

## Reply to Reviewer #2

We would like to thank reviewer #1 their valuable input and thorough review, which significantly helped us to improve the manuscript.

### **General Comments:**

*This study explores a relevant topic that requires further attention from the scientific community. I acknowledge the considerable effort made by the authors in organizing and analyzing the data. However, I have significant concerns and questions regarding various aspects of the study, as outlined in my comments below.*

*The authors do not provide a clear biological definition of "forest damage." From an ecological perspective, considering that this study operates at the interface between population and community scales, forest damage should ideally be assessed by indicators such as a reduction in the reproduction or survival of tree populations or a phytosociological shift within the community (note that changes in species composition over time can also alter the NDVI response of a forest). However, none of these ecological proprieties can be inferred solely from NDVI. NDVI indirectly measures photosynthetic activity, which can be temporarily reduced by many factors unrelated to hydro-meteorological variations, such as changes in nutrient absorption or pathogen infestations affecting aerial or root systems. At the ecosystem scale, NDVI reduction could reflect changes in productivity by capturing decreases in the photosynthetic biomass of forests. If this photosynthetic biomass reduction were detected in NDVI long-term trends and supported by direct field measurements, there could be evidence of "forest damage", but the authors cannot demonstrate this in this work.*

Thank you for your relevant comment on the definition of forest damage.

While reduction in the reproduction, or survival of tree populations, or phytosociological shifts in the community would be great measures, we are not aware of datasets providing this information for all European forests, between 1981 until present.

With NDVI, we extract episodes of forest browning at unusual times in the year (in summer), which is a type of damage to the forest, although normally short-time damage for broad-leaved species. Generally, it corresponds to leaves drying in the middle of the summer. The trees usually survive, they have new leaves the next year, but they can be weakened, as written in discussion with potentially reports from one year to the next one. For conifers, browning is generally a sign of death.

Moreover, we study here NDVI changes on a short-term basis (anomalies over 2 months), which excludes a change in soil nitrogen: such a change will not impact the NDVI in a short-term basis. The ecosystems are strongly resilient to changes in nutrients, and the recycling rate of nitrogen is very high. Therefore, a deficiency in nitrogen can affect a forest only on a long-term basis. And the same argument is true for shifts in community. If the community is changing, e.g. shift of species, as broad-leaved species to conifers, at least 40 years are necessary to record such a shift, the time for a complete change of generation. Following your remark on the CORINE mask (see further response), we anyway discarded grid points experiencing a land cover change, preventing a major change in NDVI due to a shift in forest type from coniferous trees to broadleaved trees and vice versa.

We would like to highlight that we do not discard the possibility of an intermediary driver for NDVI loss, such as forest fire, insect infestation or other pathogens. The preceding adverse meteorological

conditions increase the forest fire chances and the vulnerability of trees to insect attacks/diseases of the aerial or root system (Rebetez et Dobbertin, 2004). Our goal is to quantify to what we can predict large scale vegetation browning by using solely hydro-meteorological drivers. Following your remark, we added a sentence in the discussion section: “A causality analysis (Peters et al., 2017) could explore the role of previous summers’ forest state or insect infestations as predictors, although this framework is beyond the scope of our study.”

While we do not compare our results with field measurement (beyond the scope of this study), we provide a summary of the link between observed low-NDVI events and hydro-meteorological conditions, for forests all over Europe. In the discussion section, we added a sentence to address this point: “For smaller-scale operational forecasts, such as those for small-scale forest management, we recommend following our method but training the model on higher spatial resolution data, such as station observations of hydro-meteorological data and forest field measurement, if available for long periods. Additionally, we suggest using our verification method, with both CSI and AUC, to measure accuracy.”

To clarify our definition of damage we reformulated a sentence in the introduction section: “We employ a random forest (RF) classification to predict summer forest damage, i.e. July-August low greenness events.”

In the introduction, we also highlight further our goal to predict large scale vegetation browning by using solely hydro-meteorological drivers.

“This study aims to predict large-scale summer forest browning employing solely hydro-meteorological conditions that are available from S2S forecast model output, providing valuable lead time for forest managers to anticipate and mitigate potential impacts.”

Rebetez, M., & Dobbertin, M. (2004). Climate change may already threaten Scots pine stands in the Swiss Alps. *Theoretical and Applied Climatology*, 79, 1-9.

Wermelinger, B. (2004). Ecology and management of the spruce bark beetle *Ips typographus*—a review of recent research. *FOREST ECOLOGY AND MANAGEMENT*, 202(1), 67-82.

<https://doi.org/https://doi.org/10.1016/j.foreco.2004.07.018>

*The manuscript's structure does not adhere to the conventional format (Introduction, Materials and Methods, Results, and Discussion), compromising the overall coherence and readability. For instance, Section 5.2, which follows the discussion, introduces methodological aspects and essential criteria that should be included in the description of the methods. Therefore, I strongly recommend a comprehensive reorganization of the text to enhance its clarity and logical flow.*

Thank you for pointing this out and our apologies for the confusing section labeling. To avoid any confusion regarding the aim of Section 5.2, that does not introduce new methods, we changed its title to “Dataset limitation” and added the introductory sentence : “In this subsection, we discuss the limitations associated with the datasets used in this study.”

*The information regarding forest data represents a fundamental layer of this work, but it is scattered in different sections of the text (e.g., lines 103-105, 238-242, and 297-309). The authors should gather all the forest information in a single opening subsection of the materials and methods section, providing more details for the readers. For example:*  
*- If the land cover product generated by CORINE is separated into three classes, why did the authors analyze only two?*

Thank you for this remark. The choice of not showing mixed forest, was a trade-off between number of figures and showing results between biologically significantly different forest types. The most important variable for the third forest type, mixed forest, is a mixed between the signal brought by broadleaved forest and coniferous forest.

Following the suggestion of reviewer #1, we added a complementary figure to interpret further the model's output (Fig.E1) and added the following interpretation.

“We do not identify any specific regions where temperature (in any month or season) or soil moisture (in any month, season, or depth) consistently dominate as the most influential predictors; their importance appears relatively evenly distributed across Europe (FigE1, left and middle panels). Variables from the preceding year are among the top 10 predictors at most grid points, with a moderately greater importance observed at higher latitudes (Fig E1, right panel).”

*- What is the criterion for considering 2018 as the reference year for vegetation? Wouldn't it be better to identify the areas (pixels) where there was no change in land cover over the period available in CORINE products?*

Thank you very much for your relevant comment. The primary motivation when choosing 2018 as a forest mask was to use the most recent mask available. Following your remark, and a remark from Reviewer #1, we used all the masks available for the whole Europe over the studied period and discarded 0.01° pixels that experienced a forest change between 2006, 2012 and 2018. We then reprocessed the NDVI binary time series. We re-ran the RF model for all grid points and updated all our figures. 189 out of the 1248 initial grid points were discarded with this procedure. The take home message of the figures and the manuscript in general did not change, only minor differences in the specific number of grid points including certain variables as a top 10 predictor (Figure 4).

We adapted the method section and hope this change helped consolidating the description of forests:

“The forest coverage and forest type (broad-leaved, coniferous, mixed forests, see Fig.D1) are extracted from the CORINE (Coordination of Information on the Environment) Land Cover for the reference years 2006, 2012 and 2018 (EEA, 2020a, b, c). We discard 0.01° pixels that are not classified as forest, or that experience a forest change between 2006, 2012 and 2018 (i.e. a change in land cover or forest type).”

*- Based on the literature, the authors could briefly describe the seasonal patterns of NDVI in European forests so that readers have a better understanding of the inherent fluctuations of this variable.*

Thank you for your remark. We added the following sentence at the beginning of section 2.2:

“In European broadleaved forests, the NDVI tends to follow a bell-shaped pattern, starting with a low NDVI in spring, before budburst, reaching a maximum around July and decreasing in autumn when leaves turn yellow and fall (Klisch and Atzberger, 2014). The NDVI of coniferous species is less sensitive to phenological changes. However, these forests typically include some deciduous species, particularly in the herbaceous and shrub layers, which contribute to seasonal variations in NDVI similar to those observed in broad-leaved forests, although less pronounced (Jönsson and Eklundh, 2002).”

*- In their maps, the authors should also provide a line (or another feature) clearly showing what they consider southern and northern Europe.*

Thank you for your remark. We added a grid for latitudes and longitudes on Figure D1, and we now refer to coordinates in addition to southern/northern Europe, when relevant (roughly 50 °N).

*One of the crucial issues of this work is the inclusion criterion of pixels with only 10% forest cover and the lack of control over the heterogeneity of land cover within the pixel. Some cases shown in Figures 1 and D1 include some of the large metropolitan areas of Europe (e.g., Lisbon and Naples) that can hardly be considered as forests. Additionally, many of these pixels also include agricultural or forestry areas, where managing these sites can significantly affect the NDVI result independently of climatic factors. In other words, the assumptions listed in section 5.2 (lines 297-309) do not seem appropriate to me, and the authors cannot guarantee that the variation in the NDVI of the analyzed pixels is only affected by meteorological fluctuations.*

Thank you for highlighting this point. The land cover mask was applied at a very early stage of NDVI data processing, when the AVHRR data is still on a 0.01° grid (about a km). All the 0.01° pixels that are not classified as “forest” (may it be coniferous, broadleaved or mixed) by the CORINE land cover mask, are set to missing values. When upscaling from 0.01° to 0.1°, we only take the average for NDVI in pure forest pixels. If a 0.1° box contains less than 10% of 0.01° pixels that are forest, we set the 0.1° box to missing data.

Following your remark, we added the following sentence in section 2.2:

“We discard 0.01° pixels that are not classified as forest, or that experience a forest change between 2006, 2012 and 2018 (including change in forest type).”

We also added “for 0.01° forest pixels” in step 1. of the binary forest damage definition, to highlight that we only consider pure forest pixels at the 0.01° resolution.

*This concern could be minimized in different ways, including:*

*- Utilizing NDVI products with higher spatial resolution (e.g., Landsat).*

Thank you for your comment. The advantage of using AVHRR data is a good trade-off between spatial availability (for a large-scale analysis) and temporal availability (for a robust statistical analysis/robust fitting of the random forest). Satellite products with a higher spatial resolution (10m -30m, see <https://www.sciencedirect.com/topics/earth-and-planetary-sciences/high-spatial-resolution>) could not be used because of their low temporal resolution (Landsat series) or limited time coverage (Sentinel-2). Landsat 8 has for example a revisit frequency of 16 days under clear sky conditions, which means that in case of cloudy conditions very few observations are available. Fig 1 in <https://doi.org/10.1038/s43017-022-00298-5> provides a nice overview of the different sensors available.

To address this point, we added the following sentence in section 2.1: “The advantage of the AVHRR dataset over Europe are the long period of data availability, from 1981 to 2022 (for our analysis), and the good trade-off between temporal resolution (10-day) and spatial resolution (0.01°W× 0.01°N, effective footprint of approximately 1 km<sup>2</sup>)”

*- Selecting only pure forest pixels to enhance accuracy.*

Thank you for your remark, we hope we clarified this point in our response above and the added sentence about the use of pure forest pixels

*- Throughout the analyzed period, the authors could use control areas to mitigate potential biases arising from land use and land cover changes or stochastic events (e.g., wildfires, windstorms, pest outbreaks).*

Thank you for your suggestions. We indeed included a mask to remove grid points experiencing a land cover change (see response above).

Regarding forest fire and pest outbreaks, we refer to our above reply to your first remark. We do not exclude the presence of intermediary factors, but we rather try to assess whether hydro-meteorological conditions can be a good predictor for low-greenness events, including browning and forest dieback.

Regarding windstorm, we use the definition of low greenness events from Hermann et al. (2023). With this approach, we identify large-scale and persisting low NDVI events, while the impacts of windstorm are more localized, as specified in the last paragraph of section 4.1 in Hermann et al. (2023) and the remark at the end of section 5.1 in our manuscript: “Storm damage, though potentially extreme, is more localized (Hermann et al., 2023) and should be assessed with higher resolution NDVI datasets (Giannetti et al., 2021).”

*I have some concerns regarding the hydro-meteorological predictive variables used in this study. The authors include a large number of variables but do not mention any tests to assess autocorrelation among them. Given the similarity of certain variables (e.g., temperature and dew point temperature or the three soil moisture measurements), there is a high likelihood of significant correlation. While the Random Forest method is robust, it is not entirely immune to biases introduced by highly correlated variables. Therefore, I recommend conducting an exploratory analysis to evaluate and account for potential multicollinearity before selecting the final set of variables.*

Thank you for your remark. Indeed, several predictors could be correlated. However, each of them can still carry important, independent information (dewpoint information about water content in the air, different levels for soil moisture = different intensity and persistence of droughts/very moist conditions, rather than just instantaneous precipitation). We chose a low number (3) of predictors per decision tree in the random forest model, to disentangle better individual effects of each predictor, so that the algorithm would select the most informative ones. Following your remark and remark of reviewer #1, we added the following sentence in section 3.1: “Moreover, using a low number of predictors per tree helps reduce the dominance of correlated variables in the model (Strobl et al., 2008; Gregorutti et al., 2013).”

The idea behind keeping all the variables is to have the same set of predictors across all Europe, for a better a posteriori comparison across regions and forest types. It would be possible to perform a pre-selection of specific predictors, with correlation or principal components for example, however the comparison of the Random Forest output would be less straightforward.

Moreover, we compared the performance with the LASSO regression, designed to deal with correlated predictors, and the Random Forests performed better for 80% grid points (section 4.1).



*Concerning rainfall data, did the authors consider using datasets with higher spatial resolution, such as CHIRPS (Rainfall Estimates from Rain Gauge and Satellite Observations)? It is well established that regions with rugged topography or predominant convective rainfall during the summer can exhibit significant spatial variability in precipitation volume. Therefore, caution is necessary when extrapolating these data to neighboring areas. This may have influenced the results in some way, as rainfall volume was not identified as a primary variable influencing NDVI variations. This finding is unexpected, given that water plays a fundamental role in stomatal opening and photosynthesis, which are essential for regulating metabolism and maintaining plant homeostasis. Wind is also a variable that may have influenced the results (e.g., removal of leaves in extreme events or by microclimatic alteration of the leaf boundary layer). However, I understand the reasons that led the authors not to consider the use of this variable.*

Thank you for your suggestion and remark about precipitation data. We did not consider a dataset with higher spatial resolution, as we needed resolution consistency between all the variables. In this context, ERA-5 reanalysis is a good tradeoff, with a long temporal availability for all the variables we wanted to consider, as stated in section 2.3: “Reanalysis data offers dynamically consistent variables with a large, uniform spatio-temporal coverage.”. A dataset with higher resolution could capture more variability and local information about the hydro-meteorological conditions. However, we decided to aggregate spatially the high resolution AVHRR data to 0.1° to compute large-scale low NDVI events. Moreover, CHIRPS is only available for latitudes lower than 50°N. Lastly, rather than exact precipitation volumes, we use monthly and seasonal anomalies, which reduces the impact of potential bias in ERA-5 precipitation.

We addressed your remark regarding the importance of precipitation as a predictor, with the following paragraph in the discussion section:

“Although precipitation anomalies are known to be significantly linked with forest damage (Hermann et al., 2023), this variable has a relatively low explanatory power to predict low NDVI in our model, compared to soil moisture for example (Fig. 4). While these two variables are strongly correlated, soil moisture may provide persistence information that precipitation does not. Another possible explanation is that large areas of forests at European scale are set on flat areas with important water reserve or available groundwater table. Hence, the reaction of forests to long droughts is potentially more correlated to soil water content than to precipitation. This result highlights the added value of an automatic procedure to select the most important features among a large set of potentially important variables.”

*The authors opted to present the NDVI time series as a secondary outcome of the study, including only an appendix figure (Figure B1) and omitting an in-depth analysis in the results and discussion sections. This approach creates a disconnection between the introduction, where climate change is a central theme, and the remainder of the paper. Therefore, I recommend incorporating Figure B1 as a primary figure within the main text and that the authors highlight the annual variations and long-term trends of NDVI in the manuscript.*

Thank you for your suggestion.

We added sentence regarding the seasonal cycle of NDVI (as mentioned above).

While we understand the value of visualizing the time series of forest damage (Fig.B1), our primary aim is not to propose a new method for forest damage detection, but we simply extended the temporal coverage of low NDVI as in Hermann et al. (2023). We use the same definition, a different data source, and obtain similar patterns for the overlapping period (2002-2022). Consequently, we decided to include Figure B1 in the supplementary material rather than in the main text.

We would like to clarify that our study is not an attribution study aimed at identifying causal links to climate change. Instead, we focus on quantifying the statistical relationship between forest damage and interannual anomalies in NDVI and hydro-meteorological variables. To this end, we removed long-term trends in these drivers to minimize the influence of ongoing climate change—particularly the pronounced temperature trend—and better isolate year-to-year variability that may be associated with browning events.

We added the following paragraph in the discussion section:

“We removed long-term trends from NDVI and hydro-meteorological variables to focus on interannual variability rather than gradual climate change. This allows us to isolate the statistical relationship between short-term anomalies and forest damage. While detrending reduces the influence of the mean climate trend, it does not eliminate the increased frequency of extreme events, which may still reflect underlying climate shifts and remain visible in the anomaly patterns.”

### **Specific Comments:**

**Lines 22-23:** *The phrase "In particular, dry and hot conditions can be highly detrimental to forests (Brodribb et al., 2020)" should be revised for greater clarity and accuracy. It is important to note that hot and dry weather (and climate) constitutes the natural conditions of numerous forest formations worldwide, such as the tropical seasonal dry forests of northeastern Brazil, the eucalyptus forests of Australia, and various forests and woodlands of the Iberian Peninsula, for example. While Brodribb et al. (2020) highlight the potential dangers posed by rising temperatures and reduced water availability to forest ecosystems, this crucial context is not conveyed effectively in the original sentence.*

Thank you for this precision. To be more specific in this introductory sentence, we added the adjective “anomalously”, the notion of time, to highlight that we refer to large deviation from “normal” climate conditions, and some more references:

“In particular, anomalously dry and hot conditions, especially on an unusually long period, can be highly detrimental to forests (Leuzinger et al., 2005; Senf et al., 2020; Brodribb et al., 2020).”

Leuzinger, S., Zotz, G., Asshoff, R., & Körner, C. (2005). Responses of deciduous forest trees to severe drought in Central Europe. *Tree Physiology*, 25(6), 641-650.

**Lines 28-30:** *In both phases, while these statements may be appropriate within the original context, they appear somewhat generic in this instance. It is important to note that some forest ecoregions, such as the northwestern Amazon and the La Plata Basin, are experiencing increased rainfall due to climate change. In the same way, the conclusion of Meier et al. (2021) seems restricted to the case of broad-leaved trees in Switzerland. Therefore, I recommend that the authors specify the region they are referring to. The same observation applies to other parts of the introduction section.*

Thank you very much for pointing this out. We changed structure of the sentence, to refer more precisely to the outcome of the cited paper:

“Moreover, longer growing seasons can amplify drought effects due to earlier spring leaf unfolding, as shown for example by Meier et al. (2021) in Swiss broadleaved forests.”

**Line 44:** *As this is the first occurrence of the acronym ECMWF in the text, the authors should provide its full designation. Other similar cases are scattered throughout the different sections of the manuscript, such as EUMETSAT and MODIS.*

Thank you, we changed the text accordingly by adding the definitions of these acronyms.

**Lines 44-56:** *At the end of the introduction, the authors describe the advantages of using ECMWF forecast models. However, this approach disrupts the logical flow of the text and does not effectively communicate the study's primary objectives. To enhance clarity and coherence, I recommend that the authors revise the final paragraph of the introduction to outline their research objectives and hypotheses explicitly. This revision should establish a clear connection between climate change and vegetation responses, ensuring a more structured and engaging introduction.*

Thank you for your valuable input. We added a sentence at the end of the introduction (also following your previous remark). This sentence connects back to the paragraph about protection measures and S2S forecast of hydro-meteorological drivers.

“This study aims to predict large-scale summer forest browning based solely on hydro-meteorological conditions, which are accessible through S2S forecasts, providing valuable lead time for forest managers to anticipate and mitigate potential impacts.”

We also slightly modified the previous sentence, to underline the hypothesis that the important predictors are location-specific: “From a large set of potential predictors, our goal is pinpoint crucial, location-specific periods and variables impacting European forests.”

Lastly, we moved the sentence “In addition, machine learning can leverage extensive data to predict vegetation states based on weather and climate conditions (John Nay and Gilligan, 2018; Vogel et al., 2021; Kladny et al., 2024).” to the previous paragraph, for a last paragraph focused on the aims and goals of this study, bringing together the topics brought in the other paragraphs of the introduction section.

We hope these modifications helped with the clarity and coherence in the last paragraph of the introduction.

**Lines 85-87:** *The authors need to provide more details on the reason for adopting this approach (Maximum NDVI) and its potential consequences for the dataset.*

Thank you for your comment, we apologize for the unclarities in that paragraph.

The AVHRR NDVI dataset is available as 10 days composites of median NDVI values. In other words, the NDVI composites are derived by computing the median NDVI value across 10 consecutive days. This strategy has been chosen due to the wide field of view of the AVHRR sensor (Asam et al., 2023). The maximum NDVI of several composite observation has been chosen in order to reduce the disturbing influences, such as clouds, snow, and aerosols that typically depress NDVI. We have reformulated line 80 to 87 to clarify this point (in addition to the modifications brought following reviewer#1 comments).

“As several platforms (NOAA-6 to -18 and the MetOp-series) may be available for a given composite, we selected the maximum NDVI value among the MED (median value composite) NDVI values to obtain a single time-series”

**Section 2.2:** *The authors should provide a more detailed justification for their decision to binarize the NDVI data rather than treating it as a continuous variable. Additionally, they should clarify the criteria used to determine the 80% threshold.*

Thank you for your remark. We followed the approach of Hermann et al. (2023), in which NDVI is binarized to focus specifically on explaining extreme low NDVI events, rather than modeling the full range of NDVI variability. This allows us to better identify conditions associated with potential forest browning or damage. The choice of the 80% threshold reflects a balance between capturing sufficiently extreme events and ensuring a large enough sample size for robust statistical analysis.



We added the following sentence in section 2.2:

“We follow the approach of Hermann et al. (2023), in which NDVI is binarized to focus specifically on explaining extreme low NDVI events.”

**Lines 110-115:** *The wording of this paragraph seems more appropriate for a discussion section than a methodological description.*

Thank you, we moved this part to the discussion session, in the paragraph “Dataset limitations”.

**Section 2.3 (lines 122-124):** *How were these layers with varying spatial resolutions harmonized for subsequent analyses?*

For clarity, we added the sentence “More specifically, each 0.1° gridpoint in a 0.5° gridbox was associated to the same precipitation and surface latent heat-flux per day.”

This choice was motivated by the fact that the downscaling from ERA5 to ERA5land for these two variables is statistical, in other words using ERA5land instead of ERA5 does not bring dynamical information.

**Section 2.4:** *While I understand the rationale behind this methodological step, I question whether 41 data points can support a comprehensive assessment of an area as vast as Europe.*

Thank you for your remark. If Europe were considered as a single model, where each grid point had the same relationship between predictors and low NDVI events, we would have enough data points for a robust statistical analysis. However, each small region in Europe has its own specific characteristics. Predictors can play different roles, or no role at all, depending on the region. Therefore, we need a different model for each smaller area. We consider a 0.5° region small enough to have a similar link between predictors and predictands, and we use the 0.1° resolution to ensure we have enough data points by stacking the 0.1° time series (see section 2.4 Aggregating data for longer time series, and Fig. C1 in the appendix).

**Section 4.2.1:** *The criteria for choosing the two grid points and the objective of this type of approach/analysis are unclear. The description of all analyses and their objectives and predictions must be detailed in the methods section.*

Thank you for your comment to improve the clarity of the manuscript. The important output of the RF models, namely the predictor’s importance and the partial dependance, are introduced in the method section 3.1. The point of section 4.2.1. is to illustrate this output for one gridpoint at a time, before analysing the output for all the gridpoints at the same time ( in Europe or for a given forest type, section 4.2.2.)

For a clearer introduction of section 4.2.1, we changed the “as an introduction” to “as an illustration” and added more details to this introductory sentence:

“As an illustration of the RF model output analyzed in this study, we display the mean decrease accuracy and the partial dependence plots for two GPs in Europe (their location is displayed in red in Fig. D1): one in a broad-leaved forest in France and one in a coniferous forest in Sweden.”

**Figures 2 and 4:** *Both figures are important results but challenging to interpret, particularly Figure 4. The authors should provide a more detailed explanation in the text to clarify the meaning of these results and enhance readability.*

Thank you for your remark. We added “The mean decrease accuracy and partial dependence plot are introduced in Sec. 3.1” in the caption of figure 2 (now Fig3) and “The partial dependence plot is introduced in Sec. 3.1” in the caption of Fig4 (now Fig5).

In addition, both in 4.2.1 *Example for two grid points* and 4.2.2 *Results over Europe*, we improved the interpretation of the partial dependence plot figures, providing both an interpretation in terms of probability of damage and hydro-meteorological conditions.

For example: “Negative anomalies of soil moisture (1-7cm) in JJA the year before are largely associated with an increased probability of forest damage (see the first partial dependence plot, Fig. 3, grid point (A)), indicating that dry conditions in the preceding summer are adverse for the forest at this GP.”

Other example: “Except for one GP, the ensemble of partial dependence plots shows two clear plateaus, separating negative and positive temperature anomalies, indicating that above-average June temperatures are associated with an increased probability of forest damage (Fig. 5 (A), first panel).”

We hope that the modified version of the text helps the understanding of the figures.

**Figure B1:** *Why are the analyzed pixels omitted (i.e., not colored green or orange) on some maps, as in much of Scandinavia in 1986?*

Thank you for your question. This is due to a low NDVI quality data for these years. We added more information in section 2.2:

“Due to a low NDVI quality in 1986 and in 1988, many GPs (or even all GPs for 1988) in Europe had to be discarded, implying large regions with missing data in 1986 and the absence of 1988 in Fig. B1.”

**Figure legends and resolution:** *The authors should provide more detailed figure captions, avoiding including acronyms without a complete meaning description. Additionally, some figures have low resolution or could be better formatted and expanded to enhance clarity and facilitate reader interpretation*

Thank you very much for these helpful remarks. We added more details to the figure captions, including references to relevant sections and meanings of acronyms. We improved the size and resolution of figures 4 (now 5) and B1.

#### **Technical Corrections:**

Thank you very much for your time and attention to these specific points. We implemented the changes as suggested.

**Line 15:** “202 millions hectares” Should be singular “202 million hectares”.

**Line 15:** “32 %” Remove the space between the number and the percent sign.

**Line 17:** “for a high biodiversity” remove the article “for high biodiversity”.

**Line 18:** Add a comma between “spiritual” and “and”.

**Line 20:** “have significant consequences on the economy and society” The preposition “on” should be replaced with “for”.

**Line 21:** The repetition of "and" is redundant and disrupts the sentence flow.

**Lines 25-26:** I suggest changing "photosynthesis capacity" to "photosynthetic capacity".

**Lines 26-27:** Avoid repeating terms in the same phrase. Here, "dry year" could be replaced by "one single event" (or another term) without changing the meaning of the sentence. The comma separating "dry year" and "due to" could also be removed.

**Line 32:** I think "advance" can be replaced by "advanced".

**Line 39:** "forest-fire" I think no hyphen needed "forest fire".

**Line 61:** "vegetation greenness, widely used" remove the comma "vegetation greenness widely used" or "vegetation greenness that is widely used".

**Line 136:** "We therefore artificially" change to "Therefore, we artificially...". The same observation is valid for other parts of the text, such as in lines 187 and 292.

**Line 187:** "wheat yield, in their study" remove the comma.