Note: This file contains response two both reviewers

==========Response to Referee comment 1 (RC1) ==========

We thank the reviewer for their comments. The comments have helped us to state the contribution of the study/dataset more precisely, and to strengthen the physical and methodological motivation for several choices that were presented too briefly in the submitted manuscript. Below, we address each comment raised by the reviewer. On the whole, we would like to note that ESSD's uniqueness criterion states that "any data set on a variable supposed or suspected to reflect changes in the Earth system deserves to be considered unique," and that its test for methodology is whether "a new or improved method should not be trivial or obvious." Compound dry–hot extremes are a variable of precisely the first kind. On threshold we hope the responses below make clear that the methodological steps involved are neither trivial nor obvious: two components of the index are fitted locally at every land pixel rather than assumed globally, the soil-moisture memory timescale τ, which drives the antecedent weighting in SASMI (Comment 3), and the copula representing the joint moisture–heat dependence, and the event catalogue is constructed through cell-by-cell threshold calibration with distributional testing rather than by thresholding the index (Comment 5)

Comment 1

In the review criteria, ESSD mentions significance as the most important aspect. With regard to uniqueness, it is stated that "it should not be possible to replicate the experiment or observation on a routine basis". The current manuscript, in my view, fails to clear this hurdle representing a more-or-less straightforward application of a standardized index to openly available reanalysis data, with poorly motivated choices that strongly impact the applicability of the dataset (see also below).

Author's response

We thank the referee for engaging directly with the significance criterion. We are not aware of any openly available dataset that offers a daily, globally complete, soil moisture-centric compound

dry–hot index over six decades together with a statistically filtered, event resolved catalogue. The compound indices we know of are regional, cover shorter periods, are monthly or seasonal, or rest on precipitation based drought proxies rather than soil moisture. It is that combination that is missing. The two products also serve different users: the continuous SSTCI field suits process level and attribution research, which requires a daily signal on a common grid, while the event catalogue provides a discrete, comparable inventory that impact and risk studies can use directly.

On whether the construction is routine, we would note that at each stage the index is parameterised locally rather than by a global assumption. The dry down timescale τ that sets the antecedent weighting in SASMI is estimated at every land pixel from that pixel's own observed dry-downs; the copula linking the moisture and temperature components is fitted per pixel, so the dry–hot dependence carries the hydroclimatic character of each location; and the removal and merging thresholds underlying the catalogue are searched independently at every cell and accepted only where the resulting events pass both distributional tests. Our companion paper (Aftab et al., 2026a) demonstrated the SSTCI methodology at a single basin using one uniform τ, which is a reasonable simplification for one homogeneous region but does not transfer globally, the physical basis for local fitting is set out in our responses to Comments 2 and 3.

These costs compound: τ fitted at every land pixel, a copula fitted at every pixel, the index evaluated daily across 63 years, and event identification with two goodness-of-fit tests run cell by cell at 0.1° over the global land surface. Reproducing the product is not a matter of running a standardized index over ERA5-Land, and sparing users that burden is part of what the dataset is for. We would also note that ESSD has recently published Dheed (Weynants et al., 2025), a global catalogue of dry and hot extreme events derived entirely from ERA5, which suggests the significance criterion turns on what a product delivers rather than on whether its inputs are public. We release our code alongside the data.

Author's changes in manuscript

Introduction, contribution paragraph rewritten, L96–101:

"Consequently, a continuous, consistent global daily scale index is needed to capture the swift onset, intensification, and decay dynamics characteristic of CDHEs.

In this study, we present the Standardized Soil Moisture Temperature Compound Index (SSTCI), a global, daily (0.1° × 0.1°) dataset from 1961–2023. SSTCI integrates a standardized antecedent soil moisture index (SASMI), calculated with grid-specific dry down memory of SM, and a daily standardized temperature index (STI) via a Frank copula, fitted per cell, through which the moisture and temperature signals are joined so that the dry–hot dependence carries the hydroclimatic character of each location."

And L103–109:

"Furthermore, to facilitate immediate application for risk and impact assessment, we derived a comprehensive Global CDHE Event Catalogue from the daily SSTCI series. By applying a rigorous grid level iterative optimization technique adopted from Shan et al. 2024, in which removal and merging thresholds are searched independently at every land grid cell and retained only where the resulting events are statistically independent and belong to a genuinely extreme population, we transformed the continuous signal into a precise inventory of statistically independent, extreme events. The catalogue eliminates the computational burden of processing raw time series, providing users with exact start dates, end dates, durations, and severity metrics for every compound event detected globally over the last 63 years."

Sect. 2.6, L254–259:

"Because both the threshold search and the two distributional tests are evaluated cell by cell, applying this procedure across the global land surface at 0.1° amounts to an iterative optimization with goodness-of-fit verification at several million grid cells over a 63-year daily record; its computational cost is one reason we provide the resulting catalogue as a ready to use dataset. For details of the cataloguing approach, readers are referred to Shan et al. (2024) and Aftab et al. (2026a)."

Comment 2

In their approach, the authors use soil moisture averaged over the top 7 cm of the soil profile. This seems rather arbitrary, as there is no physical reason why the top 7 cm soil moisture would influence air temperatures more than the top 2 cm or 10 cm, or even deeper. In many more humid

climates, soil moisture averaged over the root zone would be the best representation of soil moisture impacts on T through land-atmosphere interaction, since plant transpiration (of root water uptake) is the largest component of terrestrial ET during dry conditions. Deeper soil moisture would have considerably longer e-folding timescales that more closely follow timescales involved in soil moisture-temperature coupling. The authors do not provide a physical justification of this arbitrary choice, nor do they provide a sensitivity analysis. Any other arbitrary choice here would have resulted in different values of the index. The fact that the authors show a (weak) positive correlation with other indices around most of the globe is not convincing, as almost any other choice would likely do the same because of the strong impact of P.

Author's response

We thank the referee for this important point and welcome the opportunity to clarify the reasoning

behind the choices. The choice follows from what the index is designed to resolve, and we accept that the manuscript did not say so. SSTCI targets the daily scale onset and evolution of compound dry–hot conditions and the rapid land–atmosphere coupling that drives them. For that target the near-surface layer is where the coupling is most directly expressed: it controls direct soil evaporation and is most tightly tied to the daytime surface energy balance governing latent–sensible partitioning, which is the mechanism by which moisture deficits amplify near-surface heat. This is consistent with our use of daytime (06:00–18:00 local time) fields, chosen to capture the window in which the coupling is strongest. The 0–7 cm layer is also the standard and most extensively validated near-surface layer of ERA5-Land, which supports comparability with the wider literature.

On the referee's point about e-folding timescales, we would answer differently. A shallow state variable does not leave the index blind to longer timescale moisture conditions, because that is precisely what the SASMI component provides: the antecedent weighting with a grid specific memory timescale carries the accumulated moisture history into the index, while the state variable itself retains the fast response needed to resolve sub-weekly transitions. The formulation seeks memory and responsiveness together rather than trading one against the other, and the depth choice should be read alongside it rather than in isolation.

We nonetheless take the reviewer's point seriously, and our formulation does not claim to capture deep vegetation water stress, which is optimized for rapid surface coupling and near-surface conditions. This is already stated in Sect. 3.9, where we note that reliance on top-layer (0–7 cm) soil moisture may not fully capture deeper root-zone constraints that govern vegetation stress and drought persistence, and where we identify root-zone incorporation as a priority for future development. We will make this trade-off more prominent in the revision and would note that the framework is fully extensible in this direction. ERA5-Land provides soil moisture at multiple depths on the same grid, so a root zone or multi layer variant of SSTCI can be constructed directly within a similar framework, as the authors believe (at least from preliminary analyses) that subsurface development would require critical parameterization of the framework. Additionally, a systematic comparison across depths is a natural next step that the present, near-surface product is designed to enable. We also note that it was important for the authors to first receive community feedback based on the current form, based on which we would further develop the sub-surface products, which is currently being developed. Here, we are following the principled used by other major groups such as the H-CEL's GLEAM which started from the surface to the root-zone, combining lessons from the community's feedback.

Author's changes in manuscript

Sect. 2.1, L120–129.

"ERA5-Land provides multi-decadal, globally complete hourly fields and has been shown to realistically capture surface soil moisture variability and anomalies across a wide range of hydroclimatic regimes when evaluated against large in situ networks (Hagan et al., 2020; Beck et al., 2021). The 0–7 cm layer is used because SSTCI targets the daily-scale onset and evolution of compound dry–hot conditions, for which the near-surface layer is where land–atmosphere coupling is most directly expressed: it controls direct soil evaporation and is most tightly tied to the daytime surface energy balance that governs latent–sensible partitioning. It is also the layer routinely evaluated in global soil-moisture product inter-comparisons (Beck et al., 2021; Muñoz-Sabater et al., 2021), which supports comparability with the wider literature. Longer-timescale moisture conditions are not lost through this choice, since the antecedent weighting in SASMI

(Sect. 2.3) carries the accumulated moisture history into the index while the state variable itself retains the fast response needed to resolve sub-weekly transitions."

Sect. 3.9, L684-690:

"Finally, the implementation relies on top-layer soil moisture (0–7 cm). This layer is where land–atmosphere coupling is most directly expressed and is therefore appropriate to the daily-scale signal SSTCI targets, but it does not represent the deeper root-zone constraints that govern vegetation stress and drought persistence; the index is optimized for rapid surface coupling rather than deep vegetation water stress and should be interpreted with that scope in mind."

Sect. 3.9, L696:

"Furthermore, since ERA5-Land provides soil moisture at multiple depths on the same grid, a root-zone or multi-layer variant can be constructed within the present framework, and a systematic comparison across depths is a natural next step that this near-surface product is designed to enable."

Comment 3

The authors calculate the soil moisture part of their index based on SASMI which is using an e-folding timescale. The added value of this approach is not made clear. What is the benefit of using this method over using any other possibility, including using SSMI calculated using soil moisture from ERA5-Land directly? The choice again feels arbitrary, and in my opinion results in less realistic values than alternative choices based on available data or model output.

Author's response

We answer this comment together with the related point raised in the referee's follow-up, where he asks why the treatment of drought should differ from widely accepted approaches including SPI, SPEI and SSMI.

The added value is physical rather than procedural, and we agree it was stated too briefly. A plain SSMI characterises the soil-moisture state on a single day, a snapshot of the anomaly. Compound

dry–hot coupling, however, is not driven by the moisture state on any one day but by accumulated antecedent depletion over the preceding period: it is the progressive dry-down that suppresses evaporative cooling and shifts surface energy partitioning toward sensible heat, preconditioning the surface for heat build-up. Antecedent memory is therefore not an optional refinement but the mechanism through which soil moisture modulates temperature, and a memoryless snapshot omits precisely the quantity that governs the coupling. The 2010 Russian mega-event (Fig. 5) shows this directly: a sustained early summer soil-moisture decline precedes and physically enables the subsequent heat extreme, a sequence an instantaneous index would not resolve as a compound signal.

The follow-up comment sharpens this, since SPI, SPEI and SSMI are often computed at weekly and monthly timescales as well as daily. The relevant contrast is therefore not with an instantaneous snapshot but with a fixed accumulation window, and our answer is that the difference lies in whether the weighting within that window is imposed uniformly or set by the land surface itself. SPI, SPEI and multi-scale SSMI accumulate over a window chosen by the analyst and applied identically everywhere — SPEI-30 means thirty days in the Sahel and thirty days in the Congo, weighted equally throughout. That is appropriate for a general-purpose drought measure, and it is why we use SPEI in exactly that role as a reference. It is less appropriate when the drought component must represent the antecedent depletion driving land–atmosphere coupling, because the timescale over which depletion accumulates is a property of the land surface: τ falls below 5 days across arid and semi-arid regions and reaches or exceeds 20 days in the Amazon, Central Africa and maritime Southeast Asia, with temperate and managed regions intermediate at roughly 7–15 days (Fig. 1), consistent with the aridity and land-surface controls documented by McColl et al. (2017). A fast-draining semi-arid cell and a slowly buffered humid cell do not share a memory timescale.

SASMI accordingly replaces uniform weighting with an exponential decay whose constant is estimated per grid cell from that cell's own observed dry-downs. The fourteen-day window is common to all cells and bounds the summation; within it, the grid-specific decay constant determines how much of that history actually contributes, so the effective memory is short where soils drain quickly and longer where they are buffered. The window is a computational bound rather than the quantity that sets the timescale, which is the respect in which the formulation differs

from a fixed accumulation index. We note that in the most strongly buffered cells the fourteen-day bound truncates contributions that are still appreciable, so the effective memory there is shorter than τ alone would imply; extending the window in such regions is a refinement, though it would increase the computational cost of the global product We would emphasise that we are not proposing a better general purpose drought index and would not argue SASMI should displace SPEI or SSMI for general drought monitoring. Our claim is narrower: a compound dry–hot index requires a drought component matched to the physical timescale of the coupling it represents, and a spatially uniform weighting cannot provide that.

On the concern that SASMI may yield less realistic values, our evaluation does not support that reading. SASMI agrees with two independent established daily drought metrics, SPEI and DEDI, across most global land areas (Fig. 2, now rank-based and extreme value agreement following Referee #2's suggestion). We take this as evidence that the antecedent weighting preserves drought signal fidelity, it recovers the behavior those metrics identify rather than distorting it, while operating directly on soil moisture at each cell's own memory timescale, which is what the compound coupling requires and what neither reference metric provides.

Author's changes in manuscript

Introduction, L89–94:

"Working directly with soil moisture is necessary but not sufficient, since the coupling is driven by accumulated antecedent depletion rather than by the moisture state on any single day. An instantaneous standardized soil moisture index describes a snapshot and omits the progressive dry-down through which soil moisture modulates temperature, and the timescale over which that depletion accumulates varies widely between fast drying semi-arid and slowly buffered humid environments (McColl et al., 2017; Rahmati et al., 2024). A soil moisture centric compound index therefore requires both a direct soil moisture state variable and a location-specific representation of its memory along with temperature extremes. Consequently, a continuous, consistent global daily scale index is needed to capture the swift onset, intensification, and decay dynamics characteristic of CDHEs."

Sect. 2.3, L175–178:

"So, effective weights differ by region: a fast draining semi-arid cell retains only a few days of memory, while a slowly buffered humid cell integrates over most of the window."

Sect. 2.3, added after L174-180:

Where SM_{t-n} is daily soil moisture *n* days prior, N = 14 is the length of the antecedent window and $k_{grid} = e^{-1/\tau_{grid}}$ is the grid-specific decay constant (*k*) derived from the local dry-down timescale τ_{grid} (Sect. 2.2).Contributions decay geometrically as $1, k^1_{grid}, k^2_{grid}, \ldots k_{grid}$. So, effective weights differ by region: a fast draining semi-arid cell retains only a few days of memory, while a slowly buffered humid cell integrates over most of the window. Unlike SPI, SPEI and standardized soil moisture indices, which weight every day in the accumulation window equally, SASMI therefore weights antecedent days by each cell's own dry-down rate. The fourteen-day window balances responsiveness with the persistence typical of short-term droughts (Christian et al., 2024) and bounds the summation, while the decay constant determines the weighting within it."

Sect. 3.1, L279:

"This heterogeneity directly supports our modeling choice to move away from a single global memory parameter (Rahmati et al., 2024). Thus, translating τ into a grid-specific decay constant $k=e\text{-}1/\tau$ allows antecedent moisture weighting (and thus SASMI) to adapt to local dry-down rates rather than imposing uniform persistence everywhere."

Comment 4

The authors claim that "Yet, most existing compound indices are provided at monthly or seasonal scales". This does not really do justice to the large body of literature on multivariate indices for heat-drought extremes. I refer here to Miralles et al, 2012 (doi:10.1029/2012GL053703) and Shan et al. 2024 (https://doi.org/10.5194/hess-28-2065-2024) as examples of earlier and recent studies, but many other studies exist. The novelty of the study should be motivated better based on a more thorough literature review.

Author's response

We agree the sentence was framed too broadly and have qualified it, and we have expanded the Introduction to give a fuller account of both the process oriented and the multivariate statistical literature. Two clarifications on the specific references.

Miralles et al. (2012) is a foundational contribution on soil-moisture–temperature coupling and the role of soil desiccation and atmospheric heat accumulation in mega-heatwaves. It reinforces the physical rationale for a soil-moisture-centric compound index, and we already drew on the closely related Miralles et al. (2014). We would note that this line of work characterises the coupling mechanism rather than delivering a multivariate compound index or an openly available compound dataset, and that its aim was to understand the soil-moisture–temperature feedback during heatwaves rather than to extend to droughts , the formalism of Π requires a strong positive correlation of soil-moisture deficit with air temperature to be significant, so it is not optimal for tracking the full evolution of a compound event(*personal correspondence with Diego Miralles*). It therefore motivates our product rather than overlapping with it, and we now use it in that role.

Shan et al. (2024) is closely related and we adopt its removal–merging event identification procedure. It does not, however, provide a compound index in the integrated sense that SSTCI does: drought and heat are represented there as two separate standardized series, and compound events are defined through their temporal set relationships rather than combined into a single quantity. SSTCI instead integrates the soil-moisture-based drought signal and the temperature signal into one continuous standardized compound variate through a copula, so that the joint dry–hot dependence is carried within the index itself. The two approaches are complementary, and we now say so explicitly.

Taken together, these allow us to position the contribution more precisely: to our knowledge no existing openly available product combines a daily, global, multi-decadal, soil moisture-based compound index with an objectively constructed event catalogue.

Author's changes in manuscript

Introduction, L47–56:

"Process oriented studies have characterized the soil moisture–temperature coupling that underlies compound dry–hot conditions (Miralles et al., 2012), showing how soil desiccation and atmospheric heat accumulation combine to produce mega-heatwaves (Miralles et al., 2014). In parallel, multivariate statistical approaches have sought to quantify compound conditions directly. Among these, the copula function has proven particularly effective in capturing complex dependencies between climate variables, enabling the construction of joint distributions that integrate multiple indicators into bivariate or even multivariate indices (Hao et al., 2019a, 2019b; Wu et al., 2020; Wu et al., 2021; Wang et al., 2023). Others define compound events through the temporal overlap of separately standardized drought and heat series rather than by combining them into a single quantity (Shan et al., 2024)."

Miralles et al. (2012) added to the reference list; Zscheischler et al. (2018) cited

Comment 5

The authors claim that a main novel aspect in their work is the catalogue of event timing and magnitude per location. I do not consider this particularly novel, since it is rather straightforward to do once the index has been defined and calculated. A more novel aspect would be to define unique spatio-temporal clusters of heat-drought events, and catalogue these rather than the just the temporal dynamics for each grid cell.

Author's response

The catalogue is not a thresholding of the index, and we accept that the manuscript did not make the distinction clear enough.

Events are identified through an objective removal–merging procedure in which minor spells are removed and mutually dependent spells merged, with thresholds chosen so that the retained events are statistically independent (exponentially distributed inter-arrival times) and belong to a genuinely extreme population (GEV-distributed severities). These are not global settings: the thresholds are searched and the distributional tests evaluated cell by cell, so every land pixel receives its own calibrated criteria and its own goodness-of-fit verification. Across the global land surface at 0.1° this amounts to an iterative threshold search plus two distributional tests at each of

several million grid cells over a 63-year daily record. We adopt the procedure from Shan et al. (2024) but apply it to the compound SSTCI variate rather than to separate drought and heat series, so the independence and extreme-value tests are applied to the compound quantity itself rather than to its marginals, a distinction that matters, since well-behaved marginals do not guarantee that events assembled from them are independent or extreme as compound events. What this yield is a product that is operationally ready in a way a raw thresholding list is not. Thus, for every land grid cell globally, a direct record of discrete events with explicit onset, termination, duration and marginal severity, the form of information most relevant to impact assessment and early-warning applications. We have attached figure4 from Aftab et al,2026a, below for your convenience, this figure explains the merging-removal method application for 2019 CDHEs in Poyang lake basin, China.

we thank the referee for the suggestion. We would note that a similar product also exists (published in ESSD), where Dheed (Weynants et al., 2025) derive spatio-temporal clusters of dry and hot events from ERA5, and that the two constructions are in any case complementary rather than alternatives. Spatio-temporal clustering operates on per grid cell event records, which are the elements from which coherent event objects are assembled. The present catalogue is therefore not a lesser substitute for clustering but a natural input to it, and one in which the constituent events have already been filtered for independence and extremeness before any clustering is applied. Extending it in this direction involves consequential choices, spatial connectivity criteria, temporal overlap rules, minimum cluster size, that merit dedicated treatment rather than being appended here, and we will note it as a promising direction in the outlook.

We also believe that this allows the data to be easily and readily usable by non-scientific users who would need this for other applications beyond scientific research and operational services

Author's changes in manuscript

Sect. 2.6, L243–246:

"We then employed a systematic optimization loop to determine the removal and merging thresholds. These are not global settings: the threshold search is carried out independently at every

land grid cell, so that each pixel receives its own calibrated pair of criteria reflecting its local event structure."

Sect. 2.6, added after L233-:236

"To our knowledge this procedure has not previously been applied to a compound dry–hot index. Shan et al. (2024) developed it for separately identified drought and heatwave series, whereas here it operates on the joint compound signal, so that the independence and extreme-value tests are applied to the compound quantity itself rather than to its marginals."

Sect. 2.6, L254-259:

"Because both the threshold search and the two distributional tests are evaluated cell by cell, applying this procedure across the global land surface at 0.1° amounts to an iterative optimization with goodness-of-fit verification at several million grid cells over a 63-year daily record; its computational cost is one reason we provide the resulting catalogue as a ready to use dataset."

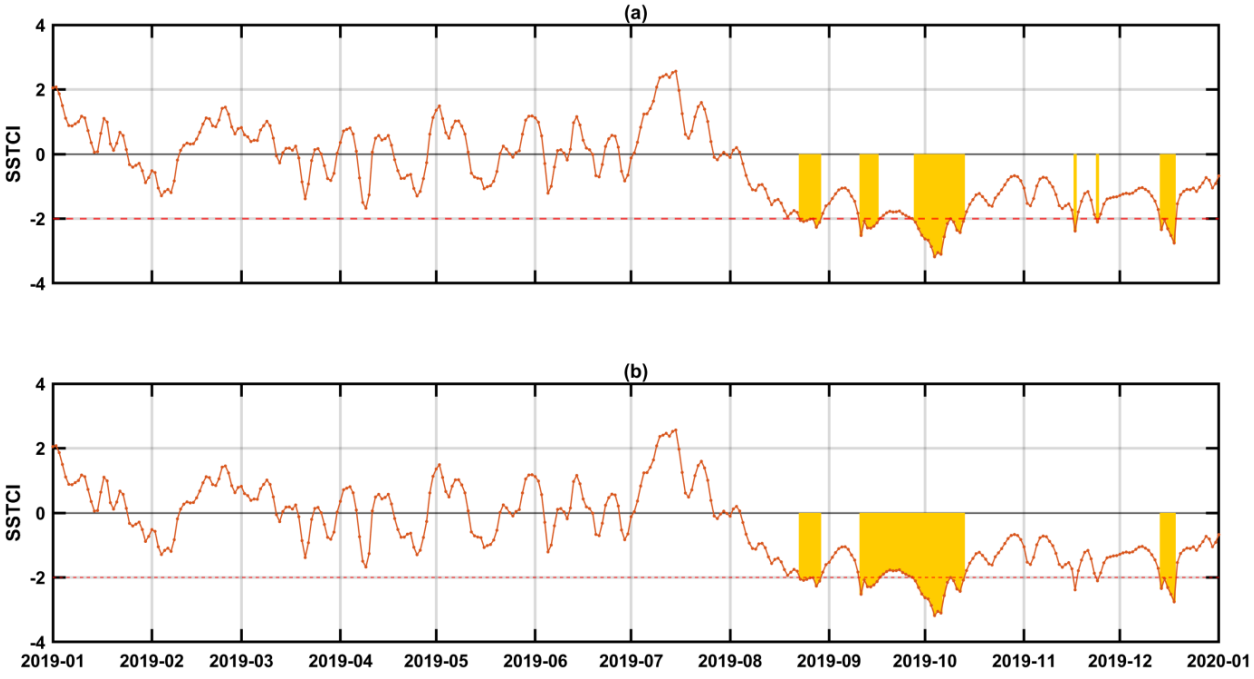


Fig. 4. Comparison of events before (a) and after (b) applying the removal and merging method.

Yellow-filled bars indicate the identified extreme compound dry–hot (CDH) events.

Comment 6

Given that many different indices exist, I would see potential value in a large ensemble of heat-drought indices, which would also allow quantification of some of the uncertainties involved through the ensemble spread. Such dataset would, in my view, classify as significant.

Author's response

We agree that a multi-index ensemble quantifying uncertainty through ensemble spread would be valuable, and we have added the direction to the outlook.

We would note, however, that such an ensemble presupposes well-documented, validated single-index products to combine. Members must be openly available, standardized and provided on a common grid before any meaningful spread can be computed, and it is precisely this that is lacking for daily, global, soil moisture based compound indices. SSTCI is released openly with its code, expressed as a standardized compound variate, and provided on a uniform global daily grid over 1961–2023, properties that make it directly usable as an ensemble member rather than only as a standalone product. We would suggest the two things are not in contradiction: providing a well-characterized single product is how the community comes to have an ensemble at all.

Author's changes in manuscript

Sect. 3.9, L698–700:

"Released openly on a uniform global daily grid, SSTCI is likewise usable as a member of a multi-index ensemble for quantifying uncertainty in compound dry–hot characterization."

Closing note on Referee #1's follow-up comment

We are grateful for the referee's follow-up and have treated its central instruction, that the reasoning belongs in the manuscript rather than in a response, as the organizing principle of the revision. The arguments given above now appear in the paper itself:

the contribution in the Introduction (L96–101), the physical justification for the near-surface layer in Sect. 2.1 (L120–129) with the corresponding limitation in Sect. 3.9 (L684–690), the rationale for antecedent memory and its contrast with fixed-window drought indices in the Introduction (L89–94) and Sect. 2.3 (L175–178), the expanded literature framing in the Introduction (L47–56), the cell-by-cell construction of the catalogue in Sect. 2.6 (L233–259), and the clustering and ensemble directions in Sect. 3.9 (L698–700).A reader who never sees this response will now encounter the reasoning in the paper.On the sensitivity of the index to the choice of soil layer, we would note that a like-for-like test is not straightforward, since a sub-surface variant requires its own parameterization of the dry-down timescale rather than a substitution of input layer. We have therefore set out the physical basis for the near-surface choice explicitly in Sect. 2.1 and its scope limitation in Sect. 3.9 and treat the systematic cross depth comparison as the natural next development of the framework, one that the release of this near surface product on a uniform global grid is intended to enable, since ERA5-Land provides soil moisture at multiple depths on the same grid.

References

Weynants, M., Ji, C., Linscheid, N., Weber, U., Mahecha, M. D., and Gans, F.: Dheed: an ERA5 based global database of compound dry and hot extreme events from 1950 to 2023, Earth Syst. Sci. Data, 17, 6621–6645, https://doi.org/10.5194/essd-17-6621-2025, 2025.

==========**Reviewer2**==========

Response to Referee Comment 2 (RC2)

Referee comments are reproduced in italics. Each is followed by the author's response and by the corresponding changes in the manuscript, with revised text quoted and line numbers referring to the revised manuscript.

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 hopefully clear a lot of confusions.

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.

Author's response

We thank the reviewer for this clear summary and have restructured the revision around the three points raised. We have sharpened 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.

The methodology is now explained in detail. 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 have 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 have moved the wavelet analysis (Sect. 3.6) to the Supplement. The 2022 Yangtze reconstruction (Sect. 3.8) has been retained in the main text; we explain this choice in an accompanying letter to the editor and defer to the editor's judgement on it. The detailed responses below describe each change

---

Specific comments

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

Author's response

We agree and have rephrased the abstract so that the procedure is described in operational terms before the technical name is used. The full technical description is retained in Sect. 2.6.

Author's changes in manuscript

Abstract, L26–29.

“To move from daily index values to physically interpretable hazards, we post-process the daily SSTCI series by discarding very short extreme spells and merging those separated by only brief interruptions, so that fragmented detections become single coherent events; the resulting catalogue carries event timing and magnitude descriptors suitable for event-based statistics.”

L43: define the CDHE acronym at first use.

Author's response

Corrected. The acronym is expanded at its first occurrence (L44).

Author's changes in manuscript

Introduction, L44.

The authors assume the reader already knows all these indices/acronyms (SCDHI, SAPEI, STI, DCDHI, …).

Author's response

Each acronym is now expanded at first use and accompanied by a brief clause identifying what the index measures and how it is constructed, so that no prior familiarity is assumed.

Author's changes in manuscript

Introduction, L62–68:

"For instance, compound dry–hot conditions have been characterized at sub-monthly scales using the Standardized Compound Dry–Hot Index (SCDHI), which combines the Standardized Antecedent Precipitation Evapotranspiration Index (SAPEI) with the Standardized Temperature Index (STI) through a copula (Li et al., 2021); the Daily Compound Dry–Hot Index (DCDHI), a daily-scale composite of precipitation, evapotranspiration and temperature anomalies (Wang et al., 2023); and a daily refined SCDHI (Zhou et al., 2025) that incorporates SAPEI (Jiang et al., 2022) with a CO_2-aware modified Penman–Monteith formulation for potential evapotranspiration (Yang et al., 2019)."

"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.

Author's response

We have made the reasoning explicit. The drought components of existing sub-monthly compound indices are derived from precipitation-based proxies that infer moisture status from the atmospheric water balance rather than from soil moisture directly. Such proxies neither represent antecedent soil-moisture memory explicitly nor accommodate the regional heterogeneity of dry-down rates and can therefore misrepresent the soil moisture state that drives land–atmosphere coupling during compound events. As the referee notes at L72–78, this argument was previously made twice; it is now made once, in full, at this location.

Author's changes in manuscript

Introduction, L68–73:

"These indices are typically limited in spatial scope (mostly are regionally computed), and their drought components are derived from precipitation-based proxies that 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 and can therefore misrepresent the soil moisture state that drives land–atmosphere coupling during compound events (Seneviratne et al., 2010)."

A complementary passage at L89–96 explains why working with soil moisture alone is necessary but not sufficient, motivating the antecedent-memory formulation.

Author's changes in manuscript

L89-96

“Working directly with soil moisture is necessary but not sufficient, since the coupling is driven by accumulated antecedent depletion rather than by the moisture state on any single day. An instantaneous standardized soil moisture index describes a snapshot and omits the progressive dry-down through which soil moisture modulates temperature, and the timescale over which that depletion accumulates varies widely between fast drying semi-arid and slowly buffered humid environments (McColl et al., 2017; Rahmati et al., 2024). A soil moisture centric compound index therefore requires both a direct soil moisture state variable and a location specific representation of its memory along with temperature extremes. Consequently, a continuous, consistent global daily scale index is needed to capture the swift onset, intensification, and decay dynamics characteristic of CDHEs”

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?

Author's response

The referee is right that the statement was too categorical, and it contradicted our own choice of reference datasets. Soil moisture anomaly products and evapotranspiration-based indices such as DEDI are widely used and are not precipitation-based. Our intended claim was narrower: that the standardized drought indices most commonly adopted as inputs to compound indices remain

precipitation-driven. The sentence has been rewritten to make only that narrower claim and to acknowledge the alternatives explicitly.

Author's changes in manuscript

Introduction, L82–87:

“Soil moisture widely used and provide direct measures of surface moisture conditions, however many of the standardized drought indices adopted to compute compound indices remain precipitation driven, such as the Standardized Precipitation Index (Liu et al., 2024),,the Standardized Precipitation–Evapotranspiration Index (Vicente-Serrano et al., 2010), the Palmer Drought Severity Index (Zhong et al., 2019), and the Standardized antecedent precipitation evaporation index (Li et al., 2020).”

Repetition of previously made points (temporal scale and precipitation vs soil moisture focus).

Author's response

The redundant restatement has been deleted; the argument is now made once, at L67–72.

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.

Author's response

The description was ambiguous and has been corrected and expanded. The two variables are not the same quantity. ERA5-Land provides only the instantaneous 2 m air temperature (2t). ERA5's mx2t is defined by ECMWF as the maximum 2 m temperature since the parameter was last archived i.e. an hourly maximum, not an instantaneous value. Because heat extremes and their

impacts are governed by daytime peak temperature, and because maximum daytime temperature is the variable used in the large majority of operational and research heatwave definitions (Perkins-Kirkpatrick and Lewis, 2020), a maximum-based variable is required. ERA5-Land does not provide one; ERA5 does. We therefore accept the interpolation from 0.25° to 0.1° as the cost of using the physically appropriate variable, while soil moisture is taken from ERA5-Land at its native 0.1°, where the land-surface fields carry demonstrable added value over ERA5 (Muñoz-Sabater et al., 2021).

We have also stated explicitly what daily quantity is formed from the hourly field, which we believe was the source of the confusion: the daily value is the mean of the hourly maxima over 06:00–18:00 local time, i.e. a daytime-averaged maximum temperature representing sustained peak daytime heat.

Author's changes in manuscript

Sect. 2.1, L130–136:

"Temperature is obtained from ERA5 as the hourly maximum 2-m air temperature (mx2t), defined by ECMWF as the highest 2-m air temperature since the parameter was last archived, i.e. an hourly maximum rather than an instantaneous value. Because heat extremes and their impacts are governed by daytime peak temperature, a maximum-based variable is required (Perkins-Kirkpatrick and Lewis, 2020); ERA5-Land provides only the instantaneous 2-m temperature (2t), whereas ERA5 provides mx2t. We therefore source temperature from ERA5 at its native 0.25° resolution and bilinearly interpolate it to the 0.1° ERA5-Land grid, accepting the interpolation as the cost of using the physically appropriate variable."

And L139–141:

"Daily values were derived as the mean of hourly data between 06:00 and 18:00 local time, computed after converting from UTC using longitude-based offsets. For temperature this yields a daytime-averaged maximum temperature representing sustained peak daytime heat rather than an instantaneous value."

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, …).

Author's response

Corrected. The subscript is applied consistently, and every symbol is now defined where the equation first appears. Following the referee's related comment on Sect. 2.5, we define symbols inline rather than in a separate notation appendix, so that each definition sits where it is needed.

Author's changes in manuscript

Sect. 2.3, L186–190:

"where F_{LL} is the cumulative distribution function of the fitted log-logistic distribution and Φ^{-1} is the inverse standard normal (quantile) function. Negative (positive) SASMI values therefore denote drier (wetter) than normal conditions for that grid cell on that day of the year. Because the reference distribution is day-of-year specific, the seasonal cycle is already accounted for, and a dry extreme arises only when D_t is unusually low relative to other years on the same calendar day."

Symbols in Eq. (2) defined.

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?

Author's response

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.

Equation (2) defines the antecedent signal D_t, a memory-weighted sum of soil moisture over the preceding 14 days with weights decaying at the grid-specific dry-down rate. This is an input feature, not an anomaly. Equation (3) then converts D_t into an anomaly by referencing it against its own climatology: for each grid cell, a log-logistic distribution is fitted to the D_t values for a given calendar day of the year across all 63 years of the record, exactly as the normal distribution is fitted per calendar day for STI. D_t is mapped to its percentile within that day-of-year reference distribution and transformed to a standard normal variate.

"Drier than normal" therefore 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, and the full record is used. It follows directly that 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 absorbs the seasonal cycle. A dry extreme arises only when D_t is unusually low relative to other years on the same calendar day.

Author's changes in manuscript

Sect. 2.3, L181–190:

"D_t is an input feature rather than an anomaly at this stage. For each grid cell, a log-logistic distribution is fitted to the values of D_t for a specific calendar day of the year across all years of the record (1961–2023), in the same way that 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 reference distribution and transformed to a standard normal variate. […] Because the reference distribution is day-of-year specific, the seasonal cycle is already accounted for, and a dry extreme arises only when D_t is unusually low relative to other years on the same calendar day."

Replace "each calendar day" by "a specific calendar day"?

Author's response

Changed as suggested, and the same phrasing is used in Sect. 2.3 for consistency.

Author's changes in manuscript

Sect. 2.4, L197: "x represents the temperature series for a specific calendar day".

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.

Author's response

We are grateful to the referee for catching the mis-citation. The pointer to Aftab et al. (2026c) was an error: that identifier is the Zenodo data record. The intended reference throughout is our companion methods paper, Aftab et al. (2026a, *Int. J. Climatol.*, doi:10.1002/joc.70306). This has been corrected wherever it occurred.

Author's changes in manuscript

Sect. 2.5, L209–216:

Where X and Y denotes SASMI and STI variables respectively, x and y are their thresholds values, u = F_X(x) v = F_Y(y), and C (u,v) is the fitted Frank copula. To make the cumulative joint probability (p) spatially comparable, it is transformed to a standard normal variate (Equation 7).

$$SSTCI = \varphi^{-1}(F(p)) \qquad (7)$$

where F is the cumulative distribution of p over the full record at each grid cell, and Φ^{-1} is the inverse standard normal (quantile) function. STCI provides a unified metric.

Citation corrected at L226 (Sect. 2.5) and L259(Sect. 2.6) from Aftab et al. (2026c) to Aftab et al. (2026a).

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 behavior.

Author's response

The reviewer is correct on both points, and we verified the standard-normal behaviour directly: pooled over all land pixels and days, SSTCI is indistinguishable from N(0,1) (mean = 0.0002, s.d. = 1.0003; P(SSTCI < −2) = 2.271% against the theoretical 2.275%). We will add this as a supplementary figure if needed (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 of the products.

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, standardization fixes the exceedance rate over the full 1961–2023 record but places no constraint on how those days are distributed in time, nor on the number of coherent events they form. What varies in space is therefore not the daily exceedance frequency but the event structure produced by the removal–merging catalogue (Sect. 2.6), the number, duration and clustering of events, which is the actual hazard product and is not pinned to 2.3%. Retained event counts range

per pixel across the regions examined in Sect. 3.5, which could not arise if a fixed exceedance rate propagated to a fixed event rate.

We have added a sentence to Sect. 2.5 reporting the pooled statistics and connecting the standardization step to the event identification procedure in Sect. 2.6. The revised text reads:

Author's changes in manuscript

Sect. 2.5, L219-226

"Pooled over all land pixels and days, SSTCI is closely standard normal (mean 0.0002, standard deviation 1.0003, P(SSTCI < −2) = 2.271% against a theoretical 2.275%). Standardization constrains the exceedance rate over the full 1961–2023 record; it places no constraint on how those days are distributed in time within the record, nor on the number of coherent events they form, which is determined by the cell-wise removal–merging procedure described in Sect. 2.6."

On the histogram: we verified the standard-normal behavior directly and reproduce the pooled distribution below for the referee's reference. We did not add it as a supplementary figure, since the verification reduces to the three numbers now stated in Sect. 2.5, but we are happy to include it if the referee or editor would prefer.

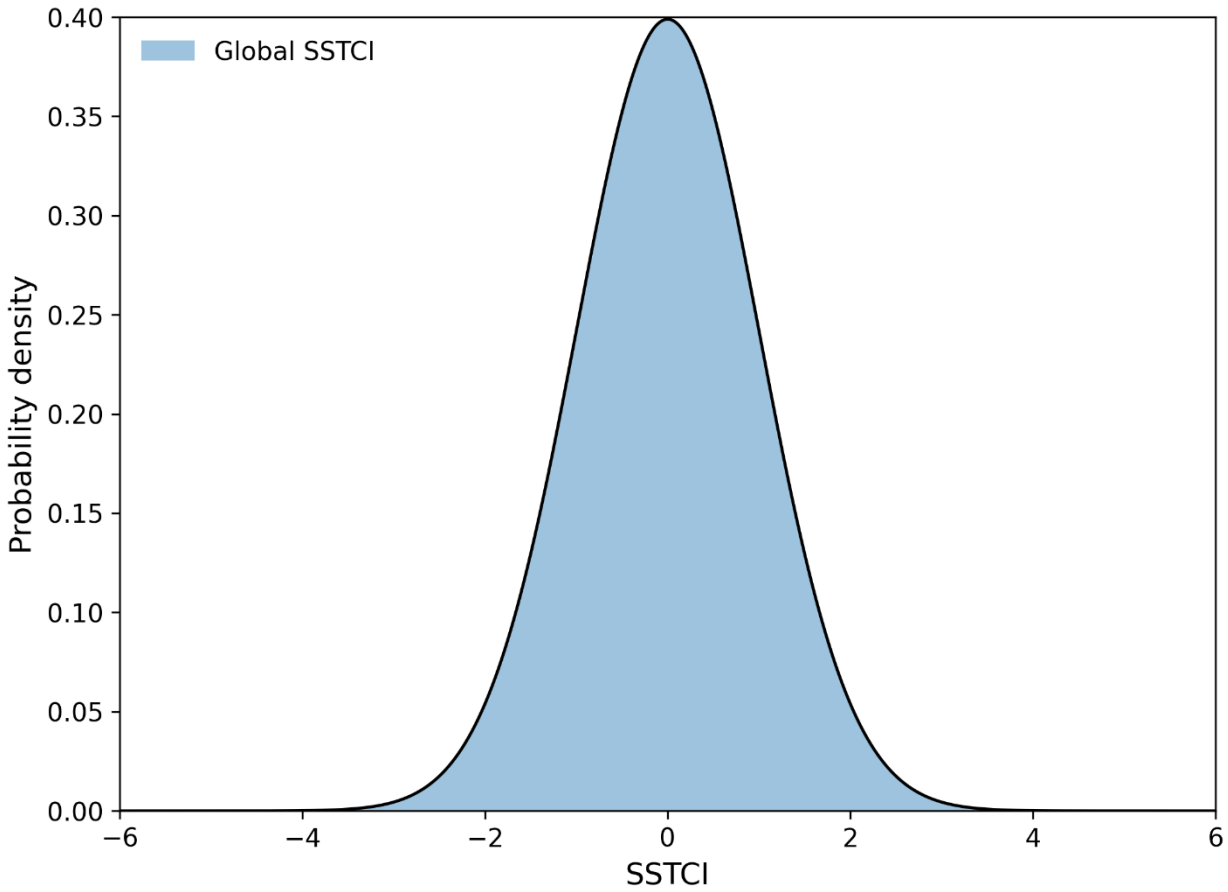


Figure: SSTCI Histogram

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

Author's response

We have moved the continuous wavelet analysis in full to the Supplement (Sect. S7, formerly Fig. 8), retaining a two-sentence summary in the main text. As explained under the general comment, we have retained the 2022 Yangtze River Basin reconstruction in the main text and set out our reasoning in a separate letter to the editor; we recognize this differs from what we indicated during open discussion and we defer to the editor on it.

Author's changes in manuscript

Wavelet text and Fig. 8 relocated to Supplement Sect. S7; summary retained at L589–593. Subsequent figures renumbered (former Fig. 9 → Fig. 8, former Fig. 10 → Fig. 9).

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

Author's response

It does not, and the original wording implied more than the model supports. We have softened the wording accordingly and added the caveat both where τ is interpreted and in the limitations section.

Author's changes in manuscript

Sect. 3.1, L279–282:

"ERA5-Land also does not include an explicit irrigation scheme, so anthropogenic water management is not represented in the soil moisture fields; τ therefore reflects natural land-surface and hydroclimatic controls as simulated by the model, and dry-down behavior in intensively irrigated regions should be interpreted with that in mind."

Repeated as a limitation in Sect. 3.9.

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

Author's response

There are two independent reasons, and we treat these regions as low-confidence.First, over hyper-arid deserts ERA5-Land soil moisture is only weakly constrained: assimilated observations are sparse and the land-surface model has little dynamic information, so much of the residual variability reflects model behaviour rather than a physically trustworthy signal. We do not think correlations built on that variability should be interpreted.Second, and independently of data quality, these regions are perennially moisture-limited: near-surface soil moisture sits at or near its residual floor year-round, with negligible dry–wet dynamic range. There is correspondingly little variability for any pair of drought indices to agree on, and the comparison is uninformative in

principle. This affects DEDI and SPEI equally and is not specific to SASMI, indeed the reference dataset papers themselves flag these regions as low-confidence.

We have stated both reasons in the revised text and noted that a targeted assessment of hyper-arid regions would be worthwhile as better-constrained soil moisture observations become available.

Author's changes in manuscript

Sect. 3.2.1, L317–320:

"Lower agreement over the arid and semi-arid belts follows from the differing accumulation windows of the three indices (a daily deficit for DEDI, 30 days for SPEI, ~14 days for SASMI), which offset the day each crosses −2, and from the low dry–wet dynamic range of these regions, which the reference-dataset papers themselves flag as low-confidence; DEDI and SPEI agree with each other at a comparable level over the same period."

Sect. 3.9, L675–677:

"hyper-arid regions have negligible dry–wet dynamic range and are correspondingly low-confidence and are masked in our evaluation."

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).

Author's response

We have adopted both suggestions, and they now form the two panels of the main text Fig. 2. We also went further than our initial reply indicated: rather than relegating the 2σ agreement map to the Supplement, we promote it to the main text, because it answers a different and stricter question than correlation and the two together give a more honest picture than either alone.

Rank correlation. Recomputed with Spearman's ρ at daily resolution over 2001–2021, SASMI correlates positively with both reference indices across essentially the entire land surface: median ρ = 0.28 against DEDI and 0.36 against SPEI, with positive values at roughly 97% and 99% of land pixels respectively. The spatial patterns match the Pearson maps, which are retained in the Supplement (Fig. S3) so that both are available. Correlation is dominated by ordinary days and establishes that SASMI tracks the same day-to-day drought variability as the established indices.

Extreme-day agreement. The 2σ metric asks whether two indices cross −2 on the identical calendar day ; a far stricter condition. Exact-day agreement is modest: median ≈ 0.015 for SASMI–DEDI and ≈ 0.004 for SASMI–SPEI. We think this is intrinsic to the comparison rather than a property of SASMI, for two reasons.

The three indices integrate over different windows, a daily deficit for DEDI, a 30-day accumulation for SPEI, a ~14-day antecedent memory for SASMI, so even when all three agree that drought is occurring, the day each crosses the threshold is offset. This mechanism makes a specific prediction that the data bear out: SASMI–SPEI has the *higher* rank correlation but the *lower* exact-day agreement, because SPEI's 30-day window is the most offset from SASMI's 14-day window.

More decisively, we computed the same metric between the two reference indices themselves. DEDI and SPEI agree with each other at a median of ≈ 0.019, essentially the same level at which SASMI agrees with either. Modest exact-day agreement is therefore a general property of comparing daily drought indices with different accumulation windows, not a deficiency of the new index.

The spatial pattern of agreement also tracks the documented reliability of the reference datasets rather than the quality of SASMI: agreement is lowest over hyper-arid regions and dense tropical forest, precisely where the reference-dataset papers flag arid PET artefacts, sub-canopy signal loss and freeze–thaw effects, and highest over temperate and humid regions where all indices are most reliable.

Author's changes in manuscript

Sect. 3.2.1 rewritten, L300-306. New Fig. 2 replaces the former two-panel Pearson figure. Method statement updated: SASMI–SPEI equivalents added as Fig. S2; Pearson maps as Fig. S3.

3.2.1 Agreement between SASMI and established drought metrics

"We evaluated the proposed Standardized Antecedent Soil Moisture Index (SASMI) against two established daily drought indices: The Daily Evapotranspiration Deficit Index (DEDI) (X. Zhang et al., 2023b) and the Standardized Precipitation–Evapotranspiration Index (SPEI) (Liu et al., 2024). All three indices are analyzed at daily resolution for 2001–2021 (based on data availability). Agreement is assessed in two ways: Spearman rank correlation, which tests whether the indices co-vary from day to day, and the fraction of agreement on extreme (2σ) days, defined as the number of days both indices fall below −2 divided by the number of days either does. Figure 2 shows both for SASMI and DEDI; the SPEI comparison and the Pearson maps are provided in the Supplement (Figs. S2–S3).

The SASMI–DEDI correlation map (Fig. 2a) exhibits widespread positive and significant correlations across large portions of the tropics, subtropics, and temperate regions, forming broad, spatially coherent belts. Weaker and locally negative values appear primarily over densely vegetated tropical rainforest, because DEDI does not remove evaporation by interception loss and therefore does not explicitly reflect evaporative stress at the soil surface, whereas SASMI is more directly related to SM stress during droughts. These regions are also persistently wet, so ordinary-day variability is small and the lower correlation reflects limited dynamic range rather than a failure to detect drought. The extreme-day agreement (Fig. 2b) supports this reading. The strongest agreement occurs over the Amazon and Congo basins and maritime Southeast Asia, precisely where the correlations are weakest, together with northern Eurasia and eastern Australia. The two metrics are therefore complementary: correlation is dominated by ordinary days, while extreme-day agreement isolates the rare occasions when both indices register a large departure. Lower agreement over the arid and semi-arid belts follows from the differing accumulation windows of the three indices (a daily deficit for DEDI, 30 days for SPEI, ~14 days for SASMI), which offset the day each crosses −2, and from the low dry–wet dynamic range of these regions, which the reference-dataset papers themselves flag as low-confidence; DEDI and SPEI agree with each other at a comparable level over the same period.

Taken together, Figure 2 points to the reliability of SASMI for drought detection across most global land areas. The consistency of these independent comparisons indicates that SASMI captures broadly similar drought variability to existing metrics, thereby supporting its use as a reliable drought indicator in our study”

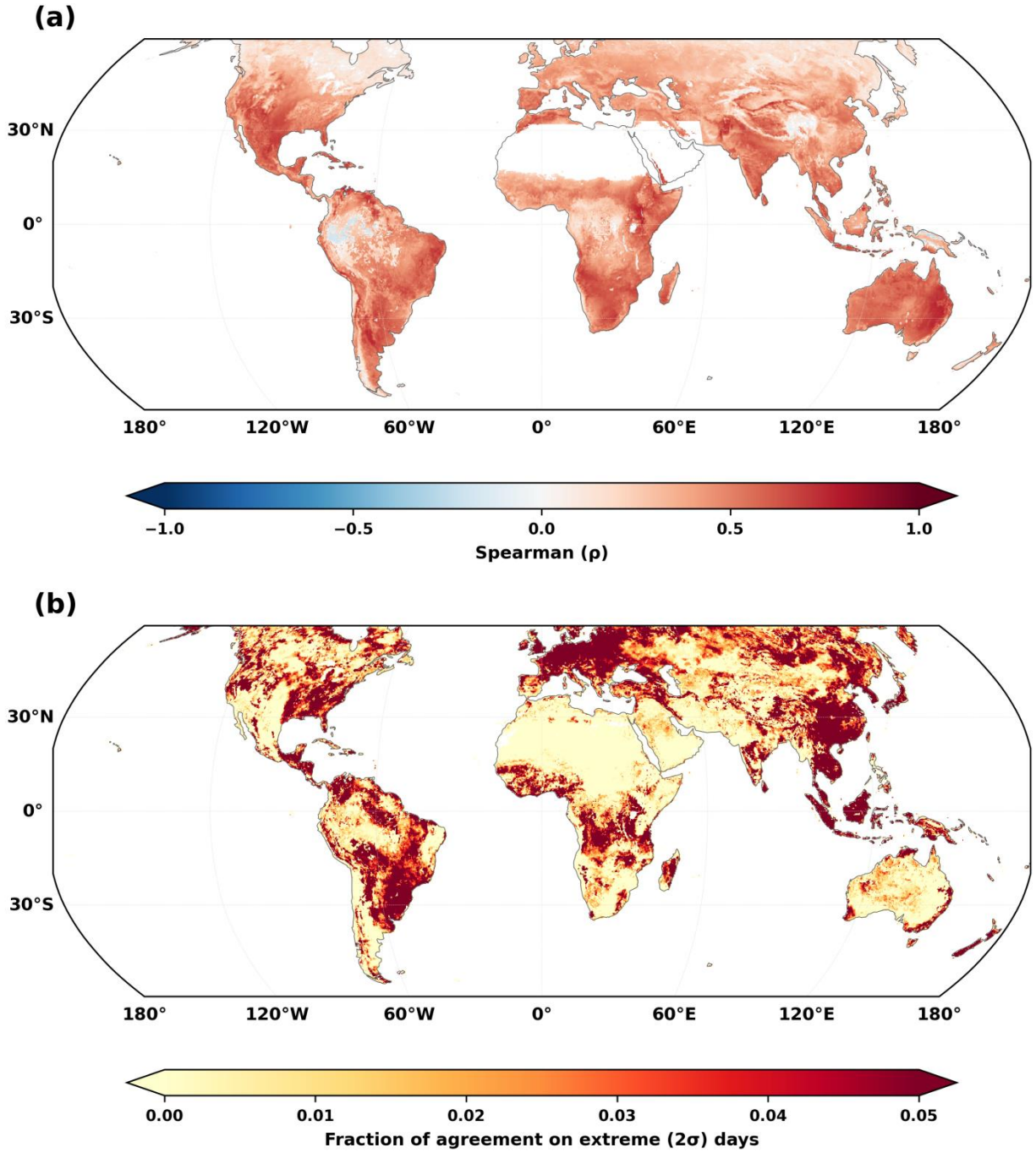


Figure 2. Agreement between the Standardized Antecedent Soil Moisture Index (SASMI) and the Daily Evapotranspiration Deficit Index (DEDI), computed at daily resolution over 2001–2021. (a) Spearman rank correlation; only statistically significant values ($\alpha = 0.05$) are shown, and non-significant pixels are masked. Warm colors indicate positive correlation and cool colors indicate negative correlation. (b) Per-pixel fraction of agreement on extreme (2σ) days, calculated as the number of days both indices fall below -2 divided by the number of days either index falls below -2.

SUPPLEMTARY: We have added the below figure in supplementary material

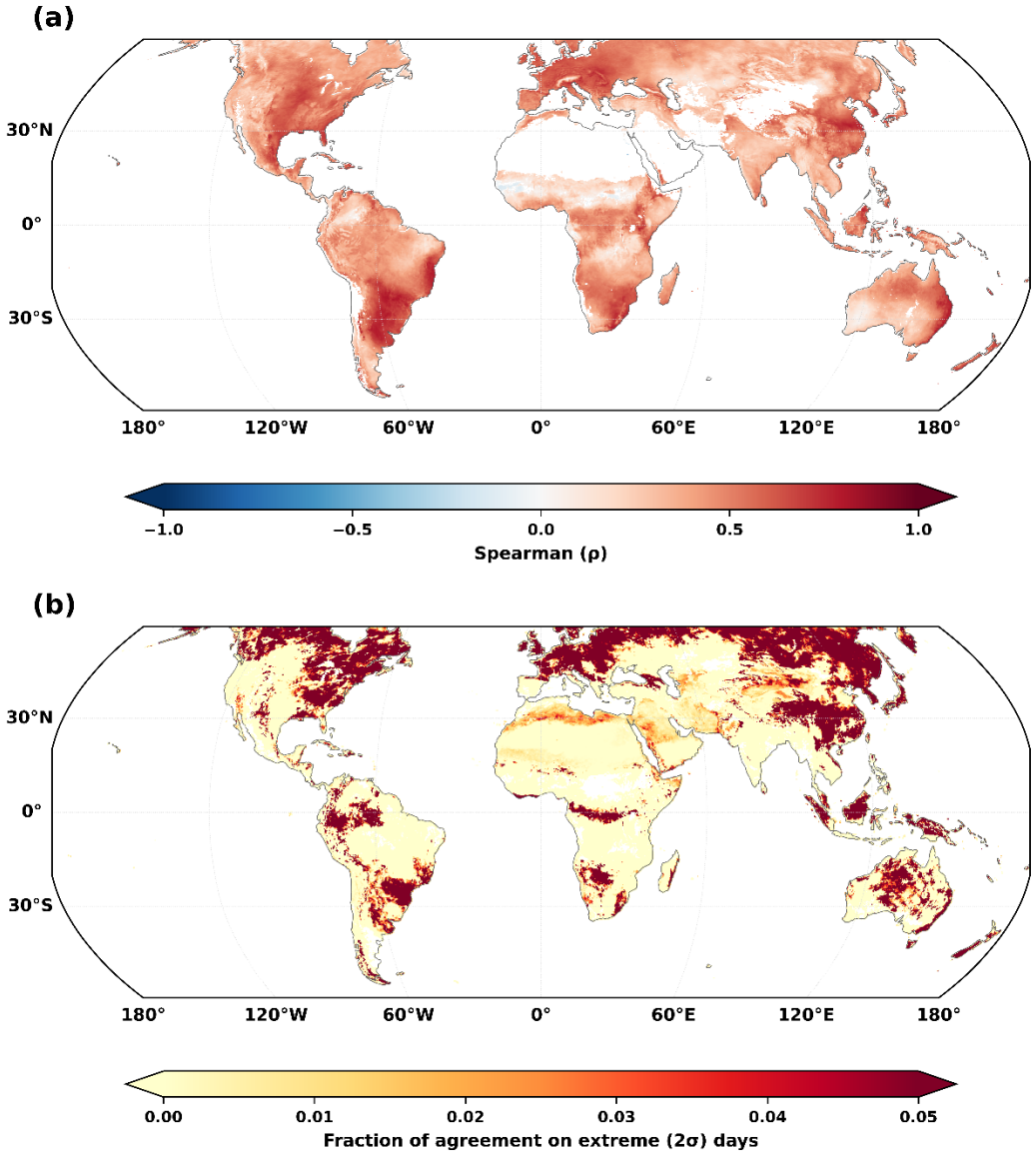


Figure S2. Agreement between the Standardized Antecedent Soil Moisture Index (SASMI) and the Standardized Precipitation–Evapotranspiration Index (SPEI), computed at daily resolution over 2001–2020. (a) Spearman rank correlation; only statistically significant values ($\alpha = 0.05$) are shown, and non-significant pixels are masked. Warm colours indicate positive correlation and cool colours indicate negative correlation. (b) Per-pixel fraction of agreement on extreme (2σ) days, calculated as the number of days both indices fall below -2 divided by the number of days either index falls below -2. Warmer colours indicate higher agreement on the timing of extreme dry days. Barren and low-variability regions masked in the SPEI product appear as no data in both panels.

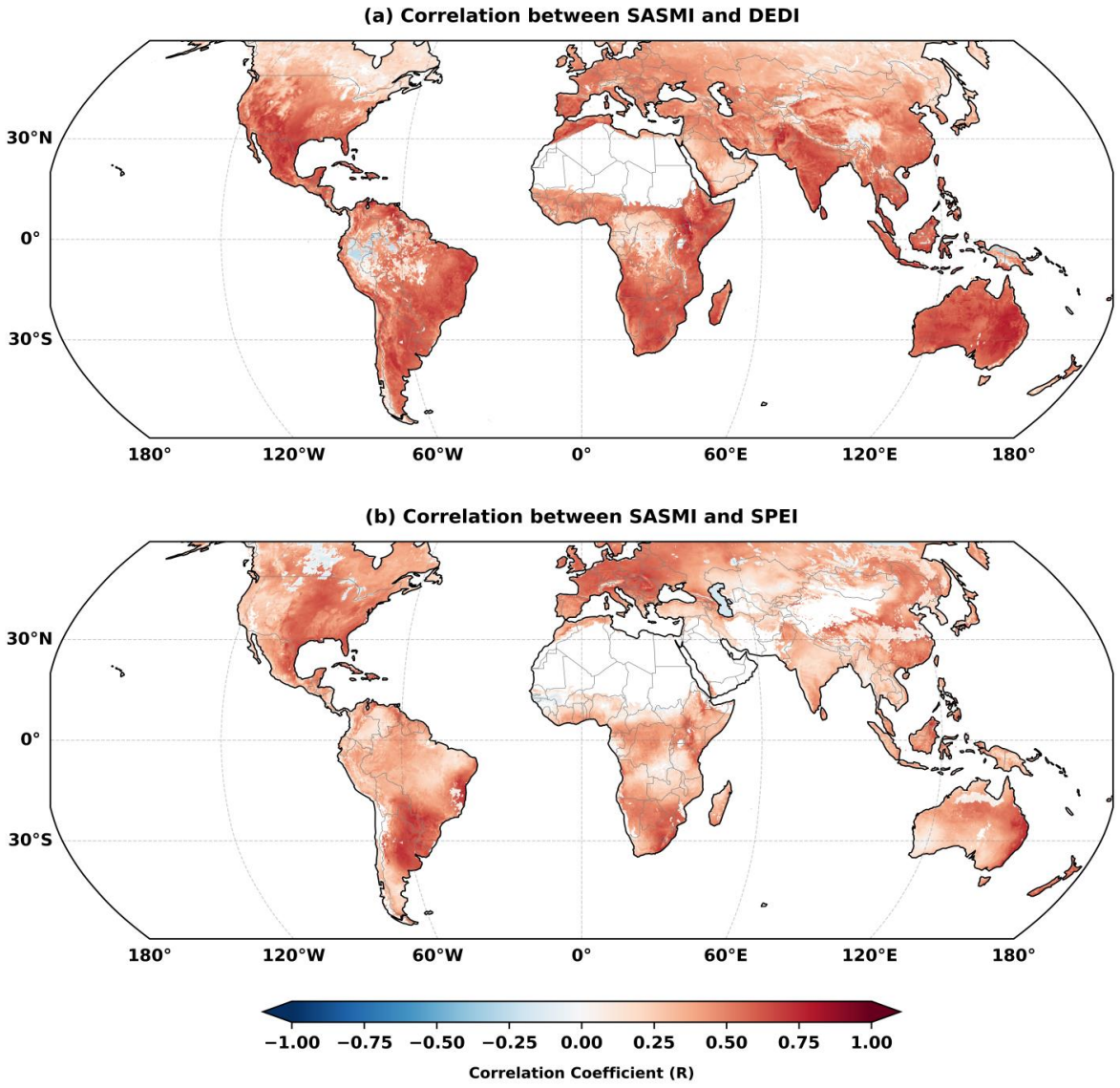


Figure S3. Spatial Pearson correlation between the Standardized Antecedent Soil Moisture Index (SASMI) and two established daily drought indicators: (a) the Daily Evapotranspiration Deficit Index (DEDI) and (b) the Standardized Precipitation–Evapotranspiration Index (SPEI). Correlations are calculated at daily resolution over the common analysis period, and only statistically significant correlations ($\alpha = 0.05$) are shown; non-significant pixels are masked. Warm colors indicate positive correlation and cool colors indicate negative correlation between the indices.

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.

Author's response

The characterisation of SSTCI's novelty needed the correction the referee implies, and we have made it. Daily resolution alone is not the contribution; the contribution is a *global*, daily, soil-moisture based compound index. SCDHI is indeed daily, but as published it is regional (China only) and its drought component is precipitation based.

We have added the requested specifications. SCDHI (Zhou et al., 2025) combines the precipitation-based SAPEI with STI through a copula and was developed over China using NASA NEX-GDDP-CMIP6 downscaled meteorological data. Because no global compound dry–hot index existed to compare against, we reproduced the SCDHI methodology at global scale following Zhou et al. (2025). We state plainly that this global SCDHI is an analysis product generated for the present evaluation rather than a pre-existing public dataset, and that the comparison period is 1981–2013.

Author's changes in manuscript

Sect. 2.1, L151–157:

"To evaluate SSTCI against an established daily compound metric, we used the Standardized Compound Dry–Hot Index (SCDHI; Zhou et al., 2025), which combines the precipitation-based SAPEI with STI through a copula. As published, SCDHI was developed regionally over China using NASA NEX-GDDP-CMIP6 downscaled meteorological data. Since no global daily compound dry–hot index was available for comparison, we reproduced the SCDHI methodology at the global scale following Zhou et al. (2025); this global SCDHI is an analysis product generated for the present evaluation rather than a pre-existing public dataset, and the period 1981–2013 is used for the consistency assessment with SSTCI."

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

Author's response

It is the spatial mean over the Sichuan–Chongqing domain, now specified in the caption. Note that the figure has been expanded to five panels in the revision, so this is now panel (e) and the maps it averages are panels (c) and (d).

Author's changes in manuscript

Fig. 3 caption, L374:

"(e) Daily time series of SSTCI and NDVI averaged over the Sichuan–Chongqing domain shown in panels (c) and (d)."

"an event historically distinguished by its unprecedented duration and intensity": this would of course make it one of the easiest events to test the metric for. Of course, it's fine to showcase with a very obvious event, but I miss insights on how the metric behaves with more intermediate (less "extreme") events. I also miss any evaluation in terms of false positives: are there any extreme events detected by the method, which were not so extreme in reality? I understand this is not easy to do without any reference, but I make a suggestion in the comment below.

***Sect 3.2.3**: Nothing on potential 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. Would be more robust than a limited number of case studies in my opinion.*

Author's response

We address these related points together.

On further examination we would like to refine the argument offered in our initial reply during open discussion. We suggested there that a confusion matrix of CDHEs against SASMI would add

little, on the grounds that an extreme SSTCI already implies an extreme SASMI. That framing does not hold once the copula construction is examined closely, and the referee's question prompted us to look at it directly. A copula does not operate as a logical AND of marginal thresholds: SSTCI is the standardized transform of the joint exceedance probability p, which can reach its extreme tail through combinations where neither driver individually crosses ±2 but the pair is jointly improbable under the locally fitted dependence structure. The two sets are therefore not nested, and a matrix against SASMI would test a relationship the index does not impose rather than isolating false positives. What guards against single-driver flags is the dependence structure itself: an extreme SSTCI requires both drivers to be jointly elevated, so a day on which only one is anomalous receives a moderate value visible across all six panels of Fig. 4, and now stated explicitly in the manuscript

We agree with the referee that the 2006 Sichuan–Chongqing event is an easy test, and the evaluation no longer rests on case studies alone. For an external check, the appropriate comparison is against an independent compound index rather than against SSTCI's own components. We therefore extended the SCDHI comparison to extreme (2σ) days across 1,339,001 land pixels over 1981–2013 (Fig. 3b, with the exact-day version in Fig. S4). Agreement exceeds the chance expectation given the two base rates by a factor of 6.9, with 99.3% of pixels above chance, so days on which SSTCI registers a compound extreme are substantially more likely than chance to be days an independently constructed compound index does so too. The full analysis is given in our response to the Figure 4 comment below. This bounds false positives only insofar as SCDHI is itself reliable, but in the absence of a global observational reference for compound events it is the strongest external check available.

On false positives more broadly: since STI is a direct temperature standardization, the drought component is where a spurious compound flag could originate. The revised Fig. 2 and Figs. S2–S3 establish that SASMI tracks two established daily drought indices and there extreme agreement across essentially the whole land surface, which is the relevant check on that half of the index.

On intermediate events: SSTCI is continuous over the full severity range, and the −2 threshold defines only the tail retained in the catalogue. Moderate compound anomalies remain fully represented in the daily field, which is distributed as part of the dataset rather than being discarded

at the thresholding stage. The index's behavior across the full range of SASMI–STI combinations is documented in the companion methods paper (Aftab et al., 2026a, their Fig. 3, for the Poyang Lake Basin)

Author's changes in manuscript

Sect. 3.2.3, L423-428

"The event signatures in Fig. 4 also illustrate a general property of the index. Because SSTCI is derived from the joint probability of low soil moisture and high temperature rather than from either component alone, extreme values require the two drivers to be jointly elevated. A day on which only one driver is anomalous is assigned a moderate value rather than an extreme one, as can be seen in all six cases, so compound extremes are not registered where a single factor is elevated in isolation; the conditional behavior of the index across the full range of SASMI–STI combinations is documented in Aftab et al. (2026a) for the Poyang Lake Basin."

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.

Author's response

Fig. 4 shows SSTCI alongside its own components, so we first considered quantifying where it diverges from SASMI and STI. That would not be an informative test: SSTCI is constructed from the two through the copula, so any divergence restates the transformation rather than testing it. The informative comparison is against an external compound index, and that is what we have added.

We agree, and we have added a systematic disagreement analysis rather than only citing individual cases.

Systematic. We computed per-pixel extreme-day agreement between SSTCI and SCDHI over 1981–2013 across 1,339,001 valid land pixels. On exact days the critical success index is 0.077. Allowing a ±5-day timing tolerance raises it to 0.177, a 2.3-fold recovery, which tells us that a substantial share of the apparent disagreement is timing offset between two differently constructed indices rather than disagreement about whether a compound event occurred. Set against a chance expectation of 0.0115 given the two base rates (0.021 for SSTCI, 0.026 for SCDHI), exact-day agreement is 6.9× chance; 99.3% of pixels exceed chance and 70.2% exceed five times chance. Cohen's κ is 0.124. The recovery under tolerance is largest in the tropics (2.84×) and smaller in the northern and southern extratropics (2.22× and 2.30×), consistent with timing offsets being widest where the drought components diverge most. We also report that the exact-day figure should be read against the ceiling imposed by the differing base rates, which is 0.789 rather than 1.Rank and linear correlation between the two indices are close (Spearman $\rho = 0.324$, Pearson $r = 0.334$), indicating a near-linear relationship, so the choice of correlation measure is not driving the result here.

Individual events. Cases where SSTCI captures conditions SCDHI underestimates do exist. In the companion methods paper (Aftab et al., 2026a, their Fig. 5, Poyang Lake Basin) SSTCI identifies the 2013 event as an extreme CDHE that SCDHI does not classify as extreme, and for the 2019 event SCDHI registers only a short episode where SSTCI resolves the full multi-week compound event.

The main text now presents both the ±5-day agreement map (Fig. 3b) and the exact-day map (Fig. S4), so that the disagreement is shown rather than described.

Author's changes in manuscript

Sect. 3.2.2, L331–374: (new addition to the section is highlighted for your convenience below (L338–345))

New Fig. 3b; new Fig. S4.

3.3.2 Consistency of SSTCI with SCDHI and impact-based evidence from vegetation response

Figure 3 provides complementary evidence for the robustness of SSTCI by demonstrating both global consistency with an established Standardized Compound Dry–Hot Index (SCDHI) (Zhou et al., 2025) and physically meaningful linkages to impact relevant land-surface stress. At the global scale, temporal correlations between SSTCI and SCDHI are predominantly positive and statistically significant across most land areas ($p < 0.05$; Fig. 3a), indicating coherent temporal co-variability between the two indices. Regions of high agreement are spatially widespread, whereas localized areas of weaker correspondence highlight geographic heterogeneity in index behavior, plausibly reflecting regional differences in hydroclimatic regimes and moisture–temperature coupling strength. We additionally compare the two indices on extreme (2σ) days, applying the same agreement metric used for the drought components in Sect. 3.2.1 (Fig. 3b). Because SSTCI and SCDHI are built on different drought components and accumulate windows, the day on which each crosses -2 during the same episode need not coincide. Agreement therefore improves substantially when a ±5-day tolerance is allowed, indicating that much of the exact day mismatch reflects timing offsets between the two formulations rather than disagreement over whether a compound event occurred; the exact day comparison is provided in the Supplement (Fig. S4). Agreement is strongest across the extratropic and weakest over hyper-arid and densely vegetated tropical regions, consistent with the pattern found in Sect. 3.2.1. These global results are reinforced by supplementary diagnostics showing broadly comparable seasonal spatial footprints and consistent event-phase evolution during the 2010 western Russia CDHE (Figs. S5–S6), supporting the ability of SSTCI to reproduce the large-scale and event-scale spatiotemporal characteristics.

To validate the SSTCI's capacity for CDHEs, we examined the catastrophic 2006 heatwave-drought in the Sichuan–Chongqing region; an event historically distinguished by its unprecedented duration and intensity (Zhang et al., 2008). The region experienced continuous extreme temperatures from mid-June to late August 2006, representing the most severe and longest heatwave in the historical record, coincident with an exceptional (≈100-year) drought. The event caused widespread societal and ecological impacts, including severe drinking-water shortages, substantial cropland losses, and numerous forest fires, prompting the highest level of local aridity warning (Wu et al., 2020). The event scale evaluation confirms that SSTCI anomalies are physically consistent with ground conditions and impact relevant. Specifically, the index effectively captures the event's temporal evolution, tracking a clear phase progression from mixed

pre-event conditions to a coherent negative phase during initiation, which subsequently intensifies and expands spatially at the event's peak (Fig. 3c). This evolution is verified by independent vegetation dynamics, where we observe that vegetation stress coincides with the SSTCI trajectory, broadening and deepening from onset to peak intensity (Fig. 3d). Notably, the regions exhibiting the most severe SSTCI values are broadly collocated with the sharpest NDVI reductions. The daily time series provides further verification (Fig. 3e), revealing that sustained negative SSTCI values, during the initiation and peak phases, drive a progressive decline in vegetation health, with the NDVI minimum lagging slightly to coincide with, or immediately follow, the peak DH stress.

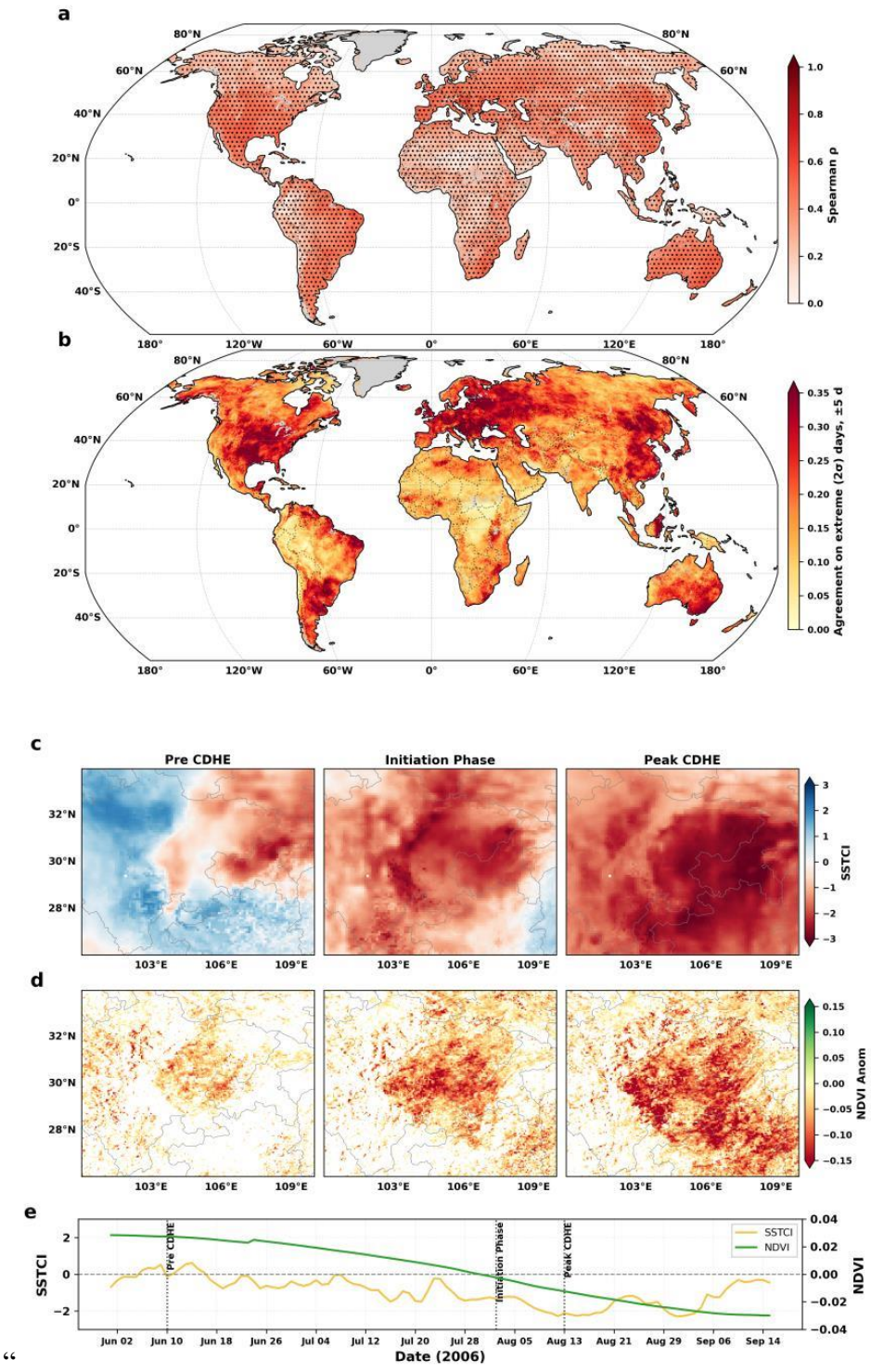


Figure 3. Evaluation of SSTCI through comparison with an existing compound dry–hot index and an independent vegetation response during the 2006 Sichuan–Chongqing extreme event. (a) Global correlation between SSTCI and SCDHI over 1981–2013, showing their temporal co-

variability across land areas; stippling denotes statistically significant correlations at the 5% level. (b) Per-pixel fraction of agreement between SSTCI and SCDHI on extreme (2σ) days over the same period, allowing a ±5-day tolerance, calculated as the number of days both indices fall below −2 divided by the number of days either index does. (c) Spatial phase maps of SSTCI during the 2006 Sichuan–Chongqing CDHE, illustrating the progression from pre-event conditions through initiation to peak. (d) Corresponding maps of negative NDVI anomalies for the same phases, showing the spatial development of vegetation stress. (e) Daily time series of SSTCI and NDVI averaged over the Sichuan–Chongqing domain shown in panels (c)and (d).

Supplementary

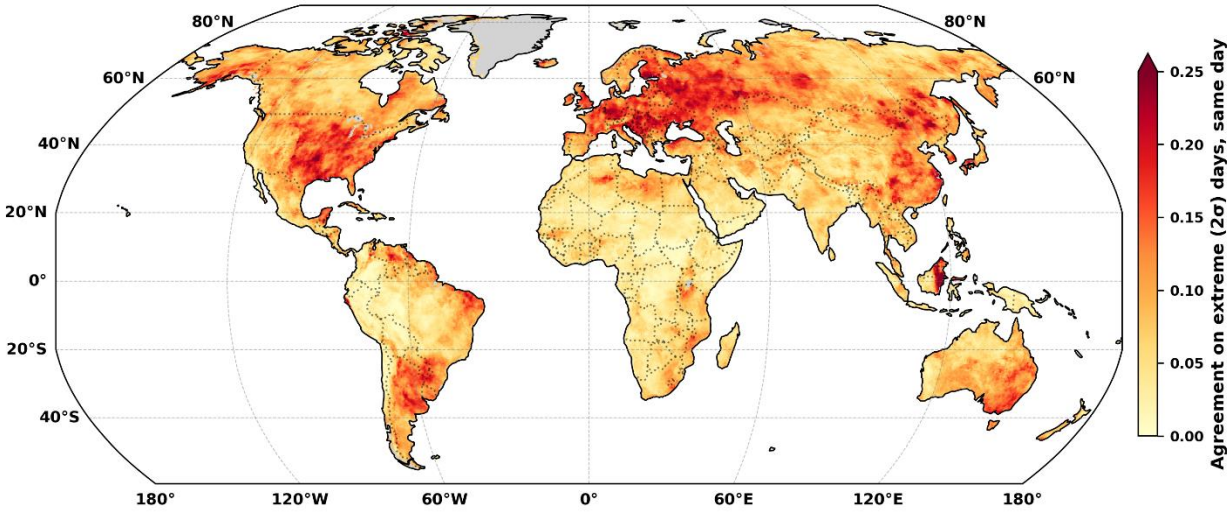


Figure S4. Per-pixel fraction of agreement between SSTCI and the Standardized Compound Dry–Hot Index (SCDHI) on extreme (2σ) days over 1981–2013, requiring exact same-day co-occurrence. Agreement is calculated as the number of days both indices fall below −2 divided by the number of days either index does. The corresponding map allowing a ±5-day tolerance is shown in Fig. 3b.

The second paragraph made me wonder what exactly the benefit of is 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?

Author's response

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.

Author's changes in manuscript

Sect. 3.3, L459–464:

"Here, the compound index is designed to capture this synergy. Dry and hot extremes do not merely co-occur, but amplify each other, and the resulting impacts on ecosystems and societies are disproportionately severe relative to either stressor alone (Zscheischler and Seneviratne, 2017; Feng et al., 2019). Because SSTCI is derived from the joint probability of soil moisture and temperature, it reaches its most extreme values specifically where dryness and heat coincide more

strongly than their independent behaviour would imply, and therefore reflects the reinforcement between the two drivers rather than their simple superposition."

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?

Author's response

They are classified as two separate events, and the distinction is physically meaningful. The first, shorter excursion corresponds to an early compound spell in which a brief warm anomaly acted on soil that was drying but not yet fully depleted, producing a compound signal of limited depth and duration. The second is the main event: by late July the soil moisture deficit had deepened, so the heat anomaly met a surface with little remaining capacity for evaporative cooling and the land–atmosphere feedback became self-sustaining. Because the two spells are separated by a sufficiently pronounced recovery interval, the compound deficiency accumulated between them exceeds the merging threshold and the removal–merging procedure retains them as two independent events. We have made this explicit in the text.

Author's changes in manuscript

Sect. 3.3, L405–456:

"The record contains two dips below −2. The first and shorter one occurs earlier, when a brief warm spell acted on soil that was drying but not yet fully depleted, producing a compound signal of limited depth and duration. The second is the main event: by late July the soil moisture deficit had deepened, so the heat anomaly encountered a surface with little remaining capacity for evaporative cooling and the land–atmosphere feedback became self-sustaining. Because the two spells are separated by a sufficiently pronounced recovery interval, the removal–merging procedure (Sect. 2.6) retains them as two independent events rather than merging them."

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

Author's response

The procedure is not specific to SSTCI. The removal–merging optimization of Shan et al. (2024) operates on any daily standardized index series and could be applied to any daily index. What is new here is its application to a compound index to build an event catalogue. We note that generating such catalogues globally at 0.1° is computationally demanding, the threshold search and two goodness-of-fit tests are evaluated at several million grid cells over a 63-year daily record, which is one reason we provide the SSTCI catalogue as a ready-to-use product rather than expecting users to regenerate it.

Author's changes in manuscript

Sect. 2.6, L233-L236.

“To our knowledge this procedure has not previously been applied to a compound dry–hot index. Shan et al. (2024) developed it for separately identified drought and heatwave series, whereas here it operates on the joint compound signal, so that the independence and extreme-value tests are applied to the compound quantity itself rather than to its marginals”

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?

Author's response

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.

Author's changes in manuscript

Sect. 3.9, L702–703:

"Likewise, since SSTCI requires only near-surface soil moisture and temperature, computing it from forecast fields is a possible route to early warning."

So, the moisture/temperature indices in themselves do not show elevated probability of an extreme event post-2000?

Author's response

Let’s 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 behavior: 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, the relevant references were missing from this passage and have been added.

Author's changes in manuscript

Sect. 3.5, L510–516:

"The event-identification procedure is applied only to the SSTCI series, so the reported increase is a property of catalogued compound events rather than of separately catalogued drought or heat events. That the individual drivers have themselves intensified in recent decades is well established (IPCC, 2021); the compound frequency, however, is governed not only by the marginals but by the strengthening dependence between them, so that joint occurrence rises faster than the independent combination of the two would imply (Zscheischler and Seneviratne, 2017), a dependence that SSTCI captures through the copula. Anthropogenically forced increases in compound dry–hot frequency has been documented directly at global and continental scales (Zhang et al., 2022)."

Figure 6

In order to verify whether the whole SSTCI distribution has shifted to more negative values post-2000 [...] or whether it has increased in variance (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?

Author's response

We appreciate the diagnostic logic, and two points 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 is not the physical counterpart of a compound dry–hot extreme. Wet–cool conditions are governed by a different set of land–atmosphere processes, and a symmetric reading of the index would not carry the same physical meaning at both ends. Such conditions are of genuine interest, but they fall outside the scope of a dry–hot dataset and we intend to treat them in dedicated follow-up work.

We have not added an analysis to the manuscript in response to this comment and note it as a response only item.

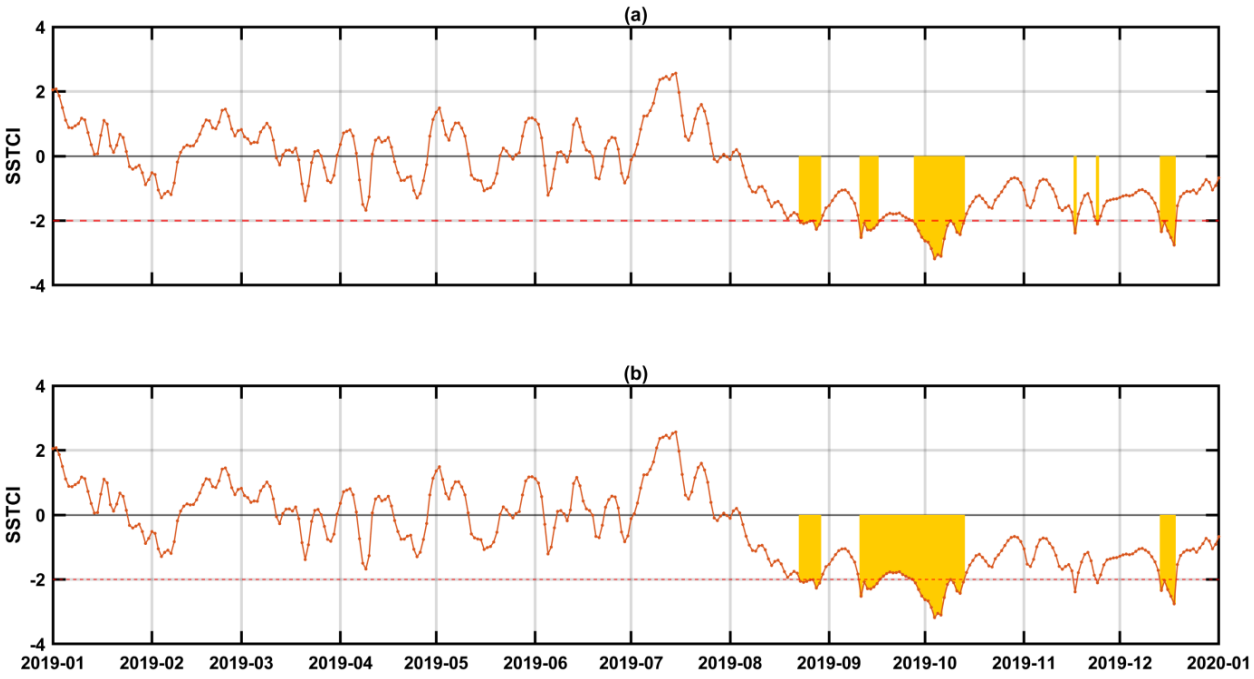


Fig. 4. Comparison of events before (a) and after (b) applying the removal and merging method.

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?*

Author's response

The description was unclear and has been rewritten. The mapped quantity is a per pixel temporal trend in CDHE duration: at each grid cell the duration of every catalogued event is regressed against the calendar year in which it occurred. Because events are discrete and irregularly spaced, with many years containing none, 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 positive values indicating that events later in the period tend to be longer.

On the box-plot suggestion: durations are not well suited to the two-period format used in Fig. 6. That format works for counts because counts take a single value per pixel and vary strongly from pixel to pixel, giving a wide and informative spread. Event durations 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, and the 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 what the slope map conveys and a two-period box plot cannot. We have therefore retained the trend map with the clarified description and an expanded caption.

Author's changes in manuscript

Sect. 3.6, L571–574:

"Figure 7 maps per-pixel temporal trends in CDHE duration for 1961–1999 and 2000–2023. At each grid cell, the duration of every catalogued event is regressed against the calendar year in which it occurred; because events are discrete and irregularly spaced, with many years containing none, the regression is fitted over the actual event years from the first to the last event at that cell."

Fig. 7 caption, L584–586:

"Per-pixel trends in CDHE duration for (a) 1961–1999 and (b) 2000–2023. At each grid cell, event duration is regressed against event year; positive slopes indicate events becoming longer. Only pixels with at least five events and significant slopes ($p < 0.05$) are shown."

Technical corrections

L43: define CDHE at first use. L91–92: section numbering. L558: figure reference.

Author's response

All three corrected.

Author's changes in manuscript

- CDHE expanded at first use in the Introduction, L44.
- Section numbering corrected at L110–114: limitations and outlook is Sect. 3.9, conclusions Sect. 4, data availability Sect. 5.
- Figure cross-reference corrected following renumbering; former Fig. 9 is now Fig. 8 (L625) and former Fig. 10 is now Fig. 9 (L660).