

Responses to reviewers' comments for EGUSPHERE-2026-1199

Global vegetation responses to wet and dry soil moisture extremes

RC1:

This study presents a global-scale analysis of vegetation greenness responses to both dry and wet soil moisture extremes using long-term NDVI data and a machine-learning-based attribution framework. The manuscript is well structured, and the separation of drivers into pre-extreme conditions, extreme characteristics, and environmental background is conceptually clear and useful. The results confirming consistent vegetation browning during dry extremes and more heterogeneous responses during wet extremes are interesting and pretty important. I include my comments below, which I hope help the authors to strengthen the paper.

Reply:

The authors thank the reviewer for their positive summary and constructive comments that have helped improve the manuscript. We have detailed replies to each comment below, marked in blue.

1. The Introduction provides a useful overview of previous studies but does not clearly articulate the specific research gap that this study aims to address.

Reply:

We thank the reviewer for this helpful comment. While previous studies have mainly focused on vegetation responses to drought or have examined wet and dry extremes separately, often using different datasets or regional-scale approaches, this study aims to provide a consistent large-scale comparison of the predictors associated with vegetation responses to wet and dry soil moisture extremes. We therefore revised the Introduction to more clearly articulate the research gap on lines 49-50 and lines 60-63. We also explicitly discuss the challenge of filling this gap on lines 64-66. We further clarify how our study addresses this challenge in the following paragraph (lines 66-74), where we describe our global, systematic assessment of vegetation greenness responses to both wet and dry soil moisture extremes using a consistent attribution framework designed to identify relevant predictors while minimizing collinearity among them.

Lines 49-50:

"While the overall relevance of dry and wet extremes for vegetation function is well established, important knowledge gaps remain regarding the spatial variability of their impacts and the mechanisms driving this variability."

Lines 60-63:

"However, it remains unclear to what extent environmental background conditions, characteristics of the extremes (e.g., severity and timing), and pre-extreme vegetation

conditions jointly shape vegetation responses across regions and biomes, particularly for wet extremes.”

Lines 64-66:

“In addition, because previous studies have often differed in vegetation and hydrological datasets, extreme definitions, and spatial or temporal scales, directly comparing the dominant predictors of vegetation responses to wet and dry extremes at the global scale remains challenging.”

Lines 66-74:

“Leveraging long-term Earth observation datasets, we provide here a global, systematic assessment of vegetation greenness responses to both wet and dry soil moisture extremes. We identify the key factors shaping vegetation greenness anomalies across regions for both types of extremes. The identification of influential factors (referred to as predictors in the following) in this study is challenged by inconsistencies among observational datasets and collinearity among considered predictors (Jiang et al., 2024). In order to account for this, we develop an effective machine learning-based attribution approach that ensures to isolate relevant predictors with the most explanatory power while minimizing collinearity between considered predictors. This way, our approach enables a clearer separation of the roles of environmental background conditions, extreme characteristics, and pre-extreme vegetation conditions.”

2. (1) The study defines wet and dry extremes based on the 5th and 95th percentiles of soil moisture, both overall and seasonally. While this ensures statistical consistency, it does not necessarily correspond to ecologically meaningful thresholds. The identified “extremes” may reflect statistical anomalies rather than true ecological stress events.
- (2) In addition, the manuscript employs two types of thresholds, but their respective roles are not clearly reflected or discussed in the Results.

Reply:

(1) We agree with the reviewer that percentile-based thresholds define statistical rather than explicitly ecological extremes. However, the 5th and 95th percentiles of soil moisture are designed to capture rare departures from the local hydroclimate background, which are often associated with conditions that are relevant for vegetation stress or enhanced water availability. While these thresholds do not guarantee ecological impacts across all regions, percentile-based thresholds are widely used to define locally rare soil-moisture conditions and provide a consistent, spatially comparable definition of hydrological extremes (Li et al., 2023).

In addition, in the attribution analyses in Figures 3-5, we only consider dry and wet soil moisture extremes during which NDVI is lower than usual. While this does not equal ecological stress, it contributes to focusing our analysis more strongly on negative vegetation responses. At the same time, we acknowledge that it is difficult to assess actual

stress at the spatio-temporal scales and extent of our analysis with the corresponding available data products. We clarified this distinction in the revised manuscript.

Lines 411-414:

“This study identifies wet and dry soil-moisture extremes relative to the local climatology and overall records at each grid cell. This percentile-based framework is well-suited for global analysis, but it does not necessarily capture absolute hydrological thresholds of vegetation’s response to extremes such as root-zone waterlogging or imply that all detected events correspond to ecologically stressful conditions (Smith, 2011).”

And we also added a future direction to better imply vegetation functioning’s response to hydrological extremes in lines 426-427:

“Future work could also benefit from emerging indicators of vegetation functioning, such as solar-induced chlorophyll fluorescence (SIF), once sufficiently long observational records become available.”

(2) The two thresholds were designed to ensure that detected hydrological extremes are both rare within the long-term soil moisture distribution and anomalous relative to the respective seasonal background conditions. We agree that their complementary roles were not sufficiently reflected in the original manuscript. We therefore clarified these roles in the Methods (lines 90-93) and added text in the Results (lines 186-188).

The overall and seasonal thresholds serve complementary purposes. The overall threshold ensures that detected events represent sufficiently dry or wet conditions within the full local soil moisture distribution, whereas the seasonal threshold ensures that these events are also anomalous relative to the expected conditions of the respective time of year. Using either threshold alone could therefore be misleading: the seasonal threshold alone may identify seasonally unusual but not hydrologically extreme conditions, while the overall threshold alone may mainly capture the normal seasonal cycle. Together, the two thresholds allow a more consistent comparison of vegetation responses across regions and seasons.

Lines 90-93:

“The combined use of overall and seasonal thresholds ensures that detected events are both rare in the context of the long-term soil moisture distribution and anomalous relative to the typical seasonal conditions. This avoids identifying regularly occurring seasonal dry or wet periods as hydrological extremes.”

Lines 186-188:

“As a result of the combined threshold approach, the detected extremes represent soil moisture that is rare within the long-term record and anomalous for the respective time of year, facilitating consistent comparisons across regions and seasons.”

Reference:

Li, X., Piao, S., Huntingford, C., Peñuelas, J., Yang, H., Xu, H., ... & Mahecha, M. D. (2023). Global variations in critical drought thresholds that impact vegetation. *National Science Review*, 10(5), nwad049.

3. The Random Forest–based attribution framework identifies important predictors but does not establish causal relationships between drivers and vegetation responses. Currently, the manuscript often interprets variable importance as indicating mechanistic control. Could you please justify it?

Reply:

We thank the reviewer for this important comment. We agree that the Random Forest framework identifies statistical associations and predictor importance rather than revealing causal relationships. Our intention in the attribution analysis is not to claim causal proof, but to identify dominant factors associated with vegetation responses and to interpret these patterns with established ecological understanding.

To avoid overstating causality directly from random forest results, we revised the manuscript to use more cautious wording throughout (e.g., replacing “controls” or “drivers” with “influential predictors”). We additionally clarified in the Method (lines 152-153) that the attribution analysis provides insights into potential mechanisms, but the mechanistic interpretation relies on ecological knowledge and should not be interpreted as direct causal inference.

Lines 152-153:

“The RF framework identifies predictors that are strongly associated with vegetation responses, while mechanistic interpretations are based on existing ecological understanding rather than direct causal inference.”

4. What about the Random Forest model performance?

Reply:

The average cross-validation (out-of-bag, OOB) scores of the selected random forest models across all regions ranged from 0.22 to 0.62 (SD = 0.11), while performance varied across IPCC regions and between wet and dry extremes. The details are shown in the supplementary (Figure S5 and S6). Considering the relevance of this information, we added the range and standard deviation of out-of-bag scores for all IPCC regions in the main text and described how this influences the reliability of results (lines 277-289).

Lines 277-289:

“The robustness and interpretability of these regional attribution patterns depend on the explanatory performance of the RF models and the degree of concavity among the selected predictors (Figure S5). The average OOB scores of the selected RF models across all regions ranged from 0.22 to 0.62 (SD = 0.11), while the average concavity values ranged from 0.23 to 0.66 (SD = 0.11). RF model performance varied across IPCC regions

and between wet and dry extremes. Lower OOB scores and higher concavity values were generally found in regions with relatively fewer retained events, such as Eastern Central Asia (ECA) for wet extremes and Russian-Arctic (RAR) for dry extremes (Figure S4). Lower explanatory power was also observed in regions where vegetation responses to hydrological extremes were comparatively weak or spatially heterogeneous, such as Western Central Africa (WAF) for wet extremes and Central Africa (CAF) for dry extremes (Figure 2). In contrast, regions with stronger and more spatially coherent NDVI anomalies during extremes generally showed higher OOB scores and lower concavity among predictors. Overall, OOB scores were slightly lower for wet extremes than for dry extremes, consistent with the more heterogeneous vegetation responses observed during wet extremes. Nevertheless, the RF models generally retained moderate-to-high explanatory power and low-to-moderate concavity across most regions, indicating that the identified attribution patterns are robust and that the selected predictors capture meaningful variability in vegetation responses to hydrological extremes.”

5. Converting 16-day NDVI to daily values may introduce artifacts. Please clarify the potential impacts.

Reply:

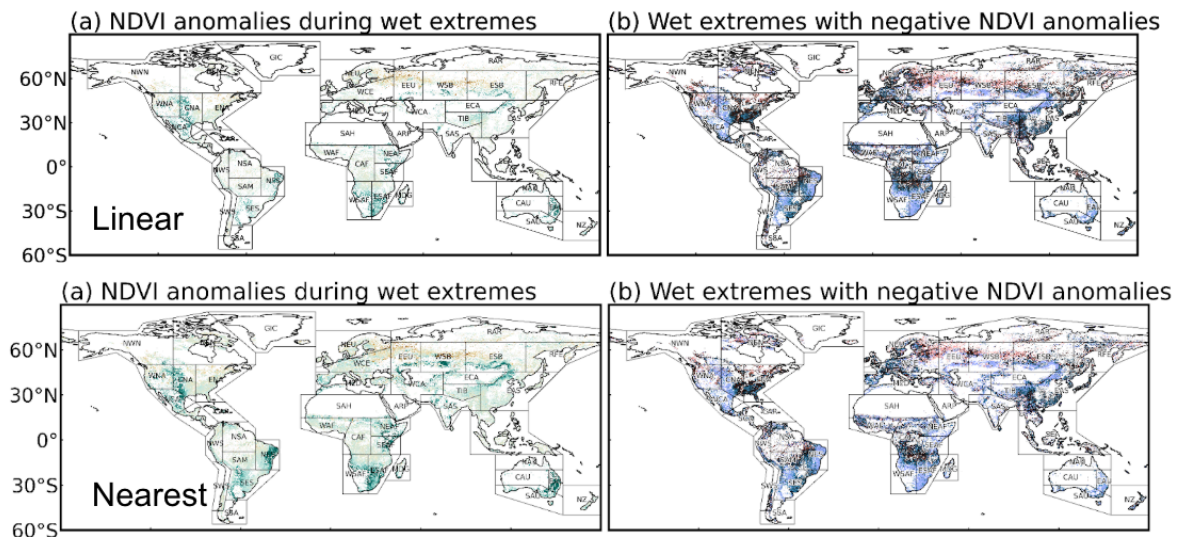
We used the 16-day MODIS NDVI product because it provides comparatively robust and spatially consistent vegetation estimates through the incorporation of multiple observations and strict quality-control procedures within each compositing period (Didan & Barreto, 2019). This choice was intended to better preserve observation-based vegetation responses during hydrological extremes while minimizing artifacts related to cloud contamination, view-angle effects, and atmospheric correction. Some nominally daily optical vegetation-related products can still rely on multi-day retrieval windows, compositing, or filtering procedures. For example, the MODIS MCD43A4 daily product is retrieved daily, but each daily retrieval is based on 16 days of input observations, with the day of interest emphasized and the product date corresponding to the center of a moving 16-day input window. We added this explanation for dataset choice concisely in the manuscript as well (lines 107-109).

Lines 107-109:

“We used the 16-day NDVI product because it provides comparatively robust and spatially consistent vegetation estimates through the incorporation of multiple observations and strict quality-control procedures within each compositing period (Didan & Barreto, 2019).”

Since the dataset user guide documentation indicates “The goal of the compositing methodology is to extract a single value per pixel from all the retained filtered data, **which will represent the pixel for the particular 16-day period**”, we applied the “nearest neighbor” method for interpolation in our actual analysis. At the same time, we note that our interpolation application can potentially introduce uncertainties. In order to quantify this potential bias, we used an alternative interpolation method (linear) to repeat the calculation

using wet extremes as an example. The spatial pattern comparison shows a limited influence of the way of interpolation (Figure R1). And we added the uncertainty of interpolation in lines 418-425.



Lines 418-425:

“In addition, the NDVI data used in this study are based on 16-day composite products that were temporally interpolated to daily resolution to align with daily soil moisture extremes. This interpolation may smooth short-term vegetation dynamics and introduce uncertainties in the timing and magnitude of rapid responses. As a result, fine-scale variability in vegetation responses may be dampened, and the exact alignment between NDVI and extreme events should be interpreted with a degree of caution. Future work could benefit from vegetation products with higher effective temporal resolution that combine frequent observations with robust quality screening and uncertainty characterization. This would help better capture rapid vegetation responses during hydrological extremes while accounting for known challenges in optical vegetation records, including cloud contamination, atmospheric effects and temporal gap filling.”

Reference:

Didan, K., & Barreto-Muñoz, A. (2019). MODIS Collection 6.1 (C61) Vegetation Index Product User Guide. The University of Arizona, Vegetation Index and Phenology Lab: Tucson, AZ, USA.

6. It is not clear which years the study period covers. Please specify.

Reply:

The results of this study are based on the period 2000-2023, noted on line 80 and also added in the caption of Figure 1.

Line 80:

*“Daily soil moisture data from ERA5-Land with a 0.1° resolution from **2000-2023** were used to identify hydrological extremes since soil moisture is directly related to water availability of vegetation (Liu et al., 2020).”*

Caption of Figure 1:

*“Figure 1: Number and average duration of wet and dry extremes during **2000-2023**.”*

7. Consider adding a table summarizing all predictors under pre-extreme condition, extreme characteristics, and environmental background for clarity.

Reply:

We appreciate this suggestion. For clarity, we have considered summarising the predictors in a dedicated table in Supplementary Table 1. And they are also described in detail in the manuscript, Method 2.3, lines 119-123.

Supplementary Table 1:

Table S1: Summary of all predictor variables used in the attribution analysis

<i>Category</i>	<i>Full name of variables</i>
Pre-extreme Vegetation Condition	Pre-extreme NDVI percentile
Extreme Characteristics	Extreme severity, seasonal timing of the extreme, accumulative shortwave radiation anomaly, accumulative vapour pressure deficit anomaly
Environmental Background	Tree cover, short vegetation cover, deciduous vegetation cover, evergreen vegetation cover, aridity index, topography, soil texture, long-term mean air temperature

Line 119-123:

“(1) Pre-extreme Vegetation Condition: The vegetation's greenness state before the current extreme event.

(2) Extreme Characteristics: Event-specific features such as extreme severity, seasonal timing of the extreme, shortwave radiation, and vapour pressure deficit (VPD) during the extreme.

(3) Environmental Background: Grid cell properties including tree cover, short vegetation cover, deciduous vegetation cover, evergreen vegetation cover, aridity index, topography, soil texture, and long-term mean air temperature.”

8. One of the main findings is that vegetation responses to wet extremes are highly heterogeneous (both positive and negative NDVI anomalies). The explanation of heterogeneous responses to wet extremes is needed.

Reply:

We thank the reviewer for acknowledging our finding that vegetation responses to wet extremes exhibit strong heterogeneity, and we agree that further discussion would be beneficial. We have therefore expanded the Discussion section to clarify two complementary reasons for this heterogeneity.

First, wet extremes are closely linked to precipitation characteristics, including its frequency, intensity and spatial distribution. These characteristics can vary substantially across regions and have been shown to induce context-dependent vegetation responses (Feldman et al., 2024). In contrast, dry extremes are often associated with larger-scale atmospheric anomalies and enhanced evaporative demand (Zhou et al., 2019). Consequently, wet-extreme characteristics are generally more spatially variable, contributing to more heterogeneous vegetation responses across regions (lines 328-332).

Second, the ecological consequences of excess soil moisture are strongly mediated by local environmental conditions, including soil drainage and water-storage capacity, topography, climatic background, and vegetation cover types. Therefore, similar wet soil-moisture anomalies may enhance vegetation greenness in some regions but induce stress in regions where excess moisture leads to waterlogging, oxygen limitation, or reduced root functioning. This interpretation is consistent with our attribution analysis, which shows that environmental background conditions have comparatively stronger relevance for vegetation responses during wet extremes than during dry extremes (lines 349-359).

Lines 328-332:

“Wet extremes, by contrast, are closely linked to precipitation-event characteristics, including event frequency, intensity and spatial distribution, which can vary substantially across space and contribute to context-dependent vegetation responses (Feldman et al., 2024). Together with local differences in soil drainage and water-storage capacity, this event-scale variability can lead to more spatially variable wet conditions and helps explain the more regionally heterogeneous vegetation responses observed during wet extremes.”

Lines 349-359:

“These mechanistic differences explain why environmental background exerts a stronger influence on vegetation responses during wet extremes in some regions. The consequences of excess moisture depend heavily on factors such as drainage capacity (linked to soil texture and topography), climatic constraints on evapotranspiration, and the thermal environment that shapes recovery potential (reflected in long-term mean air temperature and aridity). Vegetation cover types and conditions also modify how strongly waterlogging affects root functioning, leading to divergent greenness responses (Figure S9). In contrast, the direct water-stress exerted by dry extremes leaves less room for

background conditions to mediate the impacts. Vegetation responses primarily reflect the characteristics of the dry extreme itself. As a result, wet-extreme responses are substantially more heterogeneous across regions than dry-extreme responses, because the ecological consequences of excess soil moisture are more strongly modulated by local environmental conditions.”

References:

Feldman, A. F., Konings, A. G., Gentine, P., Cattri, M., Wang, L., Smith, W. K., ... & Poulter, B. (2024). Large global-scale vegetation sensitivity to daily rainfall variability. *Nature*, 636(8042), 380-384.

Zhou, S., Williams, A. P., Berg, A. M., Cook, B. I., Zhang, Y., Hagemann, S., ... & Gentine, P. (2019). Land–atmosphere feedbacks exacerbate concurrent soil drought and atmospheric aridity. *Proceedings of the National Academy of Sciences*, 116(38), 18848-18853.

9. In Figure 1, please clarify whether “Number of events” refers to annual counts or totals over the study period.

Reply:

It is the total amount of events during the whole 2000-2023 period. This information has been added to the caption of Figure 1.

Caption of Figure 1:

*“Figure 1: Number and average duration of wet and dry extremes **during 2000-2023.**”*

RC2:

This study uses NDVI anomalies as a proxy for vegetation response to wet and dry soil moisture extremes and uses random forest feature importance to relate anomalies to various drivers. I agree with the authors that comparing the response of vegetation to wet and dry extremes is necessary and important. The results reiterate previous studies which show that wet extremes generate a more heterogeneous response than do dry extremes (and often increase, rather than decrease, NDVI). The results also highlight the complex factors that influence NDVI response to both wet and dry extremes. My detailed comments are below:

Reply:

We thank the reviewer for the constructive comments and for recognizing the importance of comparing vegetation responses to wet and dry soil moisture extremes. We appreciate the positive assessment of the main findings, particularly regarding the heterogeneous nature of vegetation responses to wet extremes and the complex factors influencing these responses. Detailed responses to each comment are provided below, marked in blue.

Methods:

1. Line 71: Is 0-100cm of soil thickness the best choice for both dry and wet extremes? I could imagine that the surface soil moisture might be more relevant for wet extremes in particular.

Reply:

We thank the reviewer for this important comment. We aim to analyse vegetation responses to root-zone soil moisture, because sustained vegetation responses to hydrological extremes are expected to depend on water availability within the rooting zone rather than on near-surface soil moisture alone. This also applies to wet extremes. While we agree with the reviewer that surface soil moisture may reveal more wet extremes, we focus on deeper soil moisture in order to detect wet extremes where waterlogging can potentially affect most plant roots rather than only shallow roots.

Due to the lack of accurate global rooting-depth information, we used soil moisture integrated over the upper 1 m as a commonly applied proxy for root-zone soil moisture (e.g., Guo et al., 2026). Previous studies have also shown that soil layers below the surface layer are of most relevance for vegetation productivity (Li et al., 2021), and estimated rooting depths across much of the globe typically extend beyond the uppermost soil layers (Stocker et al., 2023). We also added this explanation in the manuscript lines 81-84.

Lines 81-84:

“We used the thickness-weighted average soil moisture of the top three layers (0-100 cm) as a proxy for root-zone soil moisture because vegetation responses are more strongly associated with sub-surface than near-surface soil moisture (Li et al., 2021), and rooting depths commonly exceed the uppermost soil layers across much of the globe (Stocker et al., 2023).”

References:

Guo, R., Wu, X., Wang, P. et al. Increased spread of global flash droughts threatens vegetation productivity resilience. *Nat Commun* 17, 4050 (2026). <https://doi.org/10.1038/s41467-026-70417-z>

Li, W., Migliavacca, M., Forkel, M., Walther, S., Reichstein, M., & Orth, R. (2021). Revisiting global vegetation controls using multi-layer soil moisture. *Geophysical Research Letters*, 48, e2021GL092856. <https://doi.org/10.1029/2021GL092856>

Stocker, B.D., Tumber-Dávila, S.J., Konings, A.G. et al. Global patterns of water storage in the rooting zones of vegetation. *Nat. Geosci.* 16, 250–256 (2023). <https://doi.org/10.1038/s41561-023-01125-2>

2. Line 84: Did you disregard events shorter than 5 days because of the 16-day NDVI resolution? If so, this reasoning (and its implications) is worth stating explicitly, since ideally shorter events could also be examined (particularly wet extremes). However, this 5 day threshold may still be too short given that you use 16-day NDVI and interpolate by repeating the same value for each day (Line 90). I would recommend further discussion of the implications and limitations of this resolution choice.

Reply:

The exclusion of events shorter than 5 days was not determined by the temporal resolution of the NDVI dataset, but was introduced to reduce the influence of very short soil moisture fluctuations that are less likely to induce measurable vegetation responses. For example, vegetation impacts of excess soil moisture, such as waterlogging stress, generally require sustained anomalous conditions rather than short-lived fluctuations. We have clarified this in the revised manuscript (lines 98-100).

Lines 98-100:

“Events shorter than 5 days were excluded to reduce the influence of transient soil moisture fluctuations that are less likely to induce measurable vegetation responses, particularly for processes such as waterlogging stress that require sustained anomalous conditions.”

We also acknowledge that the temporal interpolation of NDVI to daily resolution may limit the representation of rapidly evolving vegetation dynamics, particularly for short-duration extremes. We therefore clarified this limitation in the revised manuscript (lines 418-425).

Lines 418-425:

“In addition, the NDVI data used in this study are based on 16-day composite products that were temporally interpolated to daily resolution to align with daily soil moisture extremes. This interpolation may smooth short-term vegetation dynamics and introduce uncertainties in the timing and magnitude of rapid responses. As a result, fine-scale variability in vegetation responses may be dampened, and the exact alignment between NDVI and extreme events should be interpreted with a degree of caution. Future work could benefit from vegetation products with higher effective temporal resolution that combine frequent observations with robust quality screening and uncertainty characterization. This would help better capture rapid vegetation responses during hydrological extremes while accounting for known challenges in optical vegetation records, including cloud contamination, atmospheric effects and temporal gap filling.”

3. The RF description (Section 2.4) would benefit from several clarifications:

(1) Describe/define concurvity and its relevance (first mentioned in Line 142)

Reply:

We clarified the benefit of applying concurvity in the revised manuscript, lines 161-163. In our study, concurvity is used as a measure of collinearity among predictors, quantifying the extent to which one predictor can be explained by all other predictors (explained in lines 160-161).

Lines 161-163

“In contrast to correlation measures, concavity can additionally capture non-linear dependencies among predictors, which is particularly relevant for environmental variables with potentially complex relationships.”

Lines 160-161:

“..., where concavity (approximated using second-degree polynomial functions) measures the extent to which a predictor can be explained by all other predictors (Jiang et al, 2024).”

(2) Write ‘out of bag’ score the first time the acronym OOB is used

Reply:

We explicitly wrote “out-of-bag (OOB)” when the acronym is first introduced (line 159).

(3) How did you treat the wet vs dry extremes? Did you repeat the entire training/prediction process twice, once for dry extremes and once for wet extremes? (Line 152)

Reply:

Wet and dry extremes were analysed separately throughout the Random Forest attribution framework. Specifically, the entire predictor selection, model training, and attribution procedure was performed independently for dry and wet extremes. We clarified this explicitly in the revised manuscript (line 171).

Line 171:

*“We applied the algorithm to all identified wet or dry extremes **separately** in a climate reference region defined by the Intergovernmental Panel on Climate Change (IPCC) (Iturbide et al., 2020).”*

(4) You say you only considered negative NDVI anomalies? (Line 155) Doesn't this eliminate the possibility of detecting wet or dry extremes that cause an increase in NDVI relative to expectation? (Such as what you say in Line 181 about greening during wet extremes?)

Reply:

We thank the reviewer for this important comment. We have taken both perspectives on any vs. only negative greenness anomalies in our study. While the analysis of the vegetation greenness response to wet and dry extremes in Figure 2 considers both increased and decreased greenness responses, our attribution analysis focused specifically on negative NDVI anomalies during hydrological extremes. This is in order to isolate events where vegetation functioning was adversely affected. We agree that wet or dry extremes can also coincide with positive NDVI anomalies, for example, when additional water availability enhances vegetation greenness. However, combining positive and negative vegetation responses within the same

attribution framework would mix fundamentally different ecological processes and could obscure the interpretation of predictor effects. For instance, greater wet-extreme severity may promote greening in some extremes where vegetation benefits from more water (e.g., in dry regions), while causing vegetation stress and reduced greenness in other extremes (e.g., through waterlogging in wet regions). We therefore focused the attribution analysis on negative NDVI anomalies to specifically investigate the factors associated with vegetation stress responses during hydrological extremes. We clarified this in the revised manuscript, lines 175-177.

Lines 175-177:

“This was done to specifically isolate vegetation stress responses, as combining positive and negative NDVI anomalies would mix fundamentally different ecological processes and complicate the interpretation of predictor effects.”

And the attribution results for considering extremes with both positive and negative NDVI responses are shown in supplementary Figure S10-11 for comparison.

Results:

4. Figure 1. Would the number of events be proportional to the timeseries length, if events are chosen based on percentile thresholds? Does this mean that the different number of events per IPCC area is due to different numbers of pixels in each region or due to the number of events that could not be detected due to poor quality NDVI data?

Reply:

Since hydrological extremes were identified using percentile-based thresholds, the number of extreme days depends on the length of the soil moisture time series (2000-2023 here) and the chosen thresholds. However, event counts are also influenced by how long extremes persist, because consecutive extreme days are grouped into events and short events (<5 days) are excluded. In addition, some events were removed when sufficient reliable NDVI data were not available.

Figure 1 presents event counts at the grid-cell level rather than aggregated statistics for each IPCC region; the IPCC boundaries are shown mainly to facilitate comparison with the regional attribution analyses. Therefore, differences in the number of events shown in Figure 1 do not directly reflect different numbers of pixels within each IPCC region.

To clarify this point, we added the following text to the Results section, lines 193-194 and lines 196-197.

Lines 193-194:

“Because consecutive extreme days were grouped into individual events, spatial differences in event counts also reflect how soil moisture anomalies persist and recur over time.”

Lines 196-197:

“Consequently, spatial differences in the number of retained events also reflect regional differences in NDVI data availability.”

5. Additionally, you mention in Line 173 that “the most common reason for excluding an event was insufficient NDVI data availability.” Did you see that wet extremes were associated with reduced NDVI data quality to a greater degree than dry extremes?

Reply:

The spatial patterns of NDVI data availability were generally similar for wet and dry extremes (Figure S8), suggesting no systematic reduction in NDVI data quality specifically associated with wet extremes spatially. However, the final number of events retained for the attribution analysis differed between wet and dry extremes across regions. This difference mainly reflects differences in the occurrence frequency of the detected extremes and the event filtering criteria applied in the study. Specifically, for each event, we required sufficient NDVI data availability before, during, and after the extreme peak, with at least one-third of the expected observations available in each period. The same quality-control criteria were applied consistently to both wet and dry extremes to ensure comparable reliability of the analysed events. The final numbers of retained events and selected RF models for each IPCC region are added in Figure S4.

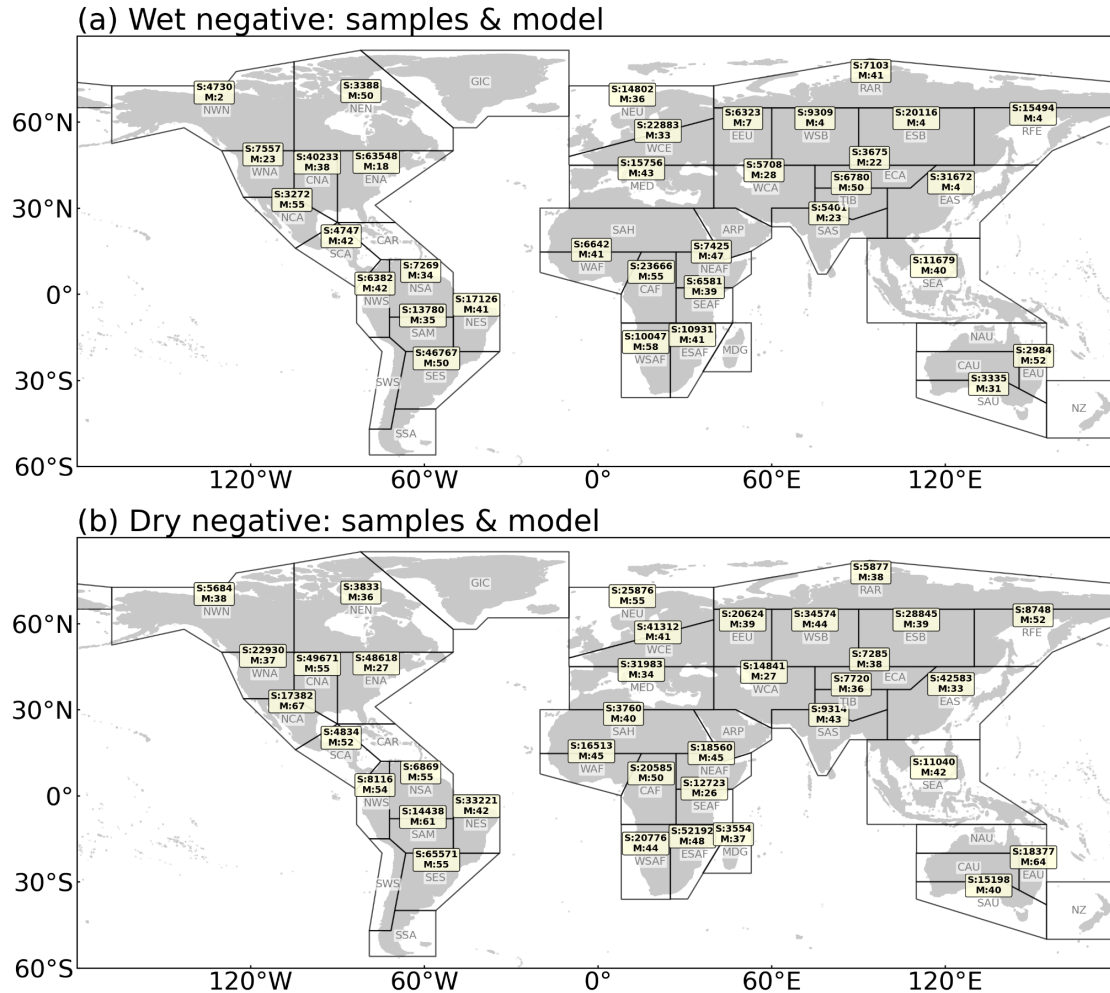


Figure S4: Number of extremes(S) and number of models that are finally selected (M) in each IPCC region for (a) wet and (b) dry extremes.

6. You show the data availability in Figure S7 and it appears to be the majority of the land surface area, so discussing these issues is important for contextualizing the frequency of extreme events you report.

Reply:

We agree that NDVI data availability is important for contextualizing the reported spatial patterns and frequencies of vegetation responses to hydrological extremes. Regions with low data availability (shown in yellow in Figure S8) should not be interpreted as regions with fewer hydrological extremes, but rather as regions where a substantial fraction of detected events could not be included because insufficient high-quality NDVI observations were available for the required analysis periods. Consequently, the representativeness of the analysed events varies spatially and is lower in regions with limited NDVI availability. To make this clearer, we revised the Figure 2 caption and corresponding text to explicitly note that spatial differences in event representation partly reflect differences in NDVI data

availability and event filtering criteria, and we direct readers to Figure S8 for reference (lines 205-206).

Caption of Figure 2, lines 231-232:

“Spatial representativeness may vary across regions due to differences in data availability and event filtering; corresponding data availability flags are shown in Figure S8.”

Line 205-206:

“The spatial representativeness of these patterns may vary across regions because events with insufficient NDVI data availability were excluded from the analysis (Figure S8).”

7. You provide the OOB score and concurrency information in Figure S5, but the RF model performance is very important for interpreting your results. Could you report on the performance more explicitly in the main text? I would like to know how the RF performance varies across IPCC regions and across the wet vs dry extreme models and how this impacts your results.

We thank the reviewer for this important comment. We agree that the RF model performance is essential for interpreting the attribution results. We therefore additionally clarified that RF performance varied across IPCC regions and between wet and dry extreme analyses, reflecting differences in the predictability of vegetation responses and the extent to which the selected predictors explain regional variability in lines 277-289.

Line 277-289:

“The robustness and interpretability of these regional attribution patterns depend on the explanatory performance of the RF models and the degree of concurrency among the selected predictors (Figure S5). The average OOB scores of the selected RF models across all regions ranged from 0.22 to 0.62 (SD = 0.11), while the average concurrency values ranged from 0.23 to 0.66 (SD = 0.11). RF model performance varied across IPCC regions and between wet and dry extremes. Lower OOB scores and higher concurrency values were generally found in regions with relatively few retained events, such as Eastern Central Asia (ECA) for wet extremes and Russian-Arctic (RAR) for dry extremes (Figure S4). Lower explanatory power was also observed in regions where vegetation responses to hydrological extremes were comparatively weak or spatially heterogeneous, such as Western Central Africa (WAF) for wet extremes and Central Africa (CAF) for dry extremes (Figure 2). In contrast, regions with stronger and more spatially coherent NDVI anomalies during extremes generally showed higher OOB scores and lower concurrency among predictors. Overall, OOB scores were slightly lower for wet extremes than for dry extremes, consistent with the more heterogeneous vegetation responses observed during wet extremes. Nevertheless, the RF models generally retained moderate-to-high explanatory power and low-to-moderate concurrency across most regions, indicating that the identified attribution patterns are robust and that the selected predictors capture meaningful variability in vegetation responses to hydrological extremes.”

8. Paragraph starting at Line 195: It would be helpful to clarify here whether the results refer to wet or dry extremes.

We added which extremes the texts refer to.

Lines 221-224:

*“While Figure 2a and 2c provide an overview of mean vegetation responses, the low consistency of NDVI anomalies across individual events in regions such as eastern North America (ENA), central Africa (CAF) **for wet extremes** (Figure 2b) and boreal forests in Eurasia **for dry extremes** (NEU, EEU, WSB) (Figure 2d) highlights that even under similar environmental conditions, vegetation responses to hydrological extremes can differ substantially.”*

9. Figure 2. The map is great, but because you are showing the average NDVI anomalies, the information in Figure S6 (with the distribution of the anomalies) is possibly more relevant for the main text, especially because so many of the IPCC regions have both negative and positive anomalies.

Figure 2 (map) and Figure S7 (boxplot) were designed to highlight complementary aspects of vegetation responses to hydrological extremes. Figure 2 primarily emphasizes the spatial patterns of mean NDVI anomalies and the spatial distribution of events associated with negative NDVI anomalies, thereby providing a global overview of where vegetation greening or browning responses are more prevalent. In particular, panels (b) and (d) explicitly show that both positive and negative NDVI anomalies can occur within the same regions and even within nearby grid cells. In contrast, Figure S7 focuses on the statistical distribution of NDVI anomalies within each IPCC region and is therefore intended to complement the spatial information shown in Figure 2 rather than replace it. We nevertheless agree that Figure S7 provides important additional context for interpreting the regional heterogeneity of vegetation responses and now refer to it more explicitly in the main text (lines 224-225).

Lines 224-225:

“The regional distributions of NDVI anomalies further illustrate this heterogeneity and the coexistence of positive and negative responses within most IPCC regions (Figure S7).”

10. Line 208-210: You describe the importance of antecedent conditions for wet extremes very definitively here, but in the abstract the takeaway seems to be that all of the factors you looked at are important. It would be helpful to clarify the main paper takeaways (for example if there are consistently important drivers or differences in drivers between the wet vs dry extremes) and make this message consistent between the text and the abstract. In general, it is a little bit hard to follow in the text which factors you describe as important for controlling the negative vs positive anomaly attribution for the wet vs dry extremes.

We thank the reviewer for this important comment. We revised the manuscript to clarify the hierarchy of influential factors across regions as a summary at the end of the Result section 3.3 (lines 251-253). In particular, we emphasize more explicitly that pre-extreme vegetation conditions and extreme characteristics are generally the dominant predictor categories,

whereas environmental background variables usually play a secondary but regionally important role, especially for negative vegetation responses during wet extremes. We also revised the Abstract (lines 16-19) to better distinguish the dominant controls for wet versus dry extremes and to ensure consistency with the regional attribution results presented in Figure 3.

Lines 251-253:

“Overall, pre-extreme vegetation conditions and extreme characteristics constitute the two most influential categories across most regions, while environmental background generally plays a secondary but regionally important role, particularly for negative vegetation responses during wet extremes.”

Abstract, Lines 16-19:

“Negative vegetation responses during wet extremes are most consistently associated with pre-extreme vegetation conditions, while environmental background variables such as climate (e.g., long-term mean air temperature, aridity) and topographic variability show comparatively stronger relevance than for dry extremes. Extreme characteristics also contribute substantially, though with greater regional variability than for dry extremes.”

11. Figure 4: It would be helpful to see how far ahead the ‘leading variable’ is.

We thank the reviewer for this helpful suggestion. To better illustrate the relative dominance of the leading predictor within each category, we added an indicator in Figure 4 showing regions where the most relevant variable was selected substantially more frequently than the second-ranked variable. Specifically, an asterisk next to the region label indicates that the most relevant variable was selected at least 50% more frequently than the second most relevant variable within the respective category. This addition helps distinguish regions with a clear dominant predictor from those where multiple variables show similarly high relevance.

New Figure 4:

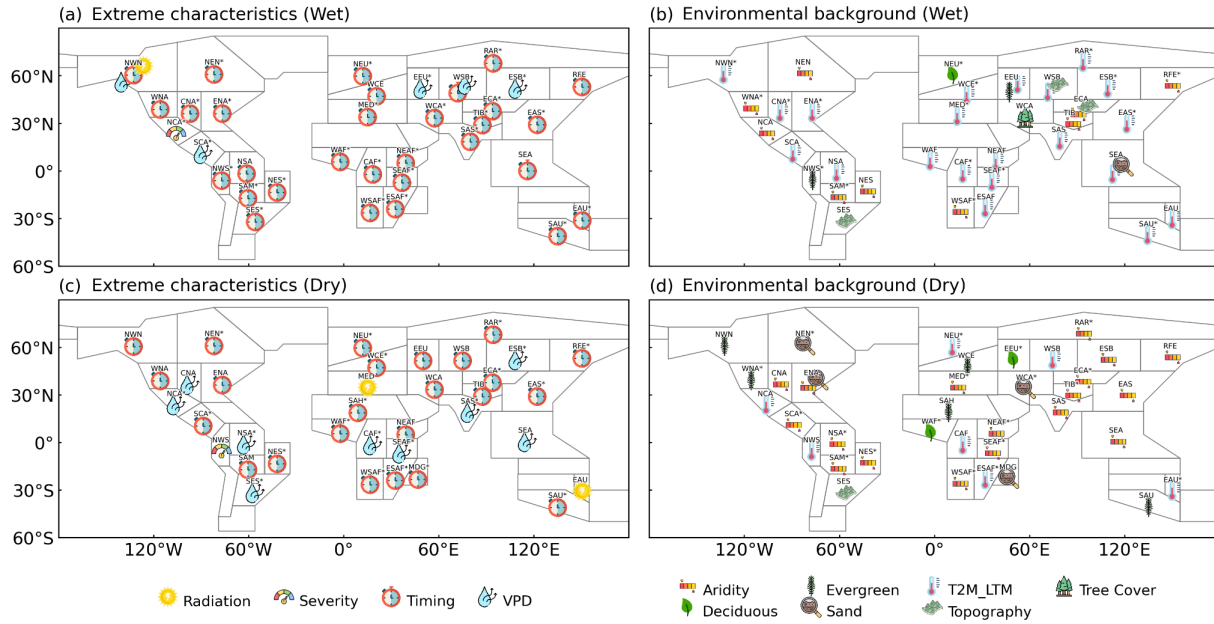


Figure 4: Most relevant variables for negative vegetation responses during wet (a,b) and dry (c,d) extremes in terms of extreme characteristics and environmental background. Multiple variables are shown in case their relevance is the same according to the ranking. **The asterisk next to each region label indicates that the most relevant variable was selected at least 50% more frequently than the second most relevant variable within that category.** Note that the attribution analysis is based only on extremes with negative NDVI anomalies.

12. Section 3.4 (Line 246ish): When you look at the two regions with the most negative NDVI anomalies during wet and dry extremes, is it fair to look at just the negative NDVI response for wet extremes, when you found such a heterogeneous response to wet extremes, even for the sign of the NDVI anomaly? It would be worth mentioning this.

We agree that vegetation responses to wet extremes are substantially more heterogeneous than responses to dry extremes, including differences in the sign of NDVI anomalies. In Result 3.4, we specifically focused on negative NDVI anomalies because the objective was to investigate vegetation stress responses associated with hydrological extremes. The selected regions with the strongest negative NDVI anomalies during wet extremes were used as representative examples of such stress responses and also exhibited relatively high proportions of wet extremes associated with negative NDVI anomalies (e.g., 53% in Eastern Europe (EEU) and 46% in Northern Europe (NEU)). We nevertheless clarified in the revised manuscript that these regions still contain substantial event-to-event variability and do not represent uniformly negative vegetation responses to wet extremes (lines 294-297).

Lines 294-297:

“Although vegetation responses to wet extremes are regionally heterogeneous and can include both positive and negative NDVI anomalies, EEU and NEU also exhibited relatively

high proportions of wet extremes associated with vegetation browning (53% in EEU and 46% in NEU), supporting their selection as representative case-study regions.”

13. Small comments:

Line 54: You use several different phrases to refer to antecedent conditions. Consider picking one term.

We kept using “pre-extreme vegetation conditions” now in the revised manuscript.

Line 61: I wouldn’t necessarily describe this RF approach as “new,” although it does seem effective.

Revised to “effective” in line 71.

Line 117: Which version of GLEAM did you use?

GLEAM 4.1a (now added in line 135)

Line 135:

“..., and evapotranspiration from the Global Land Evaporation Amsterdam Model (GLEAM **4.1a**).”

Grammar/punctuation errors:

Line 35, “...stomatal closure, decrease...”

Line 40, “oxygen deficiency add nutrient uptake”

We have revised the manuscript correspondingly.

Lines 39-40:

*“Dry extremes affect vegetation by inducing water stress, leading to stomatal closure (Knipfer et al. 2020) **and decreased carbon uptake.**”*

Lines 43-44:

*“Excessive soil moisture can directly impair vegetation through waterlogging, which can increase susceptibility to pathogens or fungal infections, cause root oxygen deficiency, and **limit nutrient uptake** (Li et al., 2019).”*