–48 h) cool-season quantitative precipitation forecasting: Recommendations from a USWRP workshop. *Bull. Amer. Meteor. Soc.*, 86, 1619–1632.
- Ross, T., N. Lott, S. McCown, and D. Quinn, 1998: The El Nino winter of '97-'98. NCDC Tech. Rep. 98-02, National Climate Data Center, 28 pp.
- Rudolf, B., and U. Schneider, 2005: Calculation of gridded precipitation data for the global land-surface using in-situ gauge observations. Proc. Second Workshop of the International Precipitation Working Group IPWG, Monterey, CA, EUMETSAT, 231–247.
- Sheffield, J., A. D. Ziegler, E. F. Wood, and Y. Chen, 2004: Correction of the high-latitude rain day anomaly in the NCEP– NCAR reanalysis for land surface hydrological modeling. *J. Climate*, **17**, 3814–3828.
- —, G. Goteti, and E. F. Wood, 2006: Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *J. Climate*, **19**, 3088–3111.
- Widmann, M., and C. S. Bretherton, 2000: Validation of mesoscale precipitation in the NCEP reanalysis using a new gridcell dataset for the northwestern United States. J. Climate, 13, 1936–1950.
- Xie, P., and P. A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.*, **78**, 2539–2558.
- —, and —, 1995: An intercomparison of gauge observations and satellite estimates of monthly precipitation. J. Appl. Meteor., 34, 1143–1160.
- Yang, Z., and R. W. Arritt, 2002: Tests of a perturbed physics ensemble approach for regional climate modeling. *J. Climate*, 15, 2881–2896.