Authors' Response to Reviews of

Using statistical models to depict the response of multi-time scales drought to forest cover change across climate zones

Yan Li, Henning W. Rust, and Bo Huang

Hydrology and Earth System Sciences (HESS)

RC: Reviewers' Comment   AR: Authors' Response

RC: The authors use linear models to explore the influence of forest cover, temperature, and precipitation on the drought indices in various climate zones. The study's motivation and goal are exciting for the community. The exploratory data analysis used in this study is robust and could be interpreted very well. However, I have a few major comments, which shall be clarified/discussed further:

AR: Thanks for the appreciation of our work and the useful feedbacks to improve the quality of the manuscript.

RC: How do you isolate the local effects of forest cover and drought from the global drivers and large- scale atmospheric patterns? For example, increase/decrease in precipitation, anthropogenic global warming, jet-stream shift, ITCZ, etc.? The tree growth dependency on T and P depends on the biomes (Boreal forest, Temperate seasonal forest, etc.). Each tree has its characteristics.

AR: Isolating the local effects of forest cover and drought from global drivers and large-scale atmospheric patterns is a challenging task. It is difficult to completely separate the effects of global drivers, such as anthropogenic global warming and jet-stream shifts, from local effects. We employed several statistical models in this study to explore the potential impact of forest cover change on drought across different climate zones. However, it is important to note that statistical models can only provide an estimation of the true relationship between forest cover change and drought change. The coefficient of determination (Adjusted R2) of the statistical models varied between 0.23 to 0.97 (as shown in Table 2), indicating that these models can only partially explain the relationship between forest cover change and drought change. In future studies, it is crucial to consider multiple factors and utilize a variety of methods, such as statistical models, remote sensing, and ground-based measurements, to obtain a more comprehensive understanding of the dynamics between forest cover and drought.

Regarding tree growth, it is true that tree growth can depend on temperature and precipitation, and the specific characteristics of each tree species. However, in this study, we focus on the overall effect of forest cover change on drought, rather than the specific effect of each tree species. We use a satellite-based forest cover dataset, which provides information on the extent of forest cover across different regions, rather than information on the specific tree species present in each region.

RC: It should be described why the authors used linear models for their analysis.

AR: We include in the text "We use linear models because of their great flexibility, versatility and robustness. They are characterized by linearity in parameters to estimate; relations between predictand and variables in the predictor can still be formulated in a non-linear way. Furthermore, they easily allow to describe joint effects of different variables (temperature and forest cover) on the predictand (interactions), a feature made extensive use of in this study." in Line 192-195.

RC: One suggestion which might be considered to add value to the results: Using the linear regression model is an excellent approach to analysing the interactions between the variables and features. However, as mentioned by the authors, the interplay among precipitation, temperature, soil, land cover and drought might be complex and non-linear. Authors could add some complexity to the model by using decision-tree-based models already implemented in R and comparing the results with the linear model. On the other hand, simple/shallow decision tree models are also interpretable.

AR: Indeed, a regression tree model could be an interesting extension to our approach. However, we did not see any indication for the need of more complex models based on explorative line plots analysis made prior to building the regression models. Furthermore, a further splitting of data — which a tree-based model brings analog — would probably lead to less robust results. As we also assume relations to be at least smooth (if not linear), we suggest using additive models for further studies. And we add the text "A generalization to additive models (not necessary linear) might reveal more subtle effects. However, an initial explorative analysis with line plots did not suggest these based on the data used here." in Line 422-423.

Other comments:

RC: Lines 1-5: What do you mean by forest cover change? Do you mean human-made changes or natural changes?

AR: In this study, we are not making a distinction between whether the changes in forest cover are a result of human activities or natural processes. We are solely examining the impact of changes in forest area on drought.

RC: Line 6: Hard to understand: ""to explore the changes in forest fraction and drought from 1992–2018."". Do you mean to find a kind of relationship between those two? Or exploring them separately? And why those 27 years?

AR: Our aim is to investigate the impact of changes in forest area on drought across different timescales and climate zones. To achieve this, we begin by examining the average changes in forest cover and drought indices across four distinct climate zones. Subsequently, we construct a series of statistical models to identify the relationship between changes in forest area and drought. We changed the word "changes" Line 6 to "relationship". For our analysis, we utilize the ESA CCI land cover data, which was available for the period of 1992-2018 at the time of our study.

RC: Line 7: which various factors? Please clarify! Are they natural factors or management factors, etc.

AR: In this study, the term "various factors" refers to forest cover change, precipitation, and temperature, which we clarified in Line 8-9. It is important to note that we only considered forest area changes based on the ESA CCI map and did not distinguish between changes caused by natural factors or management practices.

RC: Lines 8-9: Is precipitation the dominant one among the two variables? Please mention!

AR: The impact of forest cover change, precipitation, and temperature on the change of droughts varies across different regions. Precipitation appears to be the dominant factor affecting the change of droughts in the equatorial, temperate, and snow regions, whereas in the arid region, temperature is the dominant factor among the three. It is clarified in the manuscript (Line 8-9).

RC: Lines 9-10: It needs to be clarified: You mention precipitation and temperature (which describe the climate state), then forest cover and finally, short and long-term drought. The reader needs to catch up on the clear goal. Please clarify which relationships or driving effects you will explore in this manuscript.
Some chains like: T, P => forest cover => drought?

AR: In this study, we aimed to investigate the influence of forest cover change and meteorological factors, such as precipitation and temperature, on droughts at different time scales. In Section 4.1, we analyzed the effects of forest cover change and meteorological factors on droughts based on Analysis of Variance mentioned in Section 3.2. In Section 4.2, we focused on how meteorological factors influence the impact of forest cover change on droughts. Finally, in Section 4.3, we examined the effects of meteorological factors on droughts under extreme values of forest cover area, specifically the maximum and minimum values. The expression has been clarified in the revised manuscript (Line 91-95).

RC: Line 30: "–500 million hectares up to +1000 million hectares" what do -500 million hectares mean? And all the SSPs show the same trend, or do they differ from each other?

AR: Popp et al. (2017) reported a systematic understanding of the Shared Socio-Economic Pathways (SSPs) and their potential impacts on land-use changes, agriculture, food security, greenhouse gas emissions, and related issues. The authors employed five Integrated Assessment Models, each equipped with distinctive land-use modules, to convert the SSP narratives into quantitative projections. We give details to explain the future forest changes under SSP scenarios. (Line 31-34)

RC: Lines 35-45: maybe you could also mention that extensive forests like the Amazon are the sink of CO_2 and are predicted to become a source of CO_2 under the recent trend of climate change we are following:

Boulton, C.A., Lenton, T.M. & Boers, N. Pronounced loss of Amazon rainforest resilience since the early 2000s. Nat. Clim. Chang. 12, 271–278 (2022). https://doi.org/10.1038/s41558-022-01287-8

AR: We add this reference in the new version (Line 24-26).

RC: Line 57-58: "in this region," which region mention again.

AR: Clarified. The region is temperate region. (Line 61)

RC: Line 59-60: Is drought a condition or a phenomenon? Clarify? There are also many definitions for drought, like meteorological, agricultural, etc.... Please clarify how you define the drought. Which index do you use? Is it based on temperature and precipitation, or other variables, like soil moisture, evaporation, etc. are, involved?

AR: Our primary focus is on the changes in drought conditions. It should be noted that there are various definitions of drought, including meteorological drought, agricultural drought, hydrological drought, and others. For our study, we use two specific indices to measure drought: the Standardized Precipitation Evapotranspiration Index (SPEI) and the self-calibrated Palmer Drought Severity Index (scPDSI). The SPEI is based on the difference between precipitation and potential evapotranspiration, assuming a Gaussian distribution. And the scPDSI is based on a two-layer soil model, incorporating not only meteorological variables but also soil condition factors, making it an appropriate index for measuring hydrological drought. The information is give in details in Section 2.2. And we also changed the word "phenomenon" to "condition" in Line 63.

RC: Line 62-63: What do you mean by "forest structure and carbon content"? Please clarify.

AR: Changes in drought patterns can affect the forest structure by altering the growth and survival of different species, leading to changes in biodiversity and ecosystem functioning. In addition, changes in forest structure can affect carbon storage by altering the distribution and types of vegetation that store carbon. Meanwhile, drought can lead to increased mortality and decreased growth of trees, which can reduce the amount of carbon stored in the forest ecosystem. We added the explanation in the new manuscript (Line 65-67).

RC: Line 65: How much increase in the frequency and intensity? Is it significant? With respect to which period? Please describe in more detail!

AR: The scope of this study was limited to the analysis of the changes in drought conditions (dry or wet), and did not include an investigation of the frequency change and intensity. Further research can explore the impact of forest cover area change on the frequency and intensity of drought across different climate zones. And we added the explanation in detail in the new manuscript (Line 68-75).

RC: Lines 79-80: The word ""change"" is used frequently.

AR: Revised.

RC: Lines 87-88: Why didn't you use the newer version of the data with a higher resolution or cite this study:

Beck, H., Zimmermann, N., McVicar, T. et al. Present and future Köppen-Geiger climate classification maps at 1-km resolution. Sci Data 5, 180214 (2018). https://doi.org/10.1038/sdata.2018.214

AR: Several versions of the Köppen-Geiger climate classification maps have been generated over time. In this study, we utilized the main climate zones from the Köppen-Geiger climate classification system. We compared the main climate zones in the latest version and the version we used, and found no significant differences. Therefore, we decided to keep the current climate zone classification and provided an explanation for it in the revised manuscript (Line 105-108).

RC: Figure 1: Please insert the number of grid points belonging to each main climate classification.

AR: We added the grid points for the five main climate zones in the new version (Line 109-110).

RC: Line 116: Does the potential evapotranspiration data have a reference?

Please include a table with the characteristics of the data used in this study to have a better overview. For example, it is boring to know when and where you downloaded each dataset. A table would be enough, which describes all the datasets. And please include the citation of each dataset in the table.

AR: The potential evapotranspiration in this study was calculated using the FAO-56 Penman-Monteith method, and the necessary datasets for the calculation were obtained from the CRU TS3.24.01 dataset. More details regarding the data used in this study can be found in the supplementary document, specifically in Table A1.

RC: Figure.2. Given that the scPDSI values between -1 and 1 are considered normal, how significant are the annual trends shown in the drought indices in Fig.2?

AR: In Fig. 2, we present the scPDSI values that have been averaged for each region and year. So these values are more towards the average. According to Dai (2013), the scPDSI values ranging from -0.5 to -1.0, -1.0 to -2.0, -2.0 to -3.0, and -3.0 to -4.0 correspond to dry spell, mild drought, moderate drought, and severe drought, respectively. By analyzing the changes in scPDSI values, we can infer the trend of dryness or wetness.

RC: Line 172: Why not consider the precipitation sum (yearsum) instead of mean (yearmean)?

AR: Since we use standardized precipitation, temperature, and forest cover area to build our linear models, it does not matter whether we use precipitation sum (yearsum) or mean (yearmean) in our analysis.

RC: Line 175: What do you mean by complex? Clarify!

AR: Based on the observations from Fig. 3 P, Q, R, S, T, we cannot establish a clear relationship between the drought indices and precipitation or temperature in the snow region. This implies that precipitation or temperature may not be the dominant factors in this region. Other factors such as their interaction or other environmental variables may play a more important role in driving the changes in drought conditions in the region (Line 183-185).

RC: Lines 200-215: How about the problem of collinearity? There might be correlations between the forest cover change and P or T. How do you consider this? A correlation matrix might show the collinearity between the predictor variables, or the Variance inflation factors (VIF) method might help. The other concern is how many grid points you achieve for each climate zone. How big is the training dataset for each climate zone? I assume you have a more extensive training dataset for the temperate than the equatorial zone. How about the seasonality? You have an arid zone in both the North and South hemisphere. Averaging over all those grid points might mix seasons. Could one include the latitude as an extra feature in the lm model?

AR: In order to ensure the accuracy of our linear models, we also examined the collinearity among the independent variables in Equation 5 using the Variance Inflation Factor (VIF) function in R. All VIFs were found to be less than 5, indicating moderate correlations between the variables but not severe enough to require attention (added in Line 216-218).

It should be noted that for the equatorial region, there were 11,030 points; for the arid region, there were 15,673 points; for the temperate region, there were 9,587 points; and for the snow region, there were 20,734 points. In each linear model, the data length was 25 or 26 years since we only analyzed the annual change of various variables and filtered out the seasonal influence. As shown in Fig. 1, the climate classification is also latitude-related, so if we use different linear models for different climate regions, we do not need to add latitude as a variable in model building.

RC: Equation 6: Is "i" indicating the observation over different grid points and times? Or do you average the gridpoints of each climate zone at each time, and "i" is just the time? Do you train for each grid point a separate linear model? Do you train one linear model for each climate zone? Please clarify in more detail....

AR: For each region and variable, we averaged the grid points, resulting in a time series of 25 or 26 years with "i" representing the number of observation years. We utilized forest cover, precipitation, temperature, and drought indices (on different time scales) datasets to train a linear model for each region. Therefore, the study resulted in 20 linear models (4 climate regions X 5 drought indices: scPDSI and SPEIs). In Section 4.1, we analyzed the contribution of different factors to drought across different time scales and climate regions. In Sections 4.2 and 4.3, we conducted sensitive experiments based on the linear models to assess the interaction of forest cover and meteorological factors to droughts across different time scales and climate regions. Relevant information regarding this is presented in Line 130-133, Line 144-145, and Line 155-157.

RC: Line 261: Must be moved to data and methods.

AR: The sentence has been deleted in the section.

RC: Line 262: You mentioned the regions before.

AR: Deleted.

RC: Lines 283-284: How does the time deviation of forest cover look like in equatorial regions? There may be some temporal changes in tropical forest cover. This is because the trees receive enough energy (T) and moisture (P) throughout the year. Have you removed the seasonal cycle from the "lm" features, i.e., T, P and forest?

AR: The variable (forest cover) we used in the models are annual value, so it has filtered out the seasonal influences.

RC: Lines 295-297: It is a strong conclusion based on a single linear statistical model. I would be cautious about concluding solid results on this.

AR: Yes. The conclusions presented in this study are based on a set of linear statistical models with a limited number of predictors. It is important to note that this approach has its limitations and future studies can investigate more complex models to explore other potential effects of forest cover change on drought. However, we chose to use linear models because they are easy to implement, interpret and efficient to train. We would like to emphasize that this study is not perfect, but it can still be meaningful if it helps people to better understand the complex relationship between land use and climate change. We add some sentences to discuss it. (Line 419-423).

RC: Figure 6: I see green and yellow colours and not blue and red lines. Using symbols instead of colours could help readers with colour blindness.

AR: Modifying the figure from colour lines to symbols may not be feasible. However, we clarify the colours used in the figure by explicitly stating that dark green and dark yellow represent the effect of precipitation and temperature, respectively.

RC: Lines 364-366: Given that the trees' species might change in the snow and equatorial regions, how do you isolate those impacts?

AR: Our current study only focuses on the forest cover area change, and we are not explicitly accounting for the potential impacts of changes in tree species on drought conditions. However, changes in tree species could potentially affect the water cycle and ecosystem functioning in ways that could indirectly impact drought conditions. In order to isolate these impacts, more detailed studies would be needed that incorporate information on the specific tree species and their water use characteristics in different regions. This information could then be used to refine models and better understand the complex relationships between forest cover, tree species, and drought conditions in different ecosystems. It can be done in the future research.

RC: Line 368-369: I am unsure if this is the correct English: "The colour change ... should be vertical". Please re-frame.

AR: Clarified (Line 379-380).

RC: Line 382-383: Please mention that your conclusion is valid only under the assumptions you use here. There might be other models more accurate than your linear model.

AR: We add some sentences to clarify the limitation and possible application of our conclusion. (Line 419-423)

RC: Line 414: You have to spell out CMIP.

AR: The full name of CMIP is given in the main text (Line 429-430).

Reference:

Dai, A, 2013: Increasing drought under global warming in observations and models. *Nature Climate Change*, **3**, 52.

Popp, A., and Coauthors, 2017: Land-use futures in the shared socio-economic pathways. *Global Environmental Change*, **42,** 331-345.