Evaluating PRISM Precipitation Grid Data As Possible Surrogates For Station Data At Four Sites In Oklahoma

Dr. Jeanne M. Schneider and Donald L. Ford
USDA Agricultural Research Service Grazinglands Research Laboratory
7207 West Cheyenne St., El Reno, OK 73036

The development of climate-sensitive decision support for agriculture or water resource management requires long time series of monthly precipitation for specific locations. Archived station data for many locations is available, but time continuity, quality, and spatial coverage of station data remain significant issues. One possible alternative for station data are continuous, gridded monthly data produced by the PRISM Climate Group, with each grid cell roughly 4 km per side. The PRISM monthly precipitation data is evaluated against station data for four sites in Oklahoma, for possible use whenever station data is unavailable or insufficient. The station and PRISM data are found to be very similar in two key respects (specifically 30-year monthly means and certain characteristics of probability density functions), but significantly different in variance. The difference in variance is attributed to situations where precipitation varied by significant amounts over short distances, with the gridded PRISM data "smearing" the heavy rainfall over locations that received lower amounts. As a result, PRISM precipitation data would be suitable for downscaling seasonal climate forecasts or other analyses that require knowledge of the mean and central shape of the local probability density function, but not for specifying variance as input for weather generators.

INTRODUCTION

Given the significant impact of monthly, seasonal, and annual variations in precipitation on Oklahoma's agriculture and water resources, it would be prudent to develop climate-informed decision support that incorporates the best available information on those variations. The availability of seasonal climate forecasts from the National Atmospheric and Oceanic Administration (NOAA) presents that opportunity, but the climate forecasts are statements of probability for 3-month periods and large areas (see Figure 1 for an example), so the forecasts must to be downscaled to locations, and their potential practical impact interpreted through the use of crop or hydrologic models (1, 2, 3). This area of developing research requires very long duration (covering at least 3 decades, preferably 5 to 10 decades), high quality station data in essentially every location needing decision support, especially for precipitation.

Station data has been collected and archived by NOAA, in particular their Cooperative Summary of the Day data set (COOP hereafter), and continues to be used extensively. Unfortunately, the COOP data sets are irregular in location, duration, and quality, and require significant evaluation and filling of missing data before they can be used with confidence. For example, there have been a total of 364 COOP stations in Oklahoma measuring precipitation at some time since settlement, but only 137 of those stations have at least 50 years of mostly continuous data. Given the high degree of variability in precipitation amounts, this situation leaves most locations in the state under-observed in the manner required for the development of climate-informed decision support for agriculture and water resource management.

Recently, a suite of spatial climate products created by the PRISM Climate Group (PRISM is an acronym for "Parameter-elevation Regressions on Independent Slopes

Model") at Oregon State have become available via internet, and are described online as the USDA's official climatological data (4). The term "climatology" in this application refers to collections of weather data covering 30 or more continuous years; for example, the current NOAA standard climatology covers the years 1971-2000. The PRISM products were created specifically to address the issue of spatially sparse location climatologies (5, 6, 7, 8), by generating spatially and temporally continuous climate data covering the contiguous U.S. The PRISM climate data is defined on a 2.5 minute grid, so the values represent estimates over quadrants approximately 4 km (2.5 miles) on a side. The quality control and grid-filling algorithms employed in the PRISM grid data generation are a rational approach to a difficult problem, using existing station (location) data available from many sources, and accounting for most of the terrain and coastal factors that impact climate on spatial scales of a few kilometers, including altitude. Among the PRISM spatial products, the long term (1895-present) monthly precipitation and maximum and minimum temperature products (9) are of particular interest as a possible surrogate for COOP data in the development of climate-informed decision support tools.

The analysis reported in this article investigates whether the monthly gridded PRISM precipitation data (PRISM hereafter) might be useful surrogates for station precipitation data, for locations lacking station data, or in situations where the existing station data is of questionable quality. Such a use does not appear to have been intended by the PRISM group – their focus was on developing continuous gridded data suitable for use in spatially distributed models and analyses. However, the small grid size suggests that such an application might be plausible, especially given the reported quality of the PRISM data (8). Unfortunately, the statistics reported in (8) were computed over large regions (basically 1/3 of the contiguous U.S.), and do not include a number of details relative to our interest. In particular, there is no Oklahoma-specific report, and we found no published work on use of PRISM data to construct probability density functions (needed for the spatial downscaling of the seasonal forecasts).

Given the methods used to develop the PRISM data, we expect the monthly 30-year PRISM precipitation means to be very similar to co-located station means. We also expect that the processes employed to produce the PRISM precipitation data will necessarily result in some degree of spatial smoothing in the data (i.e., "smearing" the rainfall over a larger area), possibly reducing extreme values (both number of months with zero precipitation, and the largest monthly totals), variance, and skewness, effectively changing the shape of the station probability distributions. If so, how large are the differences, and do they preclude the use of monthly PRISM time series to construct surrogate station probability distributions, or to derive statistics required by weather generators associated with crop or hydrologic models?

In order to begin to answer these questions, analysis was conducted for four stations in Oklahoma, a region where the PRISM data can reasonably be expected to be highly similar to underlying station data, due to the relatively simple terrain in Oklahoma. With such a small sample, this analysis is neither exhaustive nor definitive, but may be sufficient to support initial conclusions concerning possible utility of PRISM data for use in the development of climate-informed decision support for locations.

METHODS

In the United States, most precipitation measurements are collected and archived in units of inches. This is the case for all precipitation data used in this comparison, so all values reported here are also in inches.

The longest duration station climate data available for most of the U.S. (including Oklahoma) are the NWS Cooperative Station Data (COOP). As such, the COOP data are the backbone of the PRISM analysis products. If our goal was to validate the PRISM data, use of COOP data would not be appropriate – independent data, preferably from several sites within each PRISM quadrangle, would be required. However, our goal is different – determining if the PRISM data is sufficiently similar to colocated COOP station data to be used as a surrogate.

COOP Station Data

Daily station data were acquired from NOAA's National Climatic Data Center on CD-ROMs (10) covering the period 1850s-2006. Preliminary surveys of COOP time series in Oklahoma were conducted to identify stations with a long period of record and minimal missing data. To qualify, a COOP precipitation time series had to be at least 75% complete and have data from at least 1948 to 2006. Four stations were chosen in central and eastern Oklahoma for this initial analysis: Enid, Hobart, Madill, and Tulsa. COOP station data have a number of known quality problems, including the observer problems recently reported by Daly (11), which included examples from COOP stations in Oklahoma. To check for the problems reported in (11), the *daily* time series for each station were examined for two particular types of error: underreporting of daily amounts less than 0.05"; and Daly's "5/10 bias", which is over-reporting of daily precipitation amounts easily divided by 5 or 10 (amounts more easily read on the rain gauges, i.e., 1.05", 0.2"). Of the four stations, only the Madill data showed any sign of possible under-reporting of amounts less than 0.05". With respect to the over-reporting error, the Tulsa and Hobart data showed no apparent problem, while the Madill and Enid data exhibited possible small problems. As noted in (11), if significant, these problems would impact statistics for the average number of wet days per month, and the average amount of precipitation that fell on wet days (both statistics are needed for weather generators used with crop or hydrologic models), and possibly on decade-total amounts. None of these statistics are part of this analysis, and it appears that the reporting problems for these four stations are small enough that any related impact on monthly totals can be expected to be small. Calculation of the mean average error (the mean of the unsigned difference) between the COOP and PRISM monthly time series for 1971-2000 are similar to published results for validation of PRISM data in the central U.S. (0.187" for PRISM data, 0.166" average for the 4 COOP stations). This suggests that any such reporting problems for these four stations are no worse than those for other station data used to validate PRISM data in (8).

Whenever daily values were missing, nearby station data were used to fill the gaps by calculating relationships over time between station time series. This is a relatively common (although labor intensive) approach that suffers from the potential problem of assigning precipitation on days when there was rain in the area, but not at that location. All station-based filling techniques (including those used to produce the PRISM products) have this problem. In order to isolate possible impacts from the differences between our technique and that used for PRISM data, a second set of data for each location was generated, omitting months in which three or more days (approximately 10%) were filled, or months in which the filled data comprised more than 10% of the monthly total. This is a relatively severe restriction, as many quality control methods use a 15%-missing criterion. The months identified in this way were eliminated from both the COOP data and the corresponding PRISM data in the second data set, so as to maintain a direct one-to-one comparison.

Daily precipitation data were summed into monthly values, and time series were developed for January 1901 through December 2006 for Enid, and for January 1948 through December 2006 for Hobart, Madill, and Tulsa.

PRISM Grid Data

Monthly PRISM precipitation data were acquired over the internet from the PRISM web page (9), and time series were extracted for grid locations in Oklahoma in the vicinity of the four chosen stations. Station locations did change over the decades (a common and continuing problem), necessitating the creation of location-matched COOP and PRISM time series. (The movement of station locations was automatically accounted for during the generation of the PRISM data grids.) For each individual month, the latitude and longitude of each COOP station was matched to a particular PRISM quadrangle. If a COOP station moved mid-month, the COOP station was matched to the PRISM quadrangle with the majority of days; this impacted only one month each for Hobart and Tulsa.

Analysis Approach

Several complementary approaches were used to answer the questions:

1) How similar are the COOP and PRISM time series and their probability density functions, on a monthly basis, over the entire available record and the current 30-year climatology?
2) When they are different, how (direction, magnitude), and under what circumstances?

The first approach is a classical "difference of means" test and a "ratio of variances" test (12). These tests are evaluating the assumption that the two data sets could be two different sets of samples from a long, common series of events, with equivalent statistics. Both tests require independent sequential observations in both data sets, and Gaussian distributions. If the distributions are not quite Gaussian in shape (the case for monthly precipitation in Oklahoma), then a sample size of at least 30 is required. A probability value (hereafter *p*-value) was generated for each monthly pair of COOP and PRISM data values, representing the probability of the COOP and PRISM data values being the same, scaled so that a value of 1.0 indicates a perfect match. The level of significance for tests was arbitrarily chosen to be 0.10, so *p*-values surpassing 90% indicate equivalence of means or variances within the constraints of the tests. Given the range of the resulting *p*-values for variance, a second "tier" was added at 50% in order to separate months with some degree of correspondence from months with little to none.

The *p*-value results are supported and interpreted using histograms of the differences between the PRISM and COOP monthly precipitation over the common period of record (1948-2006). The largest differences between the COOP and PRISM monthly values are investigated to determine the associated circumstances.

The second approach examines plots of probability of exceedance (PoE) distributions for individual months over the 30 years 1971-2000, the probability distribution format used for NOAA/ CPC seasonal climate forecasts (see Figure 1) and of primary interest here (13). A probability of exceedance function is a variation on a cumulative probability density function, designed to make it easy to associate odds of occurrence with different ranges of precipitation. This format facilitates a risk-based interpretation of climate forecasts and their associated impacts.

RESULTS

The means of the COOP and PRISM data are almost a perfect match post-1951, with few *p*-values below 0.9, and only one below 0.5 (January for 1977-2006 at Hobart). The *p*-value results for the 30 years 1971 – 2000 are presented in Table 1 (all data included) and Table 2 ("filled" data removed). Comparing Table 1 and 2, removal of the filled data has very limited impact on the *p*-values for the means, with the exception of the Hobart data, which was missing almost 3 years

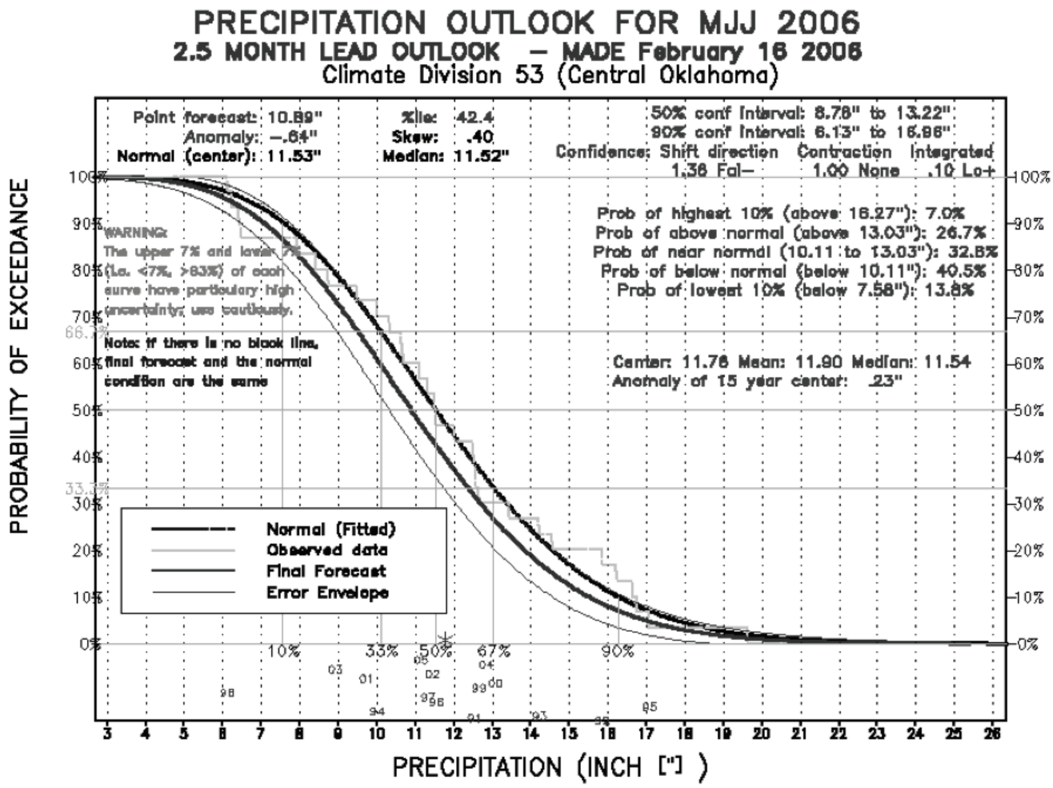

Figure 1. Example of a NOAA/CPC seasonal climate forecast in the format of a probability of exceedance (PoE) distribution. The underlying 30-year climatology for total precipitation in March-June-July (MJJ in the figure title) for this large region of Oklahoma is the stepped grey line (indicated with an arrow) roughly coincident with the smooth black curve identified as "Normal", while the "Final Forecast" is the heavy line just to the left. Other examples of these forecasts and supporting information is available via the internet at http://www.cpc.ncep.noaa.gov/pacdir/NFORdir/HOME3.shtml.

of data in the late 1990s. For the Hobart climatology in Table 2, the sample size may be too small to support any firm conclusions with these tests.

Contrary to the results for the means, the variances of the COOP and PRISM data were significantly different. As shown in the tables, variance p-values greater than 0.9 are rare, and the number of p-values less than 0.5 is discouraging. When the filled data were removed, the p-values change more than they did for the means, but the conclusions remains the same: the means are essentially indistinguishable, but the variances of the COOP and PRISM data can *not* be construed as possibly belonging to a common distribution. Given the similarity in results between the two data sets, all further discussion will refer to the filled data (Table 1).

For most months in the 4-station 1971-2000 comparison, the PRISM variances are smaller than the COOP variances, with half of the months showing differences of less than 10% of the COOP variance. This result is consistent with our expectation before performing the analysis. However, the number of months with very large differences in variance prompted examination of the differences between the COOP

Table 1. Statistics for the 30 years 1971-2000 for the four sites in Oklahoma; *p*-values indicating equivalence are in bold; *p*-values less than that are shaded, with darker shading indicating lower values.

Site	Month	Mean COOP	Mean PRISM	Mean *p*-value	Variance COOP	Variance PRISM	Variance *p*-value	Skewness COOP	Skewness PRISM
Enid 1971-2000	Jan	1.17	1.13	0.88	0.91	0.88	0.86	0.70	0.74
	Feb	1.58	1.57	**0.97**	1.68	1.59	0.76	0.62	0.60
	Mar	2.65	2.73	**0.93**	3.60	3.59	**0.99**	1.03	0.93
	Apr	3.25	3.25	**1.00**	3.62	3.41	0.74	0.55	0.54
	May	4.87	4.90	**0.99**	6.98	6.92	**0.96**	0.70	0.67
	Jun	4.41	4.39	**0.99**	5.64	5.35	0.78	0.52	0.42
	Jul	2.66	2.84	0.75	2.35	2.02	0.41	-0.15	-0.26
	Aug	3.37	3.42	**0.97**	6.23	5.91	0.78	0.72	0.85
	Sep	3.18	3.18	**1.00**	3.80	3.67	0.85	0.33	0.39
	Oct	3.37	3.19	**0.95**	12.84	8.43	0.03	2.09	1.54
	Nov	2.40	2.38	**0.97**	2.35	2.19	0.71	0.45	0.31
	Dec	1.40	1.40	**0.99**	1.41	1.33	0.75	0.88	0.87
	ALL	2.86	2.86	**0.99**	5.46	4.96	0.07	1.49	1.17
Hobart 1971-2000	Jan	0.96	0.95	**1.00**	0.70	0.68	0.88	0.61	0.61
	Feb	1.08	1.07	**0.99**	1.11	0.92	0.33	1.24	0.81
	Mar	2.00	2.03	**0.94**	1.89	2.05	0.67	0.51	0.55
	Apr	2.55	2.53	**0.98**	3.53	3.31	0.74	0.86	1.07
	May	4.55	4.70	**0.95**	9.37	9.53	**0.92**	0.59	0.69
	Jun	3.29	3.61	0.75	3.83	4.04	0.78	0.53	0.24
	Jul	2.40	2.31	**0.94**	4.53	3.83	0.37	1.98	1.96
	Aug	2.82	2.67	0.90	4.81	3.61	0.13	0.84	0.69
	Sep	3.33	3.36	**0.99**	9.06	7.95	0.48	0.81	0.67
	Oct	2.83	2.78	**0.97**	5.16	4.92	0.80	1.41	1.51
	Nov	1.55	1.64	0.83	1.72	1.85	0.69	2.03	1.60
	Dec	1.26	1.26	**1.00**	1.76	1.67	0.80	1.23	1.21
	ALL	2.38	2.41	**0.94**	5.01	4.85	0.53	1.49	1.47
Madill 1971-2000	Jan	2.16	2.09	**0.90**	2.19	1.91	0.46	0.93	0.75
	Feb	2.46	2.42	**0.93**	2.30	2.01	0.48	0.37	0.25
	Mar	3.69	3.62	**0.94**	3.84	3.38	0.50	0.45	0.30
	Apr	3.49	3.59	**0.93**	4.33	4.41	**0.92**	0.99	0.87
	May	5.32	5.24	**0.96**	7.40	6.83	0.67	0.01	0.01
	Jun	5.06	4.98	**0.98**	11.06	9.25	0.34	1.69	1.98
	Jul	2.20	2.39	0.82	2.91	3.26	0.55	0.55	0.45
	Aug	2.72	2.67	**0.97**	5.47	4.96	0.60	1.01	1.13
	Sep	4.49	4.51	**1.00**	9.88	8.95	0.60	0.86	0.82
	Oct	4.50	4.82	**0.94**	11.45	20.52	0.00	1.63	3.08
	Nov	3.17	3.12	**0.96**	4.18	3.37	0.25	0.71	0.44
	Dec	2.72	2.67	**0.96**	4.00	3.52	0.49	0.96	0.82
	ALL	3.50	3.51	**0.98**	6.89	7.19	0.42	1.42	2.33
Tulsa 1971-2000	Jan	1.59	1.61	**0.95**	1.19	1.13	0.76	0.38	0.39
	Feb	1.97	1.96	**0.97**	2.09	1.94	0.70	0.85	0.92
	Mar	3.60	3.58	**0.99**	6.03	5.79	0.82	1.32	1.28
	Apr	3.95	3.93	**0.99**	4.34	3.88	0.55	0.21	0.04
	May	6.11	5.91	**0.91**	7.02	6.07	0.44	-0.03	0.02
	Jun	4.72	4.70	**0.99**	6.04	5.57	0.66	0.27	0.20
	Jul	2.96	2.94	**0.99**	6.15	5.53	0.57	1.33	1.25
	Aug	2.85	2.82	**0.97**	3.52	3.11	0.50	0.82	0.77
	Sep	4.76	4.86	**0.98**	13.11	12.39	0.76	2.10	2.04
	Oct	4.06	4.02	**0.99**	7.00	6.82	0.89	0.47	0.45
	Nov	3.47	3.50	**0.98**	4.55	4.41	0.87	0.26	0.22
	Dec	2.48	2.44	**0.97**	4.48	4.03	0.57	1.12	0.94
	ALL	3.54	3.52	**0.97**	6.97	6.51	0.19	1.23	1.20

Table 2. As in Table 1, except after removing all "filled" data; "#" indicates the number of months remaining in the nominally 30-year analysis period.

	Month	#	Mean COOP	Mean PRISM	Mean p-value	Variance COOP	Variance PRISM	Variance p-value	Skewness COOP	Skewness PRISM
Enid 1971-2000	Jan	28	1.17	1.16	**0.97**	0.97	0.93	0.82	0.67	0.66
	Feb	28	1.66	1.63	**0.95**	1.72	1.64	0.81	0.52	0.50
	Mar	27	2.56	2.55	**0.99**	3.65	3.59	**0.93**	1.18	1.17
	Apr	30	3.25	3.25	**1.00**	3.62	3.41	0.74	0.55	0.54
	May	30	4.87	4.90	**0.99**	6.98	6.92	**0.96**	0.70	0.67
	Jun	29	4.54	4.54	**1.00**	5.32	4.89	0.66	0.57	0.54
	Jul	29	2.75	2.93	0.73	2.20	1.84	0.36	-0.17	-0.27
	Aug	30	3.37	3.42	**0.97**	6.23	5.91	0.78	0.72	0.85
	Sep	29	3.13	3.14	**0.99**	3.88	3.76	0.88	0.39	0.43
	Oct	28	3.54	3.31	**0.94**	13.30	8.79	0.04	2.01	1.45
	Nov	29	2.26	2.25	**0.99**	1.82	1.79	**0.92**	0.01	0.01
	Dec	27	1.41	1.41	**1.00**	1.28	1.28	**0.99**	0.93	0.92
	ALL	344	2.90	2.90	**1.00**	5.49	4.98	0.07	1.51	1.19
Hobart 1971-2000	Jan	27	0.91	0.92	0.83	0.71	0.70	0.80	0.74	0.72
	Feb	26	1.03	1.03	0.84	0.78	0.78	0.18	0.68	0.71
	Mar	28	1.97	2.00	0.90	1.64	1.74	0.25	0.55	0.56
	Apr	28	2.44	2.38	**0.92**	2.77	2.26	0.36	0.64	0.40
	May	28	4.67	4.83	**0.99**	9.76	9.90	0.90	0.51	0.60
	Jun	28	3.39	3.74	0.83	3.79	3.89	0.75	0.52	0.26
	Jul	28	2.44	2.36	**0.91**	4.71	4.03	0.28	1.97	1.89
	Aug	27	2.53	2.50	0.89	3.70	3.02	0.89	0.86	0.66
	Sep	27	3.47	3.38	**0.96**	9.54	8.45	0.34	0.74	0.69
	Oct	27	2.74	2.74	**0.97**	5.20	5.29	0.77	1.52	1.55
	Nov	27	1.56	1.58	0.87	1.89	1.90	0.90	1.92	1.80
	Dec	27	1.25	1.24	**0.98**	1.75	1.68	0.80	1.31	1.28
	ALL	328	2.38	2.41	**0.94**	5.02	4.91	0.52	1.54	1.51
Madill 1971-2000	Jan	28	2.22	2.17	**0.93**	2.28	1.94	0.41	0.84	0.67
	Feb	27	2.44	2.39	**0.91**	1.71	1.58	0.70	-0.19	-0.17
	Mar	28	3.57	3.55	**0.98**	3.69	3.40	0.68	0.50	0.35
	Apr	28	3.51	3.65	**0.91**	4.52	4.60	**0.92**	0.98	0.81
	May	27	5.39	5.36	**0.99**	6.92	6.67	0.85	0.06	0.05
	Jun	27	5.09	5.02	**0.98**	11.04	10.13	0.66	1.81	1.90
	Jul	29	2.28	2.44	0.84	2.85	3.29	0.45	0.52	0.39
	Aug	29	2.78	2.70	**0.96**	5.56	5.11	0.66	0.96	1.08
	Sep	29	4.60	4.64	**0.99**	9.86	8.74	0.53	0.83	0.82
	Oct	28	4.45	4.85	**0.93**	12.22	21.97	0.00	1.63	2.96
	Nov	28	3.02	3.03	**0.99**	3.36	3.26	0.88	0.41	0.47
	Dec	27	2.51	2.54	**0.97**	3.00	3.05	**0.94**	1.07	1.07
	ALL	335	3.48	3.52	**0.94**	6.76	7.37	0.11	1.51	2.41
Tulsa 1971-2000	Jan	29	1.56	1.58	**0.95**	1.22	1.15	0.77	0.44	0.44
	Feb	29	1.92	1.93	**0.99**	2.07	1.97	0.81	0.96	0.98
	Mar	30	3.60	3.58	**0.99**	6.03	5.79	0.82	1.32	1.28
	Apr	30	3.95	3.93	**0.99**	4.34	3.88	0.55	0.21	0.04
	May	30	6.11	5.91	**0.91**	7.02	6.07	0.44	-0.03	0.02
	Jun	30	4.72	4.70	**0.99**	6.04	5.57	0.66	0.27	0.20
	Jul	30	2.96	2.94	**0.99**	6.15	5.53	0.57	1.33	1.25
	Aug	30	2.85	2.82	**0.97**	3.52	3.11	0.50	0.82	0.77
	Sep	30	4.76	4.86	**0.98**	13.11	12.39	0.76	2.10	2.04
	Oct	29	4.13	4.11	**0.99**	7.07	6.83	0.86	0.41	0.40
	Nov	30	3.47	3.50	**0.98**	4.55	4.41	0.87	0.26	0.22
	Dec	29	2.38	2.32	**0.96**	4.33	3.75	0.45	1.26	1.07
	ALL	356	3.55	3.53	**0.97**	7.03	6.55	0.18	1.23	1.20

and PRISM time series to determine the causes; those differences are summarized as histograms over the common period of record (1948-2006, see Figure 2). These histograms indicate that the majority of the pairs of compared monthly totals are highly similar: the overwhelming majority of the differences for all four stations are within the measurement uncertainty of total monthly precipitation (± 0.3″). Further, there are a limited number of significant mis-matches in monthly total precipitation, but those few produce the significant difference in variance. Examination of the months with large differences in mean reveal one of two situations: either a heavy multi-day event (typically the remnants of a tropical storm), or a series of summer days with sporadic, locally heavy rainstorms, with significant variations in precipitation totals over very short distances. The spatial smoothing inherent in the PRISM data appears to have "smeared" heavy precipitation events onto locations that actually received significantly different amounts, resulting in differences in variance in both directions (larger and smaller).

Surprisingly, the statistically significant differences in variance between the PRISM and COOP data appear to have limited impact on the associated probability of exceedance functions (PoEs). The *p*-value tables were used to select two months for each location during the 1971-2000 climatology

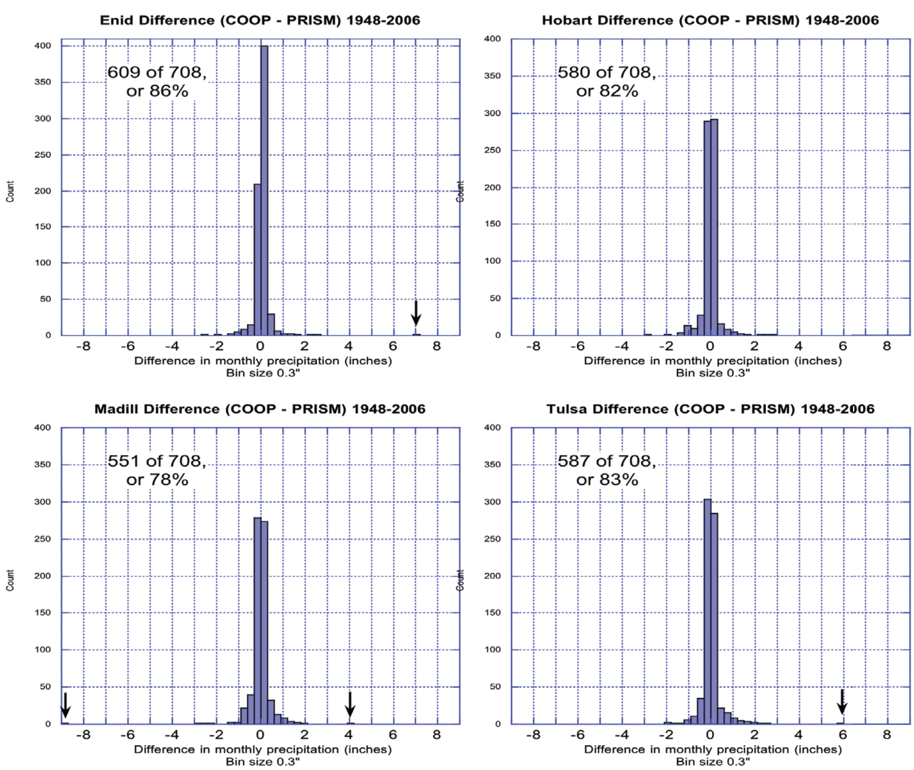

Figure 2. Histograms of the difference between the COOP and PRISM monthly totals over the period of record; if the values were identical, the difference would be zero. The bin size of 0.3″ corresponds to the measurement uncertainty of a monthly precipitation total, so all differences ≤ 0.3″ are essentially negligible. Unusually large differences are highlighted with arrows.

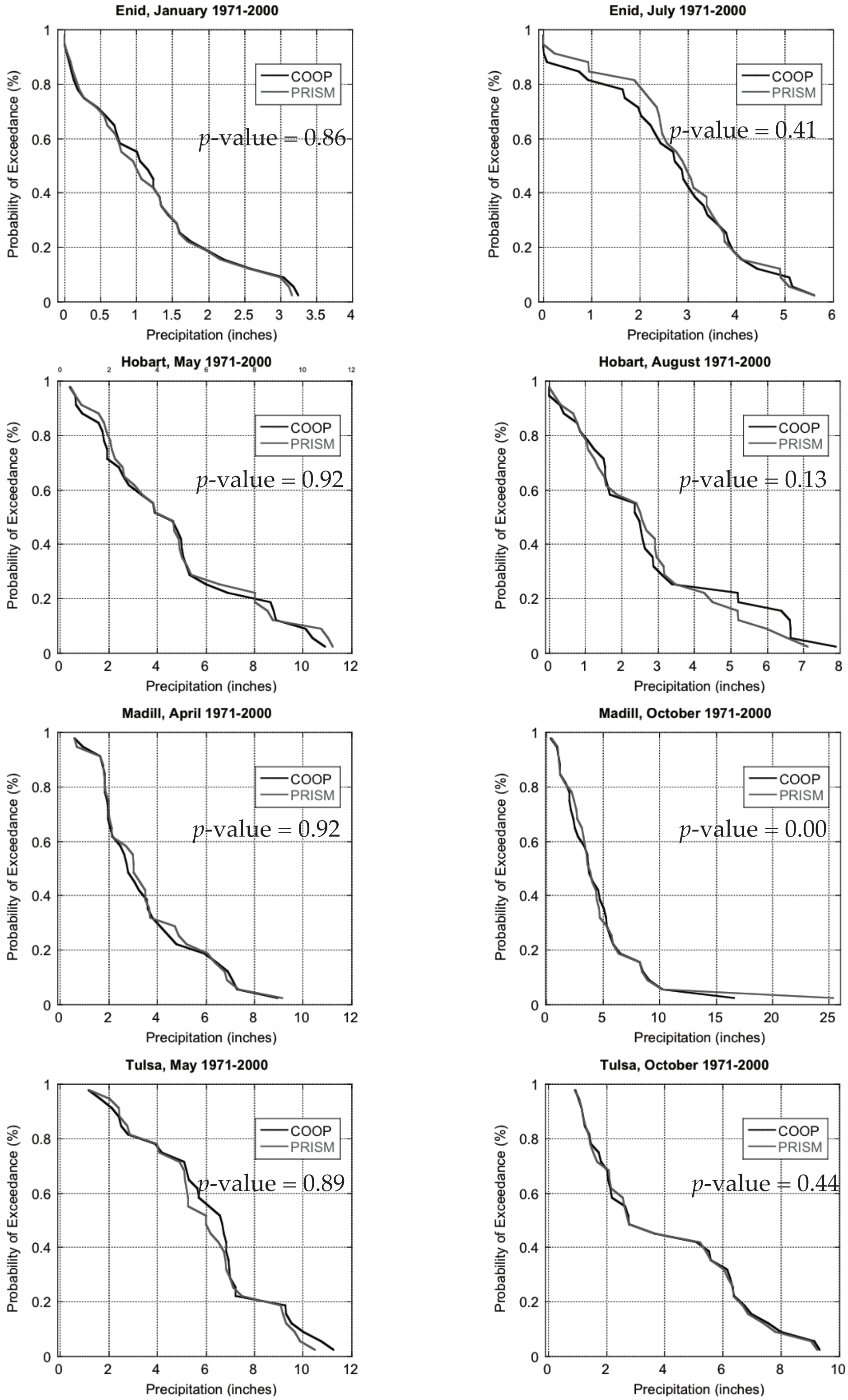

Figure 3. Example PoE distributions for months with relatively good agreement in variance, as indicated by the variance *p*-values reported in Table 1 (left panels) and with relatively poor agreement (right panels). The result illustrated here is that variance *p*-values are *not* a good indicator of the degree of similarity in the associated PoE distributions.

with relatively good versus relatively poor agreement in variance; the corresponding PoEs are presented in Figure 3, with the associated *p*-values. This collection of PoEs is a good representation of the range of variability between the PoEs for these stations.

For the purpose of downscaling seasonal climate forecasts, the slope of the center half of the distribution (events with probability of 25-75%) is of most interest, because this is where the forecast is expected to be most "confident" (14). Because PoEs are constructed from ordered data (smallest to largest values), large differences at the high end of the distribution have a significant impact on variance, but not on the shape of the distribution's center. Interestingly, it is the events where the COOP and PRISM values differ by just a few inches of rain (the shoulders of the difference histograms around the central peak) that distort the central slope of the distribution, but these do not have as large an impact on the variance. Given that estimating the central slope is an approximate process for "noisy" PoEs constructed from only 30 values (PoEs constructed from longer climatologies are much smoother), the impact of the apparent differences between COOP and PRISM data will be relatively small on downscaled seasonal climate forecasts.

In summary, the results are both good news and bad news relative to the possible use of PRISM data as a surrogate for COOP data in the development of climate-informed decision support for agriculture and water resource management. The good news is the essential identity of the means of the COOP and PRISM data, calculated over 30-year or longer periods, for the four stations examined here. This is significant for the potential use of PRISM data to downscale seasonal climate forecasts, ensuring that the location PoEs would be properly "centered" along the precipitation axis. The further surprisingly good agreement in slope across the center of the PoE distributions suggests that the PRISM data would be a useful surrogate for COOP data for downscaling and similar analyses that depend on the shape of the center of the probability distribution. This is the result of the relative insensitivity of the center portion of a PoE to differences in extreme values in the largest and smallest 25% of the distribution.

The bad news is the significant difference between the COOP and PRISM variances. Further, given the results here, it does not appear to be possible to derive a simple predictive relationship between the two that would support a "correction" to the PRISM variance. The range in variance runs from -79% (Madill, October) to 34% (Enid, October), and appears to depend on the number and magnitude of localized heavy rain events during the analysis period "picked up and smeared" by the PRISM algorithms. This factor precludes the immediate use of the PRISM variance data (and by extension, higher order statistics such as skewness, or mean wet-day incidence) to drive a weather generator, which is a common technique to generate daily data from monthly statistics to use as input for crop and hydrologic models.

These results further suggest that a formal validation of the PRISM variance values may be in order, using an independent data source composed of multiple gauges within PRISM quadrangles, assuming such can be identified.

REFERENCES

(1) Schneider JM, Garbrecht JD. 2002. A blueprint for the use of CPC precipitation climate forecasts in agricultural applications. 3rd Symp Environ Appl; 2002 January 14-17; Orlando, FL. Boston (MA): Am Meteor Soc [CD-ROM].

(2) Schneider JM, Garbrecht JD. 2006. Dependability and effectiveness of seasonal forecasts for agricultural applications. Trans Am Soc Ag Biol Eng 49:1737-1753.

(3) Garbrecht JD, Schneider JM, Van Liew MW. 2006. Monthly runoff predictions based on rainfall forecasts in a small Oklahoma watershed. J Am Water Res Assoc 42:1285-1295.

(4) PRISM Climate Group, Oregon State University. [on-line]. Available from http://www.prism.oregonstate.edu/. (Accessed September 28, 2009.)

(5) Daly C, Taylor GH, Gibson WP, Parzybok TW, Johnson GL, Pasteris P. 2001. High-quality spatial climate data sets for the United States and beyond. Trans Am Soc Ag Eng 43:1957-1962.

(6) Daly C, Gibson WP, Doggett M, Smith J, Taylor G. 2004. Up-to-date monthly climate maps for the conterminous United States. 14th AMS Conf App Climatol; 2004 January 13-16; Seattle, WA. Boston (MA): Am Meteor Soc [CD-ROM].
(7) Daly C. 2006. Guidelines for assessing the suitability of spatial climate data sets. Int J Climatol 26:707-721.
(8) Daly C, Halbleib M, Smith JI, Gibson WP, Doggett MK, Taylor GH, Curtis J, Pasteris, P. 2008. Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States. Int J Climatol 28:2031-2064.
(9) PRISM Products. [on-line]. Available from http://www.prism.oregonstate.edu/products/. (Accessed September 28, 2009.)
(10) Cooperative Summary of the Day TD3200/NCDC Cooperative Station Data [CD-ROM]. NOAA National Data Centers; Available on-line from http://ols.nndc.noaa.gov/plolstore/plsql/olstore.prodspecific?prodnum=C00447-CDR-S0001.
(11) Daly C, Gibson WP, Taylor GH, Doggett MK, Smith JI. 2007. Observer bias in daily precipitation measurements at United States Cooperative network stations. Bull Am Meteor Soc 88:899-912.
(12) Freund JE. 1997. Mathematical Statistics. 6th ed. Upper Saddle River (NJ): Prentice-Hall. 624 p.
(13) Wilks DS. 1995. Statistical Methods in the Atmospheric Sciences. San Diego (CA): Academic Press. 467 p.
(14) Probability of Exceedance Outlook. [on-line]. Available from http://www.cpc.ncep.noaa.gov/pacdir/NFORdir/NHOME3.html. (Accessed September 28, 2009.)

Received: September 30, 2009; Accepted August 17, 2010.