Response to Reviewer 2

We thank Reviewer 2 for the exceptionally careful and constructive reading of our manuscript. We are especially grateful for the reviewer catching that key methodological details were deferred to Aftab et al. (2026c), which is a Zenodo data record rather than a methods paper. The intended reference is our companion paper Aftab et al. (2026a, *Int. J. Climatol.*), reading which will clear a lot of confusions hopefully.

We respond to each comment below; reviewer comments are in italics and quoted manuscript changes are indented.

General comment

Aftab et al. present a new index for quantifying compound dry-hot extremes, the SSTCI... the paper requires further motivation on what makes this particular index novel compared to (and outperforming) the current state-of-the-art. The authors also need to elaborate more on the methodology used to construct the metric. I recommend to replace some of the case studies in the paper with a more thorough evaluation.

We thank the reviewer for this clear summary and have restructured the revision around the three points raised. We will sharpen the motivation to make the contribution explicit. It rests on two pillars. First, the index: SSTCI is a global, daily, 0.1° index that is soil moisture centric, its drought component (SASMI) is derived directly from soil moisture rather than from a precipitation based proxy (as in SCDHI/DCDHI) ;and it uses a grid specific dry down memory (τ) so that antecedent moisture is weighted according to each location's own land-surface behaviour. Since soil moisture is the variable most directly tied to the land–atmosphere feedback driving compound events, and the daily resolution resolves the onset, peak, and decay that monthly indices miss, this combination captures conditions that coarser precipitation proxy indices cannot. Second, the catalogue: rather than simple fixed-threshold exceedance, which fragments events and retains spurious spells, we apply the objective removal–merging optimization of Shan et al. (2024), which retains only statistically independent (Poisson inter-arrivals) and genuinely extreme (GEV severities) events. To our knowledge this is the first application of this procedure to compound dry–hot events, and no global, daily, Soil moisture-based compound dry–hot dataset and companion catalogue of this kind currently exist ; providing them as ready-to-use products is itself a central contribution of this data paper.

Methodology will be explained in detail in revision. The revision provides a self-contained description of the SASMI, STI, and SSTCI formulations with all symbols defined, so the index can be reproduced from this paper alone (see responses to Sect. 2.3, 2.5, and L144).Along with this, we are willing to strengthened the quantitative assessment and reduced reliance on illustrative case studies, as recommended: we add rank-based (Spearman) agreement maps against two established daily drought indices and an extreme-event agreement analysis (see Fig. 2 response), and we will move the more demonstrative subsections (wavelet analysis, Sect. 3.6; 2022 Yangtze reconstruction, Sect. 3.8) to the Supplement. The detailed responses below describe each change.

Specific comments

L26: *"removal–merging refinement to threshold based spells": this sentence can surely be rephrased using more comprehensible language.*

We agree. In the revised abstract we will rephrase this in plain language, e.g. "we post-process the daily SSTCI series by discarding very short extreme spells and merging those separated by only brief interruptions into single events." The full technical description is retained in Sect. 2.6.

L54–55: *The authors assume the reader already knows all these indices/acronyms (SCDHI, SAPEI, STI, DCDHI, ...).*

We agree. In the revision we will expand each acronym at first use and add a brief clause identifying what each index measures, so the reader is not expected to know them in advance. For example, SCDHI (Standardized Compound Dry–Hot Index), SAPEI (Standardized Antecedent Precipitation Evapotranspiration Index), and DCDHI (Daily Compound Dry–Hot Index) will each be spelled out and briefly characterised where they first appear.

L57: *"drought components in these indices may not fully capture soil moisture variability": explain why, perhaps the current state-of-the-art needs to be introduced in a bit more detail.*

We will clarify this in the revision. The point is that the drought components of existing sub-monthly compound indices (e.g. SCDHI, DCDHI) are derived from precipitation based proxies such as SAPEI or SPEI, which infer moisture status from the atmospheric water balance rather than from soil moisture directly. Such proxies do not explicitly represent antecedent Soil moisture memory (McColl et al., 2017) or the regional heterogeneity of dry-down rates, so they can misrepresent the actual Soil moisture state that drives land–atmosphere coupling during compound events. We will add a short passage making this reasoning explicit and expanding the description of the current state-of-the-art. We note this is closely related to the reviewer's comment on L72–78 (repetition of the precipitation-vs-Soil moisture point); we will consolidate the argument here so it is made once, in full, and remove the redundant restatement later in the Introduction.

L68: *"most traditional drought metrics rely on precipitation signals": really? What about simple maps of soil moisture anomalies which are very common? And the ET-based metrics you use for comparison further on in the paper?*

The reviewer is correct that this statement was too categorical. Soil moisture anomaly products are indeed widely used, and evapotranspiration based indices such as DEDI, which we employ for validation are not precipitation based. Our intended point was narrower: that many of the standardized, widely adopted drought indices used operationally and in compound event studies (e.g. SPI, PDSI, SPEI) are precipitation driven, and that the sub monthly compound indices we build on specifically inherit precipitation based drought components. We will rephrase to make this narrower claim, explicitly acknowledging soil moisture anomaly and ET-based approaches as existing alternatives.

L72–78: *Repetition of previously made points (temporal scale and precipitation vs soil moisture focus).*

As noted and we will remove the redundant restatement at L72–78.

Sect 2.1 (mx2t vs 2t): *It was not very clear what the difference is between mx2t and the more standard 2t, since both appear to be hourly (so it's not the daily maximum temperature). This is particularly relevant since 2t is available in ERA5-Land at higher resolution, so there would be no need to interpolate it to a finer grid.*

We thank the reviewer for catching this; the description was ambiguous and we will correct and expand it. Each variable is taken from the product that best represents it, for a specific reason.

The two variables are not the same quantity. ERA5-Land provides only the instantaneous 2 m air temperature (2t), defined by ECMWF as the "temperature of air at 2 m above the surface of land, sea or in-land waters." ERA5's mx2t, by contrast, is the "maximum 2 m temperature since previous post-processing," defined as "the highest temperature of air at 2 m above the surface … since the parameter was last archived in a particular forecast" that is, an hourly maximum rather than an instantaneous value. Because heat extremes and their impacts are governed by daytime peak temperature rather than by instantaneous or mean temperature, a maximum based variable is required; ERA5-Land does not provide one, whereas ERA5 does (mx2t). Maximum daytime temperature is also the standard variable used to define heat extremes across the large majority of operational and research heatwave definitions (Perkins-Kirkpatrick and Lewis, 2020). We therefore source temperature from ERA5's mx2t and bilinearly interpolate it to the 0.1° ERA5-Land grid to match soil moisture product, accepting the interpolation as the cost of using the physically appropriate variable.

Soil moisture, in contrast, is taken from ERA5-Land, whose land-surface fields (including soil moisture) are produced at 9 km (0.1°) and are demonstrably improved relative to ERA5, with added value in the hydrological cycle shown against in-situ and satellite references (Muñoz-Sabater et al., 2021).

We will also clarify precisely what daily quantity is used, which was the source of the confusion. From the hourly mx2t field we compute a daytime temperature as the mean of the hourly maxima over 06:00–18:00 local time at each pixel i.e. a daytime averaged maximum temperature representing sustained peak daytime heat, rather than an instantaneous value. We will state this explicitly in the revised Sect. 2.1 and adjust the variable description accordingly

Sect 2.3: *For consistency with the other symbols, D in equations 2 and 3 should get a t subscript index. Also, not all symbols are explained (e.g., F_LL, phi, …).*

In the revision we will write do all the corrections . F_{LL} is the cumulative distribution function of the fitted log-logistic distribution; and Φ^{-1} is the inverse standard-normal (quantile) function. These definitions will be given where the equations first appear and collected in the appendix.

L144: *"Drier than normal": I don't understand how this works based on the metric definition; from the text, I understand that the N in equation 2 equals 14, so we only look two weeks into the past? Does this mean the complete time series is not used to compute the SASMI metric? What if it has been dry for two weeks? Or, will the onset of the dry season always appear as a dry extreme?*

We thank the reviewer for this comment, which showed us that our description conflated two distinct steps. We will revise Sect. 2.3 to separate them clearly.

There are two uses of the record in SASMI. First, Eq. (2) defines the antecedent signal D_t which is a memory weighted sum of soil moisture over the preceding ~14 days, where the weights decay geometrically with the grid specific dry-down timescale. This is only the input feature; it is not itself the anomaly. Second, and separately, this D_t value is converted into an anomaly by referencing it against its own climatology (Eq. 3): for each grid cell, the log-logistic CDF is fitted per calendar day-of-year across all years of the record (1961–2023), exactly as the normal distribution is fitted per calendar day for STI (Sect. 2.4). D_t is mapped to its percentile within that day-of-year distribution and transformed to a standard normal variate.

"Drier than normal" means drier than is typical for that grid cell on that day of the year, evaluated over the full record; not drier than the preceding two weeks. The 14-day window sets only how far back antecedent moisture is accumulated; it does not restrict the reference climatology. Consequently, the onset of a normally dry season does not register as an extreme: a low D_t that is typical for that calendar day maps to a near median SASMI, because the day-of-year reference distribution already accounts for the seasonal cycle. A dry extreme arises only when D_t is unusually low relative to other years on the same calendar day. We will make this two step structure, and the per-day-of-year fitting, explicit in the revised methods and appendix.

L151: *Replace "each calendar day" by "a specific calendar day"?*

we will make this change.

Sect 2.5: *Again, symbols used in the equations are not properly explained... What are X, Y, u, v? F is a marginal cumulative distribution, but of what? The authors refer to Aftab et al. (2026c) for more details, but this is a Zenodo dataset... it should be extended here, in an appendix or in a supplement, but not by referring to someplace else.*

We will fix both issues. We will define every symbol in Eqs. (6)–(7): X and Y are the SASMI and STI variables respectively; $u = F_X(x)$ and $v = F_Y(y)$ are their probability-integral transforms to uniform margins; $C(u,v)$ is the fitted Frank copula; p is the resulting joint exceedance probability of concurrent dry–hot conditions; and F is the cumulative distribution of p across the record, used to remap it to uniform before the inverse normal transform yields SSTCI. As the reviewer correctly notes, the earlier pointer was to a Zenodo record; the intended methodological reference is Aftab et al. (2026a), it was a referenced wrong.

L188: *Is the threshold of −2 chosen because the unit of the SSTCI is 'standard deviations from a normal distribution'? In this case, do you not expect to always detect 2–3% extreme events in each pixel, both in areas where you regularly encounter CDHEs, and those that never experience them? Additionally, it would be good to see some SSTCI histograms to verify the standard normal behaviour.*

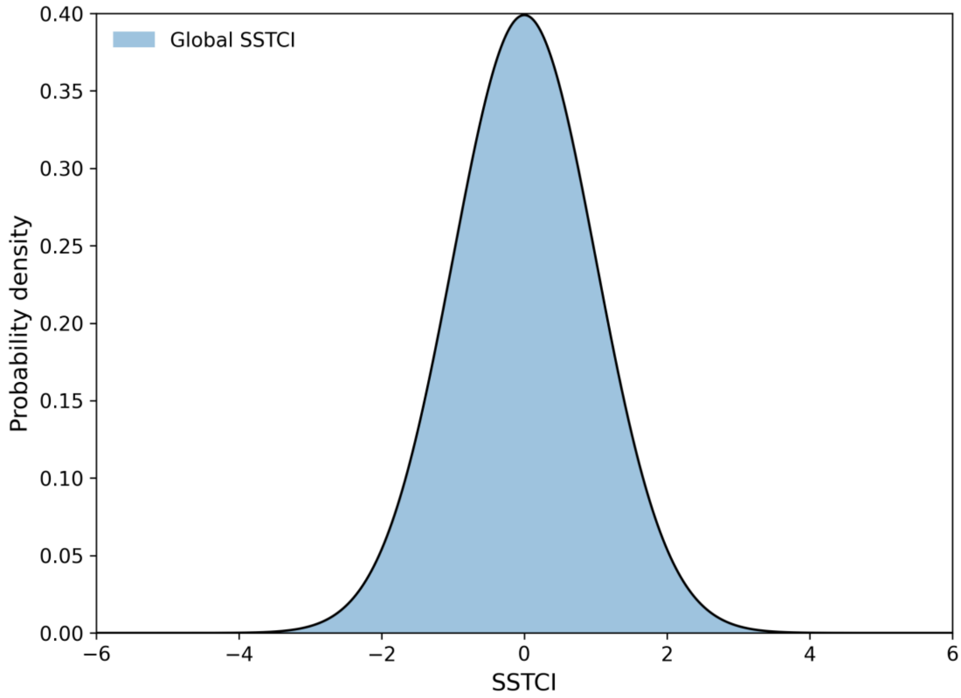
The reviewer is correct on both points, and we verified the standard-normal behaviour directly: pooled over all land pixels and days ($N \approx 3.1 \times 10^{10}$), SSTCI is indistinguishable from N(0,1) (mean = 0.000, s.d. = 1.000; P(SSTCI < −2) = 2.27% vs. the theoretical 2.28%). We will add this as a supplementary figure (attached below for your reference: the pooled global SSTCI distribution overlaid on the standard normal).

We agree that, being standardized, the daily exceedance rate below −2 is fixed at ~2.3% by construction, as for all standardized indices. Two points clarify why this does not undermine the −2 choice or the spatial information content:

First, the −2 threshold is not arbitrary but is empirically justified in our companion methodological paper (Aftab et al., 2026a): a conditional probability analysis there (their Fig. 3) shows that at SSTCI ≤ −2, moderate single driver conditions are almost entirely excluded (e.g. probability 0.03 for STI=1.5, SASMI=−1.5), while only jointly extreme dry–hot combinations are retained (probability 1 for STI=2, SASMI=−2). The threshold therefore isolates genuine compound extremes rather than cases where only one factor is elevated.

Second, although the rate of exceedance is spatially fixed, the index being thresholded is locally informed: SASMI weights antecedent soil moisture by a grid specific dry down timescale τ (Sect. 2.2). Each pixel's compound signal thus encodes its own land surface memory.

What varies in space is not the daily exceedance frequency but the event structure produced by the removal–merging catalogue (Sect. 2.6); the number, duration, and clustering of coherent events, which is not pinned to 2.3% and is the actual hazard product. We will make both points explicit in the revised Sects. 2.5–2.6.



Section 3: *Some of the later subsections can perhaps go into a supplement.*

In the revision we will move the continuous wavelet analysis (Sect. 3.6) and the detailed 2022 Yangtze River Basin event reconstruction (Sect. 3.8) to the Supplement, retaining brief summaries in the main text. This streamlines the results and creates room for the additional quantitative evaluation requested elsewhere in the review.

L216/220: *"managed/agricultural regions", "land surface controls": does ERA5-Land account for irrigation/land surface controls?*

The reviewer raises a valid point. ERA5-Land does not include an explicit irrigation scheme, so anthropogenic water management is not directly represented in the Soil moisture fields; the land surface responds only to modelled precipitation, radiation, and land-cover-dependent processes. We will soften the wording at L216/220 accordingly, replacing the implication that the index captures managed/irrigated dynamics with a clarification that τ reflects natural land-surface and hydroclimatic controls as represented by ERA5-Land, and we will note the absence of irrigation as a caveat where relevant.

L256: *Non/significant and small correlations also clearly visible over desert areas like the Sahara. Thoughts?*

We appreciate this observation, and there are two reasons, why we regard these areas as low-confidence and mask them.

First, over hyper-arid deserts such as the Sahara, ERA5-Land soil moisture is only weakly constrained: assimilated observations are sparse and the land-surface model has little dynamic information to work with, so much of the residual variability there reflects model noise rather than a physically trustworthy signal. We do not consider it defensible to interpret correlations built on this noise, and they are consequently shown as masked (white) in Fig. 2.

Second, and independently, these regions are perennially moisture-limited: near-surface soil moisture remains at or near its residual floor year-round, with negligible dry–wet dynamic range. There is therefore little meaningful variability for any pair of drought indices to agree on, and the comparison is uninformative in principle , this affects DEDI and SPEI equally, not only SASMI.

For both reasons, we deliberately focus the evaluation on higher confidence regions where the Soil moisture signal is physically meaningful and better constrained. We will add a brief statement to this effect in the revised text, and note that a more targeted assessment of hyper-arid regions would be a worthwhile direction as better-constrained Soil moisture observations become available.

Figure 2: *The Pearson correlation is not well-suited for time series in which you expect extreme events. Consider using Spearman's rank correlation as an alternative. Also, a map with an alternative metric could also be useful, e.g., the fraction of times both metrics agree on what is an 'extreme event' (by the 2 sigma definition)..*

We thank the reviewer for both suggestions, which we have adopted; together they provide a two-part evaluation of SASMI against the two established reference indices (DEDI, SPEI).

we recomputed the agreement using Spearman's ρ over the full daily record. SASMI correlates positively and significantly with both indices across essentially the entire land surface (SASMI–DEDI: median $\rho = 0.28$, 97% of pixels positive, 98% significant; SASMI–SPEI: median $\rho = 0.36$, 99% positive, >99% significant at $\alpha = 0.05$), and the spatial patterns match the Pearson maps . This establishes that SASMI tracks the same day-to-day drought variability as the established indices. In the revision we will present the Spearman maps as the main Fig. 2 and note that Pearson gives consistent patterns. Figure attached below for your reference.

We also computed the reviewer's suggested metric, the per-pixel fraction of agreement on extreme (2σ) days and will add it as a supplementary map as its not main topic of discussion. Read together with the correlation maps, the two metrics answer different questions and give a coherent picture. The correlation maps show that the indices co-vary strongly (variability agreement), whereas the 2σ agreement maps test whether they cross the extreme threshold on the identical calendar day (exact-timing agreement); a far stricter condition. The exact day agreement is modest (median ≈ 0.015 for SASMI–DEDI, ≈ 0.004 for SASMI–SPEI), for two well understood reasons.

First, the three indices are integrated over different temporal windows ; DEDI is a daily deficit, SPEI a 30-day accumulation, and SASMI a ~14-day antecedent memory(though all daily but the way they are calculated makes the difference) so, even when they agree that drought is occurring, the precise day each crosses −2 is offset. This is confirmed by the fact that SASMI–SPEI shows the higher rank correlation yet the lower exact-day agreement: SPEI's 30-day window is the most offset from SASMI's 14-day window, so its threshold-crossing timing diverges most, exactly as the mechanism predicts. Second, the spatial pattern of agreement tracks the documented reliability of the reference datasets rather than the quality of SASMI. Agreement is lowest over hyper-arid regions (Sahara, Arabia, interior Australia), and dense tropical forest (Amazon, Congo), precisely the regions the reference-dataset papers themselves flag as unreliable (arid PET artefacts, sub-canopy signal loss, and freeze–thaw effects in Zhang et al., 2023; barren/low-variability masking in Liu et al., 2024) and is highest over the temperate and humid regions where all indices are most reliable.

Finally, to show that modest daily-2σ agreement is intrinsic to comparing daily drought indices and not specific to SASMI, we computed the same metric between the two established reference indices themselves: DEDI and SPEI agree at essentially the same level (median ≈ 0.019) as SASMI does with either. Taken together, the correlation maps establish that SASMI captures the same drought variability as established indices, while the agreement maps show that exact day extreme co-occurrence is limited by integration window timing and reference data reliability, effects common to all daily indices. We will add both maps and this interpretation to the revised Sect. 3.2.1.

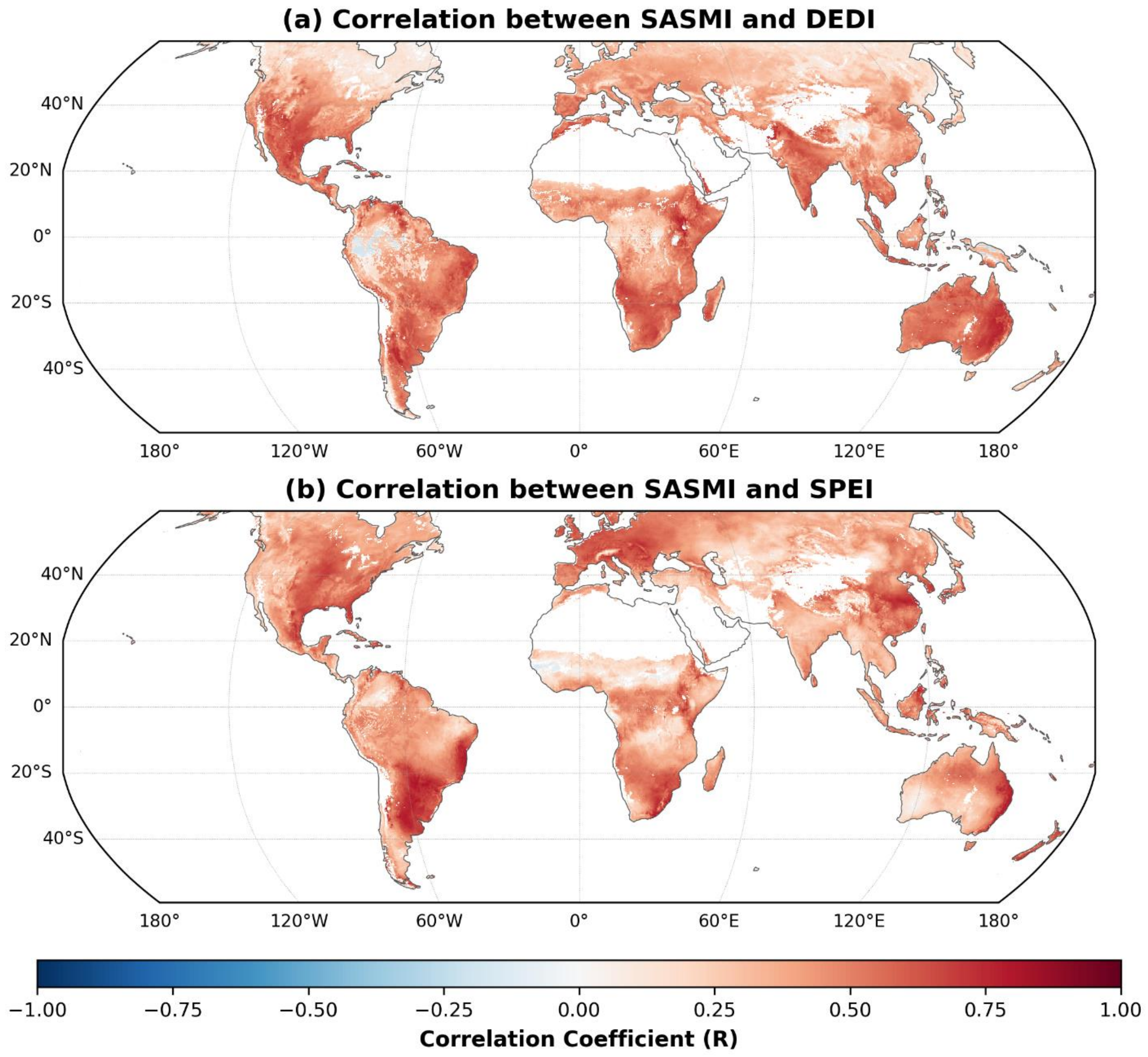


Figure 2. Spatial Spearman rank correlation between the Standardized Antecedent Soil Moisture Index (SASMI) and two established daily drought indices: (a) the Daily Evapotranspiration Deficit Index (DEDI) and (b) the Standardized Precipitation–Evapotranspiration Index (SPEI). Correlations are computed at daily resolution over the common period 2001–2020, and only statistically significant values ($\alpha = 0.05$) are shown; non-significant pixels are masked (white). Warm colours indicate positive correlation and cool colours negative correlation.

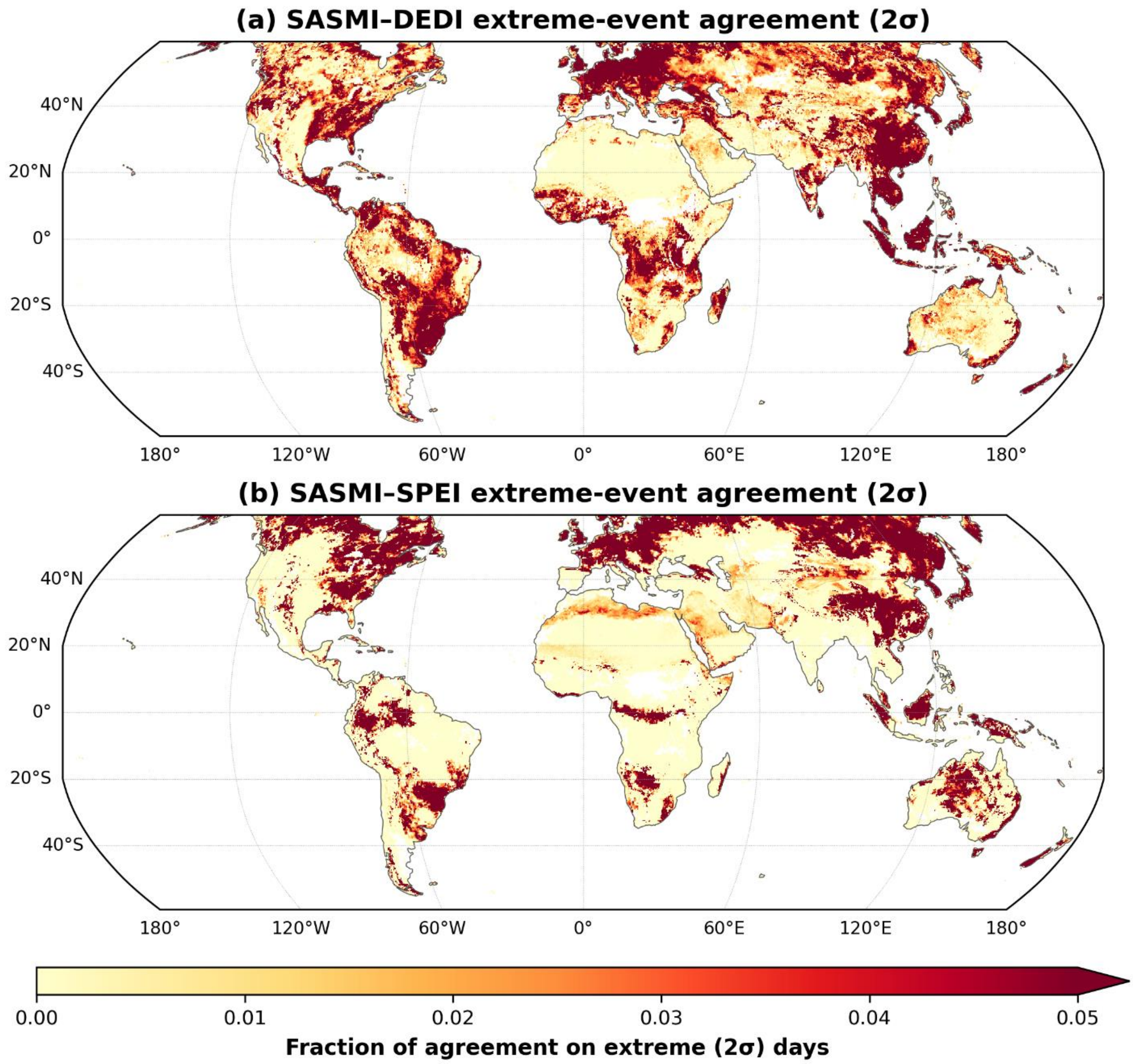


Figure S. Per-pixel fraction of agreement on extreme (2σ) days between SASMI and (a) DEDI and (b) SPEI, computed over 2001–2020 as the number of days both indices fall below −2 divided by the number of days either index falls below −2 (the threat score). Warmer colours indicate higher agreement on the timing of extreme dry days. Pixels with fewer than five extreme days in the combined record are masked (white).

Sect 3.2.2: *So the SCDHI is also a daily index? Earlier in the paper, it seemed that the daily resolution was the main novelty of SSTCI. Some more details on SCDHI (e.g., does it also use reanalysis data or is it observation based, what is the spatial/temporal resolution) would be welcome.*

This will be clarified in the revision. SCDHI (Zhou et al., 2025) combines the precipitation-based SAPEI with STI through a copula and, as published, was developed regionally over China only, using NASA NEX-GDDP-CMIP6 downscaled meteorological data. As no global compound dry–hot index was available to compare SSTCI against, the SCDHI methodology was reproduced at the global scale for the present comparison. This global SCDHI is an analysis product generated in-house (note:Zhou us co-author in the paper) for this validation rather than a pre-existing public dataset, and its construction and specifications will be described in the revised Sect. 2.1/3.2.2 if needed.

Figure 3: *Does panel d show the average of the domain in panels b/c, or a single pixel? Specify in the caption.*

Panel (d) shows the spatial mean over the Sichuan–Chongqing domain displayed in panels (b) and (c). We will specify this in the caption.

L280 / Sect 3.2.3 *(false positives, intermediate events, confusion matrix): [...] I miss insights on how the metric behaves with more intermediate (less "extreme") events. I also miss any evaluation in terms of false positives [...] It would be useful to have something like a confusion matrix for detected CDHEs w.r.t. SASMI, showing their agreement and disagreement percentages*.

We address these related points together. A CDHE is flagged only when SSTC is extreme, and SSTCI is the copula-based joint index of SASMI (drought) and STI (heat); an extreme SSTCI therefore requires both components to be jointly extreme. A false compound positive could only arise if the drought component indicated drought where none occurred, since STI is a direct temperature standardization. Our comparison against established drought indices (revised Fig. 2) shows that SASMI provides a reliable drought signal, so the drought half of each flagged CDHE is trustworthy and false compound positives are not expected.

On a confusion matrix of CDHEs with respect to SASMI: because every CDHE is by construction a subset of SASMI drought days (SSTCI < -2 requires SASMI to be extreme), such a matrix would not tell us anything beyond what the definition already implies.The relevant check is whether the drought detection itself is reliable, which the SASMI-vs-DEDI/SPEI agreement in Fig. 2 provides.

On intermediate events: SSTCI is continuous and characterizes the full severity range; the -2 threshold defines only the extreme tail retained in the catalogue. The index behaviour across the full range of dry–hot combinations is documented in the companion methodological paper as well (Aftab et al., 2026a, their Fig. 3).

Figure 4: *Are there any known events that are not captured by the other metrics, but only SSTCI? [...] it would be useful to assess the inconsistencies between the metrics as well, not only the consistencies.*

We thank the reviewer for this point, and such cases do exist. In our companion methodological paper (Aftab et al., 2026a, their Fig. 5) we compared SSTCI directly against SCDHI for two documented events over the Poyang Lake Basin. For the 2013 event, SSTCI identifies an extreme CDHE that SCDHI does not classify as extreme; for the 2019 event, SCDHI registers only a short-duration episode, whereas SSTCI resolves the full multi-week compound event. In both cases SSTCI captures conditions that SCDHI underestimates.

To assess inconsistency more systematically, we can additionally compare the compound index against its own components (SASMI and STI) if the reviewer wants, showing where a compound extreme is registered only when drought and heat coincide, as opposed to where a single driver is extreme without producing a compound event.

Sect 3.3: *The second paragraph made me wonder what exactly is the benefit of tracking the compound effect, rather than the two individual effects. A farmer for example would be mostly interested in the drought itself, I think? Perhaps this can be motivated more. Is the SSTCI perhaps in some way quantifying the positive feedback effect drought and heat have on each other?*

This is an insightful question, and the reviewer's intuition about the feedback is correct. The motivation for a compound index rather than the two individual indices rests on the fact that dry and hot extremes amplify one another rather than simply co-occurring. Miralles et al. (2014) showed that mega-heatwaves develop when heat is preceded by, and coincides with, soil-moisture depletion. The resulting compound extremes have been shown to impact ecosystems and societies far more strongly than either drought or heat alone (Zscheischler and Seneviratne, 2017; Feng et al., 2019). Quantifying the compound condition is therefore essential and cannot be recovered from a drought index or a heat index considered separately.

SSTCI is designed to capture exactly this. It is derived from the joint probability of low soil moisture and high temperature via a Frank copula, so it reaches its most extreme values specifically when dryness and heat coincide more strongly than their independent behaviour would imply. In this sense it does reflect the positive feedback the reviewer describes: SSTCI is most negative precisely in the scenarios where drought and heat reinforce one another, as in the 2010 event analysed in Sect. 3.3.

We would also note that the two component indices (SASMI and STI) are obtained as intermediate products of the methodology presented here, so a user interested only in the drought or heat signal can derive it as well as the compound one. We will expand the Sect. 3.3 motivation along these lines

L384: *"SSTCI provides a complete lifecycle analysis of the event": in the figure, it drops below −2 twice. It's not clear to me whether this would therefore be classified as a single event or not. Can the authors elaborate?*

The two dips are classified as two separate events, and this distinction is physically meaningful within the sequence described for the 2010 Voronezh event. The first, shorter excursion (~6 days) corresponds to an early compound spell in which a brief heatwave acted on soil that was already drying but not yet fully desiccated. The second, much deeper and more prolonged excursion (SSTCI approaching −3.5) is the main mega event, in which the late-July mega-heatwave coupled with by-then fully depleted soil moisture to drive the self sustaining land–atmosphere feedback. Because the two spells are separated by a sufficiently long and pronounced recovery interval, the compound deficiency accumulated between them exceeds the merging threshold, and the removal–merging procedure (Sect. 2.6) correctly retains them as two independent events rather than merging them into one.

The two events are already shaded separately in Figure 5; we will clarify the accompanying text so that the two events are described explicitly.

Sect 3.4: *Is this procedure only applicable to the SSTCI, or could it be applied to other indices for comparison?*

The procedure is not specific to SSTCI; the removal–merging optimization (Shan et al., 2024) operates on any daily standardized index time series and could be applied to other indices. Though it should be noted that this is applied here first time on Compound index to create an event catalogue We note, however, that generating such catalogues globally is computationally intensive, which is one reason we provide the SSTCI catalogue as a ready to use product.

L412: *Are there plans to identify these cases (e.g., conditions that correspond to extreme events), such that they perhaps also can be predicted? Could the SSTCI be 'forecasted' in the future using weather forecasts, providing early warnings?*

Yes, and this is now technically feasible and we thank the reviewer for the idea. SSTCI requires only two prognostic fields, near-surface soil moisture and temperature; both of which the recent operational release of ECMWF's machine-learning forecasting system (AIFS) provides, having added prognostic near-surface soil moisture (swvl1) explicitly to support drought forecasting (Lang et al., 2026). Since the standardization distributions are fitted once on the long reanalysis record and only a short antecedent window is needed at run time, SSTCI could be computed from forecast fields to provide probabilistic early warning of developing compound dry–hot conditions.

We highly appreciate this idea and would contemplate to work on in future, Thank you.

L424–425: *So the moisture/temperature indices in themselves do not show elevated probability of an extreme event post-2000?*

Lets clarify few things. The removal–merging event-identification procedure (Shan et al., 2024) was applied only to the SSTCI series, to construct the compound-event catalogue from which the post-2000 elevation is deduced; we did not apply it separately to the SASMI or STI series, so the reported increase is specifically a property of catalogued compound events rather than a claim derived from marginal event catalogues. (We also note that this event-identification procedure is computationally intensive to run globally, which is a further reason we apply it to the compound index and provide the resulting catalogue as a ready to use product.) That the individual drivers, warming temperature and declining soil moisture , have themselves intensified in recent decades is already well established (e.g. IPCC, 2021; Zscheischler and Seneviratne, 2017). Our contribution concerns the compound behaviour: the frequency of compound dry–hot events is governed not only by the individual drivers but by the strengthening dependence between them, so that their joint occurrence rises faster than the independent combination of the two marginals would imply (Zscheischler and Seneviratne, 2017), with anthropogenically forced increases in compound dry–hot frequency documented directly at global and continental scales (Zhang et al., 2022), a dependence that SSTCI captures through the copula. We realize the relevant references were omitted from this sentence, and we will add them and rephrase so that the compound specific mechanism is stated explicitly.

Figure 6: *In order to verify whether the whole SSTCI distribution has shifted to more negative values post-2000 (likely, and the hypothesis put forward here), or whether it has increased in variances*

(which would give the same results), the authors could also compare the occurrence of the "anti-CDHE" conditions, i.e., SSTCI>2. Have these decreased over time?

We thank the reviewer for this thoughtful suggestion, and we would like to clarify two points that bear on it.

First, the post-2000 increase shown in Fig. 6 is not a change in the raw daily exceedance of SSTCI < -2. The counts are the number of catalogued events produced by the removal–merging identification procedure (Sect. 2.6), which discards short spells, merges closely spaced ones, and retains only events satisfying the GEV-severity and Poisson-arrival criteria,(Figure 4 Aftab et al., 2026a can clarify this, for your reference the figure is added below) . The reviewer's shift versus variance diagnostic is framed in terms of the shape of the daily distribution, whereas our result concerns the number of extreme events; a symmetric change in distributional variance would not translate directly into this event level signal, so the raw positive tail comparison would not be a like-for-like test of our claim.

Second, the positive tail of SSTCI (> +2) a totally different physical phenomena, which are governed by a different set of land–atmosphere processes and are not the physical counterpart of a compound dry–hot extreme. Such compound conditions are an interesting subject in their own right, but they fall outside the scope of the present dry–hot dataset; we intend to examine them in dedicated follow-up work.

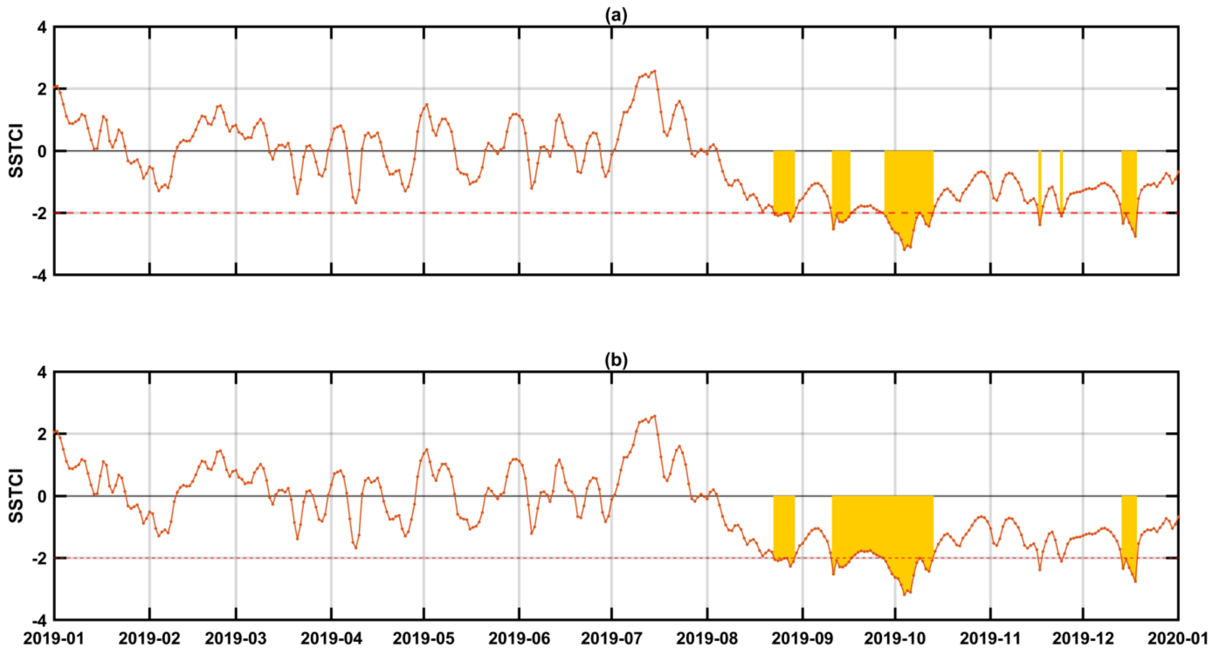


Fig. 4. Comparison of events before (a) and after (b) applying the removal and merging method. Yellow-filled bars indicate the identified extreme compound dry–hot (CDH) events

Figure 7: *I didn't understand what the slope represents. Why not show box plots like Fig. 6, but with the duration of the events rather than the count?*

We apologise for the unclear description and will rewrite it. The mapped quantity is a per-pixel temporal trend in CDHE duration: at each grid cell, the duration of every catalogued event is plotted against the calendar year in which it occurred, and a linear regression is fitted through these points across the period. Because events are discrete and irregularly spaced in time (many years contain no event), the regression is fitted over the actual event years from the first to the last event at that cell; the slope therefore expresses the rate of change of event duration over time, with a positive slope indicating that events later in the period tend to be longer.

Regarding the box-plot suggestion: duration is not well suited to the period wise box-plot format used in Fig. 6. That format works for event counts because counts has a single value at each point and it vary strongly from pixel to pixel (from a handful to several tens), giving the distribution a wide, informative spread. Event durations, by contrast, vary far less across pixels, so a period wise box plot of durations (or of per-pixel mean durations) would collapse into a narrow, low contrast distribution that obscures rather than reveals the signal, and averaging would additionally hide the temporal evolution within each period. The duration signal is fundamentally temporal, how event length changes through time and where, which is precisely what the per pixel slope map is designed to convey and what a two period box plot cannot. We therefore prefer to retain the trend map, with the clarified description above.

Technical corrections

We thank the reviewer and will make all three corrections:

- **L43:** we will define the CDHE acronym (compound dry–hot extreme event) at first use.
- **L91–92:** we will correct the section numbering for the limitations/outlook, data availability, and conclusions sections.
- **L558:** we will correct the figure reference from Figure 10 to Figure 9.