REVIEWER COMMENTS

The authors evaluate the diurnal cycle characteristics of several satellite/gauge data sets and reanalysis data sets over three continental-scale regions in the Northern Hemisphere against gauge data, and more generally against each other globally. The computations and conclusions are complete and largely convincing, but the text is in need of more consistency and completeness. I have not commented on simple grammatical issues, but someone needs to do a sweep through the entire manuscript to address this problem.

➔ We thank the reviewer for the careful evaluation and for the encouraging assessment that the computations and conclusions are complete and convincing. Following the reviewer's guidance, we have revised the manuscript to improve its consistency and completeness. In particular, Section 2.2 and Table 2 have been reworked so that each dataset is described with comparable information, acronyms and written-out terms have been unified across the text, tables, and figures, and a full proofreading pass has been carried out to correct grammatical issues throughout the manuscript. We respond to each comment point by point below, and the corresponding revisions are indicated by line numbers in the revised manuscript.

Major comments

1. Satellite data set descriptions: Section 2.2 needs to be reworked to give the same information for each data set – pointing at Table 2 from the start can shorten the descriptions a bit? As well, the descriptions are not necessarily accurate:

- IMERG now starts in 1998, even if the dataset started in 2000 at the time that you downloaded it. Also, V07 propagation vectors are computed from multiple reanalysis and forecast fields (but not winds, and certainly not IR). "IR estimates" is really "estimates based on the global collection of GEO-IR Tb's".

➔ IMERG: We thank the reviewer for these corrections. Three revisions have been made: (1) the temporal coverage has been corrected to 1998 to present, and the analysis has since been reprocessed over the full record, so the period used is now 1998–2023 as stated in Section 3; (2) the V07 propagation vectors are no longer attributed to geostationary IR imagery but are now correctly described as derived from the MERRA-2 PRECTOT, TQL, and TQV fields; and (3) the imprecise "IR estimates" has been replaced with the global GEO-IR brightness-temperature formulation following Janowiak et al. (2001). We additionally identified and corrected an error in the gauge reference, previously mislabeled as GPCC Version 3, to the GPCC V2022 analysis used in the IMERG Final Run. The revised passage reads (Lines 263–270):

"The algorithm propagates PMW-retrieved precipitation fields forward and backward in time through quasi-Lagrangian interpolation, in which the motion vectors are derived from precipitation (PRECTOT), total precipitable liquid water (TQL), and total column water vapor (TQV) fields of the MERRA-2 analyses (Huffman et al., 2023c). A Kalman filter subsequently merges these propagated fields with the estimates derived from the global collection of geostationary IR brightness temperatures (Janowiak et al., 2001) to yield weighted half-hourly precipitation estimates (Joyce and Xie, 2011). The Final Run incorporates monthly gauge adjustments from the Global Precipitation Climatology Centre (GPCC), using the Version 2022 gauge analysis (Schneider et al., 2014, 2022), applying spatially varying ratio multipliers that remain constant throughout each month to correct the bias in satellite-based estimates."

- TMPA should be described in the past tense, since it is no longer active. Here I'd drop sensor and GEO satellites, since it's in Table 2. Also, I think you should justify using the more-approximate TMPA-RT instead of the research-oriented TMPA, which includes gauge, as the other satellite products do. I would strongly suggest TMPA.

➔ TMPA: We are grateful for this recommendation, which we have followed. Rather than retaining the near-real-time product (3B42RT), we reprocessed the full analysis with the gauge-adjusted TMPA research product (3B42) Version 7, so that TMPA now incorporates monthly gauge information, as the other gauge-corrected satellite products do. All results from TRMM are replaced with the TMPA in the revised manuscript. Based on the reviewer's comment, the individual PMW sensors and geostationary satellites are no longer listed in the text but are given in Table 2, and the description instead explains the calibration and gauge-adjustment procedure. The correction entry for TMPA in Table 2 has been updated from "None" to "GPCC (M)" to note the monthly gauge adjustment. The revised paragraph reads (Lines 271–281):

"The Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) research product (3B42) Version 7 (Huffman et al., 2007) provided three-hourly precipitation estimates at 0.25° × 0.25° spatial resolution between 50°N and 50°S from 1998 to 2019 when its production was discontinued. TMPA combined PMW precipitation estimates from multiple low-Earth-orbit satellites with geostationary IR estimates (Table 2). The individual sensor PMW estimates were intercalibrated by histogram matching to a common reference, the TRMM Combined Instrument (TCI), which merges the TRMM Microwave Imager (TMI) and Precipitation Radar (PR), and merged into a single microwave field. Geostationary IR brightness temperatures were then converted to precipitation through spatially varying coefficients calibrated against this merged microwave field, and these IR-based estimates filled the gaps between microwave overpasses. The research product additionally applied a monthly gauge adjustment, scaling the satellite estimates to match the monthly accumulations of the GPCC gauge analysis (Schneider et al., 2014). Because this adjustment applied a single monthly ratio to every three-hourly field within the month, it constrained the monthly accumulation while preserving the relative sub-daily structure."

- If you're going to go into such detail about morphing for CMORPH, you need a comment that this approach was the starting point for IMERG's morphing, although they have since diverged.

➔ CMORPH: A statement has been added noting that the CMORPH morphing scheme provided the conceptual basis for IMERG's propagation and that the two algorithms have since diverged (Lines 293–295):

"This morphing approach provided the conceptual basis for the propagation scheme later adopted in IMERG (Joyce and Xie, 2011), although the two algorithms have since diverged, with IMERG now deriving its motion vectors from reanalysis fields rather than from geostationary IR (Tan et al., 2019)."

- For MSWEP, I don't think GridSat provides actual precip estimates. What is the algorithm? In L.298, to what is CDF matching applied? It is not unique for MSWEP to apply undercatch correction – TMPA, IMERG, and GPCP all do that – but I believe the particular scheme used is unique.

➔ MSWEP: We thank the reviewer for these points, which we address in turn. (i) GridSat is not treated as a precipitation product; the description now clarifies that GridSat-B1 infrared brightness temperatures are CDF-matched to CMORPH within air-temperature zones to obtain the pre-2000 estimates. (ii) The CDF matching at former L.298 is now specified as being applied to the merged satellite–reanalysis estimates, matched at the 3-hourly scale to the MSWX reference distribution. (iii) We agree that undercatch correction itself is not unique to MSWEP; the text now attributes the distinguishing feature to the discharge-based scheme rather than to the correction in general. The revised passages read (Lines 316–321, 323–326, and 332–335):

"At a 3-hour time step and 0.1° spatial resolution, the non-gauge component combines reanalysis output (ERA5) with satellite-based estimates, for which IMERG is used after 2000 and infrared-based estimates derived from GridSat-B1 brightness temperatures are used for the pre-2000 period. The GridSat-based precipitation is obtained by Cumulative distribution function (CDF) matching the infrared brightness temperatures to CMORPH within air-temperature zones, rather than taken as a direct precipitation product."

"The merged satellite–reanalysis estimates of each dataset combination are then CDF matched at the 3-hourly time scale to a reference distribution based on the Multi-Source Weather (MSWX) product (Beck et al., 2022) to mitigate spurious light rainfall and the underestimation of heavy precipitation."

"MSWEP further corrects systematic terrestrial biases (e.g., gauge undercatch and orographic effects) by inferring catchment-average precipitation from river discharge observations at 13,762 catchments globally, a discharge-based scheme that distinguishes it from the gauge-based undercatch corrections applied in other products (Beck et al., 2019)."

2. Table 2: Several corrections are needed:

- IMERG V07 Final Run.
- IMERG starts in 1998 (and state if you only used it starting in 2000).
- The first 4 satellite products all have the full complement of GEO-IR satellites because they all use the CPC 4 km Merged IR dataset: GMS/MTSat/Himawari, GOES-E, GOES-W, Meteosat.

➔ We thank the reviewer for these corrections, all of which have been applied to Table 2. The IMERG entry is now labelled "IMERG V7 Final Run" to identify the specific run used. The IMERG time period has been corrected to 1998–present. The analysis period actually used in this study (2000–2023) is stated in Section 3 (Lines 394), alongside the coverage of the other datasets. We confirm that the first four satellite products (IMERG, TRMM, CMORPH, and GSMaP) all take their geostationary input from the CPC 4 km Merged IR dataset, and the IR-sensor column now lists the full complement for each. MSWEP is kept distinct, as its pre-2000 infrared input derives from GridSat-B1 rather than the merged IR product.

A sentence noting this distinction has also been added to Section 2.2 (Table 2 and Lines 320).

3. TMPA-RT, not TRMM: There are multiple TRMM products, so you can't use "TRMM" as shorthand for the product.

- ➔ We agree that using "TRMM" as shorthand for the product is ambiguous, since several distinct TRMM products exist. In the revised manuscript, we changed the dataset from the near-real-time product (3B42RT, TMPA-RT) to the gauge-adjusted research product (3B42, TMPA), based on the reviewer's suggestion in comment #1. Accordingly, every reference to the product formerly labelled "TRMM" has been replaced with "TMPA". The name "TRMM" is now retained only where it denotes the TRMM satellite or its sensors (e.g., the TRMM Microwave Imager and the TRMM Combined Instrument), so that the satellite mission and the precipitation product are no longer conflated.

4. Acronyms vs. written-out names: The manuscript varies between using acronyms, some of which I'm not sure were defined, and written-out phrases. Please take a pass through the text, tables, and figures and enforce consistency.

- ➔ We have made a careful pass through the manuscript, including the main text, tables, and figure captions, to enforce consistent usage. Each acronym is now defined at its first occurrence and used consistently thereafter; written-out forms are retained only where they aid readability; and any term that previously appeared without a definition has been either defined or removed. We have also checked that the acronyms used in the tables and figures are consistent with their definitions in the text.

5. Scaling to daily totals and subdaily detail: On L.934-936 you haven't convinced me that scaling to daily totals will somehow warp the subdaily detail, including the diurnal cycle.

- ➔ We thank the reviewer for this comment and agree that a gauge adjustment which rescales sub-daily estimates to match a daily total need not, in itself, distort the sub-daily detail or the diurnal cycle. When the adjustment is a single multiplicative factor applied uniformly within the day, it rescales precipitation magnitude while leaving the relative sub-daily structure, and therefore the diurnal phase, unchanged. We accept that the original wording at L.934–936 overstated this point and could be read as implying that any daily-total adjustment suppresses the diurnal signal, which is not the case.

  We would note, for clarification, that GSMaP's gauge calibration is not uniform daily rescaling. Following Mega et al. (2019), the gauge-adjusted rate is obtained through an optimal-estimation (maximum-likelihood) procedure in which the 24-h accumulation is constrained to the daily CPC gauge total while the adjustment is determined from a per-pixel optimal weighting; it is applied exclusively over land, with ocean estimates retaining the unadjusted GSMaP_MVK rates. Because this scheme adjusts the hourly values non-uniformly rather than through a single daily multiplier, it can in principle modify the relative sub-daily structure over land.

  Nevertheless, we agree that this calibration effect is not the most robust explanation for GSMaP's reversed land–ocean contrast in sub-daily variance. We have therefore revised the statement so that the contrast is attributed primarily to GSMaP's intensity-dependent retrieval bias over land, which overestimates light-precipitation frequency and underestimates heavy-precipitation intensity and thereby suppresses the sub-daily variance carried by intense convective events; the land-only gauge calibration is retained only as a possible secondary contributor. This attribution is consistent with the corresponding discussion in Section 4.1. The revised passage reads (Lines 1039–1041 and 571–574):

“GSMaP shows a reversed land-ocean contrast in sub-daily variance, plausibly associated with its intensity-dependent retrieval bias over land, namely the overestimation of light precipitation frequency and the underestimation of heavy precipitation intensity, than with its daily-scale gauge adjustment.”

“Second, the gauge calibration of GSMaP is applied exclusively over land, where the hourly rates are optimized so that their 24-h accumulation matches the daily CPC gauge totals (Mega et al., 2019). Because this optimal-estimation scheme adjusts each hour unevenly rather than applying a single daily multiplier, it can reshape the relative magnitudes of the sub-daily fluctuations over land, while the ocean, lacking gauge constraints, retains its original sub-daily structure.”

6. Data availability: The editor gets to weigh in here on what the HESS practice is. My experience is that you should give pointers to the actual datasets that you used. Giving general addresses like https://cdc.dwd.de/ leave the reader floundering around chasing links and potentially grabbing something other than the data set you actually used.

➔ We agree that general portal addresses can leave readers unsure of which data were actually used. We have accordingly revised the Data availability statement to point directly to the specific products, versions, and access paths employed in this study, replacing the general landing pages.

Minor comments

1) significant: “Significant” is used qualitatively, and should be changed to some synonym like “large” or “major”.

➔ *Throughout the manuscript, qualitative uses of "significant" have been replaced with "large," "major," or "marked," reserving "significant" for statistical significance (e.g., Lines 46, 65, 91, 109, 643, 673, 709, 832, 844, 885).*

2) “comprehensive”: I think you should drop using this word and let the reader decide (perhaps) that this study is comprehensive. In particular, the initial wording makes it sound like all Northern Hemisphere land is covered, when it’s really CONUS, western Europe, and East Asia.

➔ We have removed "comprehensive" where it described our own evaluation. It has been deleted from the closing paragraph of the introduction and from the opening of Section 4. This has been revised in the revised manuscript (Lines 152 and Lines 1017).

“The physical mechanisms are then examined over Northern Hemisphere land under contrasting moisture-supply regimes and at two representative site, the ARM SGP site and a site in the southeastern United States.”

“This study evaluates sub-daily precipitation variability and diurnal cycle characteristics across five satellite-based products (IMERG, TMPA, CMORPH, GSMaP, and MSWEP) and four reanalysis datasets (ERA5, JRA-3Q, MERRA-2, and NARR), benchmarked against 6,824 ground-based observations over the Northern Hemisphere during boreal summer.”

3) L.50-51: the phrase "necessary moisture" leaves hanging the question "necessary for what?"

➔ Since the following sentences describe CAPE build-up and convective development, "necessary" was redundant, so we removed it (Lines 52):

"The sensible heat flux deepens the planetary boundary layer (PBL), while the latent heat flux supplies the moisture to the lower atmosphere (Dai, 2024)"

4) L.53: I'd change "low-level inversion" to "inversion at the top of the PBL" since you've introduced the PBL.

➔ Revised as suggested. "low-level inversion" has been changed to "inversion at the top of the PBL" (Lines 55):

"Concurrently, the weakening of the inversion at the top of the PBL and the upward moisture transport reduce convective inhibition (CIN), ultimately triggering convection (Dai, 2024; Dai et al., 1999)."

5) L.69 and elsewhere: "precipitation type" is a pretty vague term, subject to reader mis-interpretation (like "intense" vs. "showers" or "solid" vs. "liquid"). It's more like "storm system type" or something similar.

➔ Rather than replacing it with another general term, we have removed it throughout and refer directly to the convective (CP) and large-scale (LSP) components, which are defined explicitly and leave no room for the misinterpretations the reviewer notes. The same replacement has been applied at every other occurrence, where "precipitation type" now reads either as the CP and LSP components or as the CP-LSP decomposition. For example (Lines 72–74):

"Furthermore, diurnal cycles differ between convective and large-scale precipitation; large-scale precipitation (LSP) generally exhibits a weak diurnal amplitude with a dawn peak (approximately 06 LST), whereas convective precipitation (CP) displays a robust diurnal cycle peaking in the late afternoon (16–20 LST)."

6) L.102: I'd drop the very incomplete "GPM-based" and "microwave-morphing-based" descriptors and let the following paragraphs introduce the differences.

➔ We have removed the "GPM-based" and "microwave-morphing-based" descriptors and instead give the full product names, leaving the algorithmic differences to be introduced in Sect. 2.2 (Lines 104–108):

"With the proliferation of diverse satellite products (e.g., the Integrated Multi-satellite Retrievals for GPM, IMERG; the Climate Prediction Center Morphing Technique, CMORPH; and the Global Satellite Mapping of Precipitation, GSMaP), numerous intercomparisons of the diurnal precipitation cycle have been conducted (Afonso et al., 2020; Janowiak et al., 2005)."

7) L.106: The more-complete statement is "… (e.g., radiative emission by liquid at lower levels, radiative scattering by ice at upper levels, and cloud-top temperatures) rather …"

➔ We have completed the list of indirect proxies as the reviewer recommended (Lines 110–113):

"Furthermore, because satellite-based precipitation estimations rely on indirect proxies (e.g., radiative emission by liquid at lower levels, radiative scattering by ice at upper levels, and cloud-top temperatures) rather than direct surface measurements, they frequently exhibit a systematic phase delay in the diurnal peak relative to ground-based observations (Guilloteau et al., 2018; Petković et al., 2018; Sun et al., 2018)."

8) L.175-176: You need to be more explicit here; presumably this is a volume-preserving averaging, particularly for splitting the CMFD cells across analysis grid boxes? What happens at the edges of data coverage? Is there a minimum number of stations (or CMFD grid boxes) to register a value?

➔ We thank the reviewer for raising these points. Three clarifications are provided below, and the corresponding text has been added to Sect. 2.1.1 (Lines 197–200).

(i) Remapping of stations and CMFD cells. Both point stations and native 0.1° CMFD cells are aggregated to each 0.5° × 0.625° target grid cell by simple averaging of all stations or cells whose centre coordinates fall within it. This is an unweighted average rather than a strictly area-weighted one; however, because the latitude range spanned by a single grid cell is narrow, differences in native cell area are small and the simple average closely approximates the area-mean precipitation rate.

(ii) Edges of data coverage. A grid cell is assigned a value whenever at least one valid station or CMFD cell falls within it. Partially covered cells at the network edges are retained on the same basis, and cells containing no valid data are left undefined rather than filled or interpolated.

(iii) Minimum number of observations. No minimum station count is imposed. As demonstrated in Fig. 2, the variance-then-averaging estimate is insensitive to the number of contributing stations N, whereas the averaging-then-variance estimate follows the spurious 1/N decay. Grid cells populated by a single observation and those populated by many therefore yield comparable variance estimates, so a minimum count threshold is not required.

"The station statistics and the native CMFD cells are aggregated to each target grid cell by simple averaging of all stations or cells whose centre coordinates fall within it. A grid cell is assigned a value whenever at least one valid station or cell is present, with no minimum number imposed, as the statistics are computed at each station prior to spatial aggregation, and partially covered cells at the network edges are retained on the same basis."

9) L.276: I would suggest "Project (GPCP) Pentad (5-day) Product is used … technique in two steps: first using"

➔ Revised as suggested (Lines 299–301):

"Over ocean, the Global Precipitation Climatology Project (GPCP) Pentad (5-day) Product is used to carry out the PDF matching technique in two steps: first using climatological correction based on a 15-year dataset, then applying real-time refinement using a 31-day moving window (Xie et al. 2017)."

10) L.290-290: I think "thereby improving …" is extraneous.

➔ We agree and have removed the redundant clause. The revised text now reads (Line 314):

"Additionally, the retrieval domain has been extended toward the poles during PMW overpasses."

11) L.518: I think something is missing in "Second, GSMAP is applied exclusively over land …"

➔ We have clarified that it is the gauge calibration of GSMaP that is applied only over land. The revised text now reads (Lines 571–572):

"Second, the gauge calibration of GSMaP is applied exclusively over land, where the hourly rates are optimized so that their 24-h accumulation matches the daily CPC gauge totals (Mega et al., 2019)."

12) L.520: I'm pretty sure that "temporally fixed monthly ratio" will lose most readers. Please slow down and be a little more descriptive.

➔ We have expanded the description so that the mechanism is spelled out rather than compressed into the phrase "temporally fixed monthly ratio." The revised text now reads (Lines 575–577):

"In contrast, IMERG adjusts its half-hourly estimates so that they sum to the monthly satellite-gauge combination, applying a single ratio multiplier that is fixed throughout the month (though spatially varying) to every half-hour within that month (Huffman et al., 2023b). Because the same factor scales all half-hours equally, the relative sub-daily structure is preserved."

13) L.633: I think "sub-daily temporal disaggregation" will lose readers unless you introduce the concept in the description back in Section 2.2.

➔ We agree that the term "sub-daily temporal disaggregation" was not sufficiently explained. During the revision of Section 4.2, the product-specific discussion containing this expression was removed from the main text, together with the individual-product panels that are now provided in the Supplement. The term therefore no longer appears in the revised manuscript. Section 2.2 retains only the methodological description of MSWEP as a weighted combination.

14) L.655: "… observation is depicted in (e)-(h)."

➔ Revised as suggested. The caption now reads (Lines 731):

"The spatial correlation of diurnal phase (green solid) and amplitude (red dashed) from gridded products with the observation is depicted in (e)–(h)."

15) L.659-659: This also applies to IR, particularly for TMPA.

➔ In this round of revision, we have added that the same indirect detection applies to IR-based retrievals, particularly for TMPA. Huffman et al. (2007) reported a timing difference in the IR-dominated real-time product relative to the microwave-dominated research product, consistent with the reviewer's point that the indirect detection also affects IR (Lines 742–745):

"The same indirect detection affects IR-based retrievals, which infer precipitation from cloud-top brightness temperatures that are only weakly related to surface rainfall and tend to place the diurnal maximum later; this is particularly relevant for TMPA, which fills the gaps between PMW overpasses with calibrated IR (Huffman et al., 2007)."

16) L.668-669: The current phrasing misleadingly implies that there's some kind of algorithmic linkage among the reanalyses. You're just giving the various offsets.

➔ We agree, and we did not intend to suggest a continuum among the products. The phrasing has been changed to present the offsets individually (Lines 757–758):

"Reanalysis products exhibit premature phase biases and attenuated amplitudes, with peak precipitation timing at 16 LST in JRA-3Q, 14 LST in ERA5, and 12 LST in MERRA-2."

17) L.698-699: You're listing rows, not columns.

18) L.700: "backward and forward masking in Fig. 6" is really confusing. I think you just mean "the masking depicted in Fig. 6".

➔ Both corrected. TP, CP, and LSP are arranged by row rather than column, and "backward and forward masking" has been replaced with "the masking depicted in Fig. 9". The figure is now given as Fig. 11 (Lines 797-801)

"Figure 11: Diurnal cycle of area-averaged precipitation over land (a-c) and ocean (d-f) grid cells bounded by 60°S–60°N (only grid points after the masking depicted in Fig. 9), for JRA-3Q (a, d), ERA5 (b, e), and MERRA-2 (c, f). Bars show the partitioning of total precipitation (TP) into convective precipitation (CP, orange), and large-scale precipitation (LSP, blue) components, and the black line shows the TP from CMORPH, with the black dot marking its diurnal maximum. Background shaded bars indicate the local time at which each reanalysis reaches its TP maximum."

19) L.761-2: I think "In CMORPH" should come after "precipitation maximum" for clarity.

➔ This sentence has been removed as part of the reorganization of Section 4.3. The relationship is now described directly in the discussion of the evaporation-supplied regime.

20) L.929: Here as well, IR contributes to the delay, not just PMW ice scattering.

- ➔ We agree with your comment. In the revised manuscript, we now attribute the delay to both passive microwave and infrared retrievals (Lines 1031–1033):

  "Satellite products place the afternoon precipitation peak 1–2 hours late, which follows from indirect detection by both PMW sensors, which respond to upper-tropospheric ice-scattering signals rather than surface precipitation, and IR retrievals based on cloud-top temperatures."