

Comment: In the southeast, the PRISM LT dataset uses mostly COOP and WBAN stations so should be very close to what the author is looking for. I respectfully ask that the LT dataset be used as the PRISM dataset in this study.

Response: PRISM LT has been added to the analyses of monthly datasets. Since the study also includes daily datasets, the PRISM AN product remains in the study. In the Discussion section of the manuscript, it is now suggested that PRISM LT could be used for long-term analyses after the correction of a small discontinuity.

Comment: There is also a fairly recent paper from 2021 which covers some important issues about PRISM precipitation time series modeling and can be referenced if desired: <https://journals.ametsoc.org/view/journals/atot/38/11/JTECH-D-21-0054.1.xml>

Response: That article has been referenced in the following sentence: “PRISM LT data are similar to PRISM AN but are available only at monthly time steps and incorporate substantially fewer gauge networks than the AN product, a design choice intended to improve temporal stability for long-term analyses (Daly et al., 2021).”

Comment: 110: with a minimum of eight years required in each group. Why was eight years chosen?

Response: A minimum of eight years per group was required to ensure adequate statistical power and to reduce sensitivity to individual anomalous years.

Comment: 131: I am not sure what is meant by weather-bureau gauges. Do you mean ASOS (Automated Surface Observing System), WBAN (Weather Bureau Army-Navy), or something else? Based on the rest of the paper, I think you mean ASOS.

Response: The term “weather bureau” has been removed from the manuscript. Instead, terms such as “first-order airport surface observing stations” and “USW/WBAN” are used.

Comment: 134: PRISM, which used gauges from 15 networks, showed a similar pattern, with increasing CoCoRaHS and decreasing COOP coverage.

This suggests that the author used the PRISM AN (All Networks) dataset, rather than the LT (Long Term) dataset. AN (All Networks) is not the best PRISM dataset to use for this analysis. In contrast, the LT dataset was developed precisely for this purpose. See general comments for more information.

Response: PRISM AN was the product used in the first version of the manuscript. Both PRISM LT and PRISM AN are now analyzed. In the Discussion section of the manuscript, it is now suggested that PRISM LT could be used for long-term analyses after the correction of a small discontinuity.

Comment: 137: TerraClimate had much less gauge coverage overall, with a maximum of 25% from cooperative gauges and an abrupt decline from 22% to <1% between 2010 and 2011.

According to the text, TerraClimate used WorldClim climatologies and anomalies from CRU time series and JRA reanalyses (but I think JRA is not used in the CONUS). Is the use of COOP data until 2010 shown in the figure derived from what used in CRU? And why did it suddenly stop?

Response: The data shown in the figure were obtained from CRU. I do not know why the use of COOP data for the TerraClimate product decreased dramatically in 2010.

Comment: 138-139: Information on gauge coverage for gridMET was unavailable.

And yet Figure 3 shows “PRISM & gridMET” station usage. I believe gridMET uses PRISM grids so perhaps Figure 3 is mostly correct

Response: That sentence (“Information on gauge coverage for gridMET was unavailable.”) has been removed and gridMET was added to the following sentence. “PRISM AN and gridMET, which used gauges from 15 networks, showed a similar pattern, with increasing CoCoRaHS and decreasing COOP coverage.”

Comment: Figure 4. Percent coverage of the southeastern United States over time by gauge networks used in the five precipitation products.

I do not quite understand this figure. The text refers to Figure 4 when describing a comparison of daily and monthly versions of the datasets, but that is not what the caption refers to. Could it be that the caption should read something like: “Difference between monthly and daily precipitation totals (monthly minus daily?) for the five precipitation products.” But even so, I would not think that gridMET and TerraClimate would have the same results since I don’t think those products are related. Please clarify the figure and the text.

Response: The wrong caption was applied to the figure. The correct caption is as follows:
“**Figure 4.** Percent coverage of the southeastern United States over time by gauge networks used in the precipitation products. COOP is the U.S. Cooperative Observer Program. USW/WBAN gauges are first-order airport surface observing stations. CoCoRaHS is the Community Collaborative Rain, Hail and Snow network. RAWS is the Remote Automated Weather Stations network. With respect to the PRISM and gridMET panel, only COOP, RAWS, and first-order airport (WBAN) gauges are used to develop the PRISM LT product.”

TerraClimate, which is a monthly product, was the initial product used in earlier versions of the study. When I realized I should include daily versions of products, I included gridMET as its daily counterpart, because both datasets are produced by Climatology Lab at the University of California, Merced.

Comment: 149-151: The overall best products were combinations of products, and those products were nClimGrid-PRISM, Daymet-nClimGrid-TerraClimate, Daymet-nClimGrid-PRISM, Daymet-gridMET-nClimgrid, and Daymet-nClimGrid.

Taking the mean of these datasets at the daily or monthly time step could produce some pretty strange results, even if the combination results in an overall unbiased dataset. They also could be a result of happenstance. For example, does the combination of PRISM and nClimGrid happen to cancel out PRISM’s wetting trend from CoCoRaHS with a drying of nClimGrid by increasing its proportion of ASOS stations? I suggest using the term “most unbiased” rather than “best” since it implies that the combined dataset is superior in every way, rather than for this study’s narrowly defined purpose.

Response: Thank you for this thoughtful comment. I agree that the term “best” was too broad and could be interpreted as implying overall superiority across temporal scales and applications. My evaluation pertains only to annual precipitation totals and does not extend to performance at daily or monthly time steps. I also acknowledge that apparent bias cancellation among products may reflect differences in station composition and input data rather than an inherent advantage of combining datasets. To address these concerns and to incorporate the newly added PRISM LT product, the subsection has been substantially revised to focus on “most unbiased” performance within the narrowly defined criteria of this study. The revised paragraph is provided below.

“Several products are suitable for multi-decadal analyses of annual precipitation totals for the southeastern United States. An optimal product should have the following characteristics: (1) a relatively small total of cumulative residuals; (2) no significant inhomogeneities, and (3) a trend within 10% of the reference trend. Two product combinations—Daymet–nClimGrid and Daymet–nClimGrid–PRISM LT—satisfy these criteria. In

these combinations, the wetting bias in Daymet is offset by the drying biases in nClimGrid and PRISM LT. A limitation of these combined products, however, is the reduction in spatial resolution resulting from inclusion of nClimGrid, whose grid cells are at least 16 times larger than those of the other products. For users for whom this spatial disparity is problematic, an alternative is to adjust PRISM LT alone. For example, the post-1993 period can be homogenized by applying a multiplicative factor of 1.0095 derived using the mean-ratio method (Peterson et al., 1998), thereby producing a homogeneous series that satisfies the three criteria.”

Comment: Figure 5: I find it interesting that the PRISM AN dataset showed a discontinuity in 2002. This is the year when weather radar-aided interpolation was introduced into AN. Radar was not used in the LT dataset. See general comments for more on this.

Response: I see the comments and I also do not know why the anomalous values in 2002 exist in the PRISM AN dataset.

Comment: 201-202: CoCoRaHS gauges generally record slightly higher precipitation totals than COOP gauges, with increases of about 1–5% (CoCoRaHS, 2019; Goble et al., 2019).

I don't see that information in either of the two references cited, but you are correct. There is a conference paper by Nolan Doesken (2005) worth citing that reports on results from a 10-year comparison of the 8” SRG with the 4” gauge used by CoCoRaHS. He found that overall, the 4” gauge caught 3% more precipitation than the 8” gauge.

Doesken, N., 2005. A ten-year comparison of daily precipitation from the 4” diameter clear plastic rain gauge versus the 8” diameter metal standard rain gauge. Preprints, 13th Symp. on Meteorological Observations and Instrumentation, Savannah, GA, Amer. Meteor. Soc., https://media.cocorahs.org/docs/AMS_NJD_GaugeComparison_AppldClimate_2-2.pdf

Response: The values in those sources were converted to percentages. The 114 mm vs. 112 mm comparison in Goble et al. (2019) corresponds to a 1.8% increase. The statement in the CoCoRaHS document that CoCoRaHS gauges have a collection efficiency of 101–105% compared to the standard NWS gauge correspond to a 1 to 5% increase. Thank you for the Doesken (2005) reference; it has been added to the manuscript.

Comment: Is this 3% increase sufficient to explain the increasing precipitation over the study period? There is also the possibility that COOP observers show low biases compared to CoCoRaHS because of the difficulty in measuring light precipitation amounts with a measuring stick and an opaque gauge. See this paper for more on that: Daly, C., W.P. Gibson, G.H. Taylor, M.K. Doggett, and J.I. Smith. 2007. Observer bias in daily precipitation measurements at United States Cooperative Network stations. *Bulletin of the American Meteorological Society* 88(6): 899-912. <https://journals.ametsoc.org/view/journals/bams/88/6/bams-88-6-899.xml>

One other (maybe silly!) possibility is that CoCoRaHS observers live in areas that are wetter than the regional average, such as Florida, and are skewing the results in that manner. One would have to control for climatological precipitation conditions to see if wetter areas are now being oversampled compared to dry areas.

Response: I appreciate this comment. The potential influence of COOP observer-related biases was carefully considered. However, because Daly et al. (2007) evaluated biases specific to COOP and comparable assessments for CoCoRaHS are limited, we chose not to speculate on differential biases between the two networks without supporting evidence. The Daly et al. (2007) study is now cited in the Data section, and the following sentence has been added: “Although COOP provides the longest regional record, documented observer-related biases indicate that non-climatic variability may persist (Daly et al., 2007).”

A 3% difference in precipitation totals between CoCoRaHS and COOP is enough to cause a substantial change in a trend, so I do believe the introduction and expansion of the CoCoRaHS stations is the main cause of the wetting biases of the Daymet and PRISM AN products. Let’s use Daymet as an example. It has a discontinuity in 2012, after which the precipitation totals are higher than the totals during 1980-2011. If the totals during 2012-2024 are reduced by just 1.8%, then the discontinuity is removed.

Comment: 228-230: Although using this dataset reduces the spatial resolution inherent to Daymet, the resulting gain in temporal homogeneity makes the Daymet–nClimGrid product the most robust dataset for regional, multi-decadal precipitation assessments.

Again, robust does not seem like the proper term. I would again suggest ‘to have the most unbiased trends’ instead of robust. Also, the author would be remiss if they didn’t make some qualifying statements, here. First, the assessment was made using annual precipitation only and did not dig into the monthly or daily data, for example. The second is that the conclusion does not say anything about the overall quality or accuracy of the dataset, only that temporal trends matched up well. Lastly, the evaluation was made at one spatial scale, that of the entire SEUS, with no subregions within it. It is important that your conclusions support the methods and results of the study. For example, I believe that the PRISM LT dataset has stable temporal characteristics and is likely ideal for your purposes, but it is likely not as accurate as

PRISM AN on any given day, month, or year. Each dataset has been developed with specific goals in mind.

Response: Thank you for this thoughtful comment. I agree that the term robust was imprecise and that additional qualification is necessary to ensure that the conclusions are fully supported by the scope of the analyses conducted. I have revised the text to (1) replace “robust” with language that more accurately reflects agreement in temporal trends, and (2) explicitly clarify that the evaluation pertains only to annual precipitation trends at the regional scale and does not address daily or monthly performance or overall accuracy. The revised text is provided below.

“Several products are suitable for multi-decadal analyses of annual precipitation totals for the southeastern United States.”. An optimal product should have the following characteristics: (1) a relatively small total of cumulative residuals; (2) no significant inhomogeneities, and (3) a trend within 10% of the reference trend. Two product combinations—Daymet–nClimGrid and Daymet–nClimGrid–PRISM LT—satisfy these criteria. In these combinations, the wetting bias in Daymet is offset by the drying biases in nClimGrid and PRISM LT. A limitation of these combined products, however, is the reduction in spatial resolution resulting from inclusion of nClimGrid, whose grid cells are at least 16 times larger than those of the other products. For users for whom this spatial disparity is problematic, an alternative is to adjust PRISM LT alone. For example, the post-1993 period can be homogenized by applying a multiplicative factor of 1.0095 derived using the mean-ratio method (Peterson et al., 1998), thereby producing a homogeneous series that satisfies the three criteria.”