Comment: The author evaluates the Daymet, gridMET, nClimGrid, PRISM, and TerraClimate datasets. The author describes the dataset, showing that most products are largely based on station data. What were the reason for selecting exactly these products?

Response: These products were selected because they represent the primary high-resolution (≤4 km) gridded precipitation datasets available for the conterminous United States that extend back to 1980, which enabled the inhomogeneity and trend analyses in this study. Also included in the analysis is PRISM LT (Long Term), which meets the requirements listed above and was suggested to be added by another reviewer. The PRISM product that in the first version of the manuscript is now named PRISM AN (All Networks).

Comment: The authors mention different kinds of products as well (satellite based, reanalysis). Why did the author not test other potentially useful datasets that do not incorporate precipitation from weather stations? The evaluation of more datasets that do not or only marginally rely on station data would be interesting particularly for research in remote areas.

Response: We appreciate this suggestion. The scope of this study was limited to high-resolution (≤4 km) gridded precipitation products that extend back to at least 1980 in order to ensure a consistent multi-decadal comparison period and fine spatial representation across the southeastern United States. To our knowledge, no satellite-based or reanalysis precipitation datasets currently meet both of these criteria simultaneously. Many satellite and reanalysis products either begin later (e.g., post-1990s) or are available at substantially coarser spatial resolutions.

We agree that evaluating such datasets would be valuable, particularly for remote regions with sparse gauge coverage, but that objective was beyond the scope of the present study.

Comment: Some Figures, such as Figure 2, are hard to interpret given the continuous linear color scale. To cleary show differences between the datasets, difference rasters of precipitation totals of the products would be important, including a discrete color scale of the differences.

Response: I am very appreciative of the suggestion to create difference rasters. Those maps have been created, and that 15-panel figure is the new Figure 3. A new methods section (3.1 Evaluating Spatial Agreement among Precipitation Products) has been added and it as follows:

> "To evaluate spatial agreement and quantify inter-product differences in annual precipitation totals, several complementary comparative analyses were conducted. All six precipitation products were resampled to a

common 1-km spatial resolution to ensure direct cell-by-cell comparability across the southeastern United States. For each year, cell-specific percent differences in annual precipitation totals were calculated for all pairwise combinations of products (15 total comparisons). These calculations produced spatially explicit surfaces representing the magnitude and direction of inter-product differences. To quantify the overall level of agreement, the percentage of the 879,861 grid cells exhibiting absolute differences within ±5% was computed for each pairwise comparison."

In addition, the paragraph in the first Results section (4.1 Spatial Agreement among Precipitation Products) has been modified as follows:

"Mean annual precipitation totals and spatial patterns were broadly consistent among the six precipitation products, with Daymet producing slightly higher totals and nClimGrid slightly lower totals than the others (Fig. 2 and 3). Mean annual totals for 1980–2024 range from 1,348 mm for nClimGrid to 1,434 mm for Daymet, with precipitation totals smallest (~1,100 mm) across central Georgia, South Carolina, and North Carolina and largest (~2,000 mm) in the Blue Ridge and Cumberland Mountains (Fig. 2). Inter-product differences in mean annual totals were remarkably small: among the 15 pairwise combinations, 67% to 98% of grid cells differed by no more than ±5%, with a mean of 89% of cells meeting this threshold (Fig. 3). The largest discrepancies were concentrated in the Appalachian and Cumberland Mountains and in southeastern Florida, although differences exceeding 15% were rare even in these areas. On average, Daymet produced totals 4.7% higher on average than the other products, whereas nClimGrid produced totals 2.3% lower on average, indicating modest wet and dry biases relative to the other products."

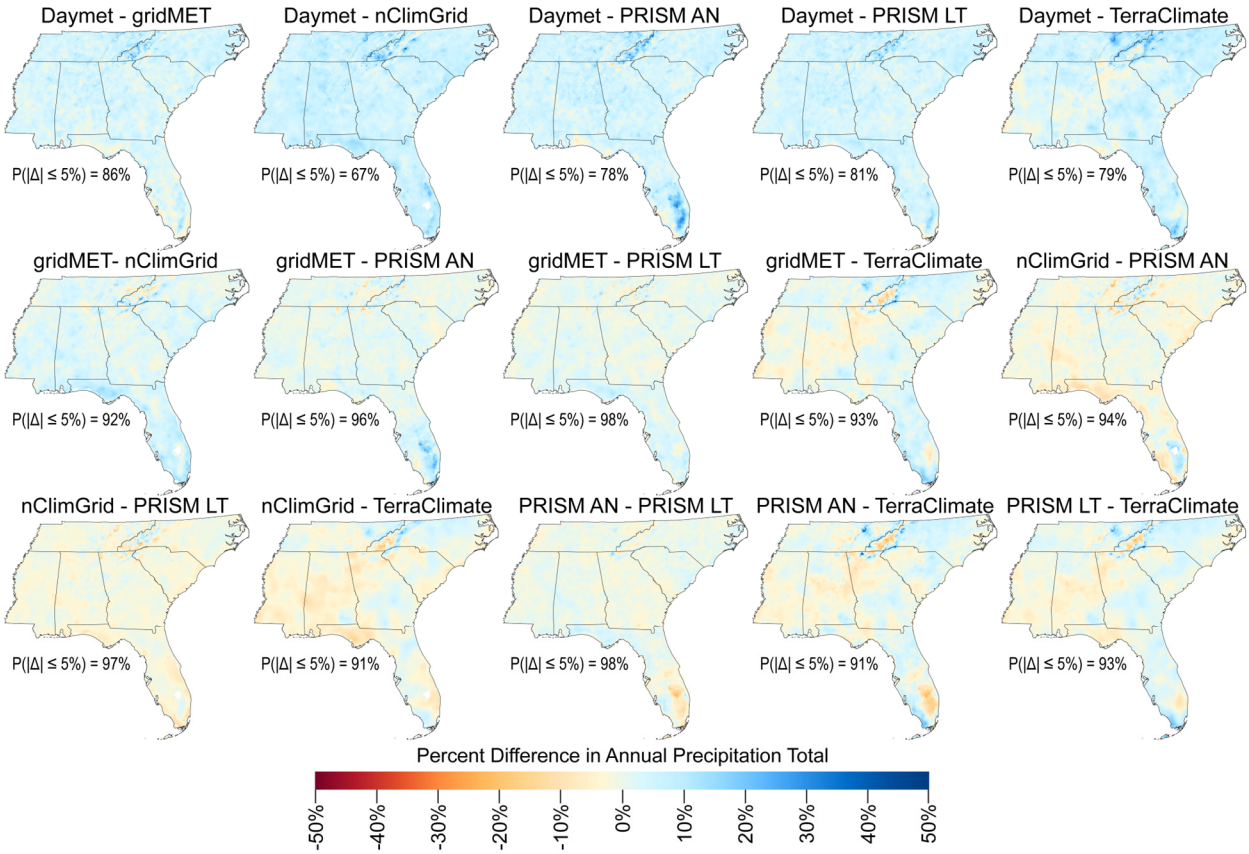


Figure 3. Percent difference in mean annual precipitation totals (1980-2024) between pairs of products. $P(|\Delta| \leq 5\%)$ is the percentage of grid cells with percent differences between −5% and 5%.

Comment: As COOP stations were used for the evaluation, datasets that integrate more of respective stations are expected to show better performance than other products. Can you include a discussion about this effect and is it possible to create an independent evaluation dataset that was not used for the creation of the respective products? Does the combination of datasets increase the proportion of evaluation stations and therefore, leads to improved results? Please discuss these issues

Response: Greater inclusion of COOP stations does not necessarily ensure improved temporal performance, particularly because the specific COOP stations incorporated into each product vary over time. Indeed, several products that rely heavily on COOP still exhibited significant discontinuities and divergent trends, indicating that shared gauge input alone does not guarantee agreement with the reference series. Moreover, the improved stability observed in certain product combinations does not reflect an increase in evaluation stations, but rather the offsetting of wetting and drying biases associated with differing network influences. Fully independent evaluation is challenging for multi-decadal U.S. precipitation products because most gridded datasets assimilate COOP observations, and no spatially extensive, high-quality network independent of the products used in the study extends back several decades.

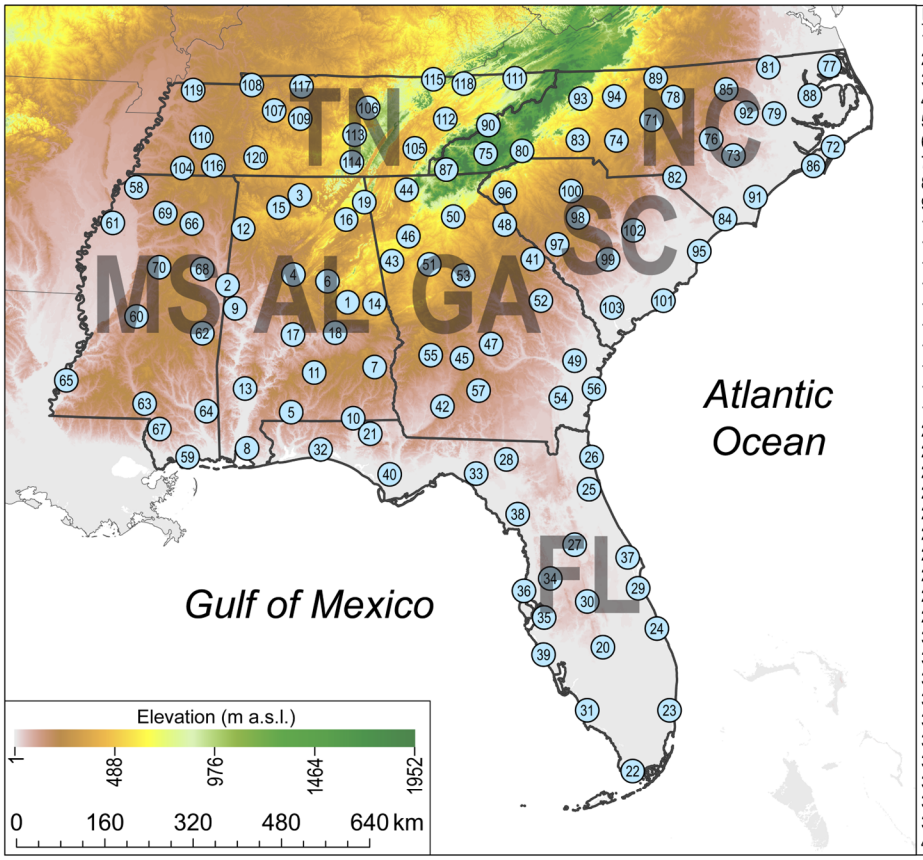
Comment: Regarding the statement that the combination Daymet–nClimGrid is superior to other datasets: this may be based on the strong inclusion of reference stations in these datasets. Without independent evaluation (e.g. test in other area with independent stations) this statement is not very well founded.

Response: Thank you for this important comment. I agree that any statement implying overall superiority must be carefully bounded, particularly given the lack of independent validation outside the study domain. My conclusions are based solely on agreement with the reference dataset for annual precipitation totals for the southeastern United States. The section has been rewritten to avoid implying superiority and to clarify the criteria used for product suitability. 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 1: Please change the color scale here. In the text it says "The gauges ranged in elevation from 1 m a.sl. to 668 m a.sl.;" The yellowish parts in the figure occur at three different elevations in the legend. This is confusing. Use clearly separable colors.

Response: A different color scheme, where yellow appears only once, is now used to show elevation gradients in Figure 1.



ID	COOP#	ID	COOP#	ID	COOP#
1	010160	41	090311	81	314456
2	010184	42	091500	82	314860
3	010655	43	091640	83	314996
4	010764	44	091863	84	315116
5	011080	45	092266	85	315123
6	011620	46	092485	86	315830
7	011725	47	092966	87	316001
8	012813	48	093060	88	316853
9	013160	49	093538	89	317202
10	013251	50	093621	90	319123
11	013519	51	094700	91	319467
12	013645	52	095863	92	319476
13	014193	53	095988	93	319555
14	014502	54	096219	94	319675
15	015635	55	097087	95	381093
16	017207	56	097808	96	381770
17	017366	57	098703	97	384607
18	018038	58	220237	98	385200
19	018469	59	220792	99	386527
20	080236	60	221389	100	387722
21	081544	61	221707	101	388405
22	083020	62	221860	102	388440
23	083163	63	221865	103	389469
24	083207	64	224966	104	400137
25	083874	65	226177	105	400284
26	084366	66	227106	106	402009
27	085076	67	227128	107	402489
28	085099	68	228374	108	402589
29	085612	69	229079	109	403280
30	085973	70	229743	110	404561
31	086078	71	310286	111	404858
32	086240	72	311606	112	404946
33	087025	73	311881	113	405882
34	087851	74	311975	114	406162
35	088782	75	312200	115	406493
36	088824	76	312500	116	408160
37	088942	77	312719	117	408562
38	089120	78	313555	118	408868
39	089176	79	313638	119	409219
40	089566	80	313976	120	409502

Figure 1. Locations of the 120 reference gauges in the southeastern United States. The seven states that comprise the southeastern United States are Alabama (AL), Florida (FL), Georgia (GA), Mississippi (MS), North Carolina (NC), South Carolina (SC), and Tennessee (TN).All gauges are part of the U.S. Cooperative Observer (COOP) network. In the table, the two leading zeros have been removed from all COOP identification numbers.

Comment: Figure 3 and Figure 4: the captions are switched and generally, should include all necessary information and readability should be improved. How were the differences calculated?

Response: I am sorry about the switched captions in the first version of the manuscript. Those figures are now Figures 4 and 5. The differences are the mean annual precipitation totals from the monthly product minus the mean annual precipitation totals from the daily product. The caption has been changed to the following: “Differences in annual precipitation between monthly and daily products, 1980–2024.”

Comment: Figure 7: please clearly elaborate how these differences were derived in the methods section. Are this pixel means covering all stations vs stations means? What is the reference area?

Response: The differences shown in Figure 7 are based on regional mean precipitation values derived from grid-cell averages, which were compared with the mean precipitation of the 120 COOP gauges. To clarify this procedure and the reference domain, the final sentence of the second paragraph of the Data section has been revised as follows: “When calculating regional means, which were compared with the mean of the 120 COOP gauges, approximately 2.5% of grid cells (those exceeding 668 m a.s.l.) were excluded to restrict the analysis to low- and mid-elevation portions of the Southeast and minimize precipitation inaccuracies associated with mountainous areas.”