nature portfolio

Peer Review File

The distribution and drivers of tree cover in savannas and forests across India

Corresponding Author: Dr Trisha Gopalakrishna

Version 0:

Decision Letter:

** Please ensure you delete the link to your author home page in this e-mail if you wish to forward it to your coauthors **

Dear Dr Gopalakrishna,

Your manuscript titled "The distribution and drivers of tree cover in India" has now been seen by 2 reviewers, whose comments are appended below. You will see that they find your work of some potential interest. However, they have raised quite substantial concerns that must be addressed. In light of these comments, we cannot accept the manuscript for publication, but would be interested in considering a revised version that fully addresses these serious concerns.

Should additional work allow you to address these criticisms and meet our editorial thresholds as outlined below, we would be happy to look at a substantially revised manuscript. If you choose to take up this option, please either highlight all changes in the manuscript text file, or provide a list of the changes to the manuscript with your responses to the reviewers.

In particular, we ask that you:

** Provide a clear and compelling support for using the MOD44B tree cover dataset to quantify savanna tree cover, by (1) using independent tree cover measurements, and/or (2) showing that potential flaws of the dataset do not impact your findings

** Provide a data-driven approach to define the thresholds used to separate tree cover zones into different bins by using the maximum climate water deficit (MCWD), or rigorously justify the choice of the visual approach

Please bear in mind that we will be reluctant to approach the reviewers again in the absence of substantial revisions.

If the revision process takes significantly longer than three months, we will be happy to reconsider your paper at a later date, as long as nothing similar has been accepted for publication at Communications Earth & Environment or published elsewhere in the meantime.

We are committed to providing a fair and constructive peer-review process. Please do not hesitate to contact us if you wish to discuss the revision in more detail.

Please use the following link to submit your revised manuscript, point-by-point response to the reviewers' comments with a list of your changes to the manuscript text (which should be in a separate document to any cover letter), a tracked-changes version of the manuscript (as a PDF file) and any completed checklist:

Link Redacted

** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first **

Please do not hesitate to contact us if you have any questions or would like to discuss the required revisions further. Thank you for the opportunity to review your work.

Best regards,

Erika Buscardo, PhD

Editorial Board Member
Communications Earth & Environment
orcid.org/0000-0003-4029-8465

Alienor Lavergne, PhD
Associate Editor
Communications Earth & Environment

EDITORIAL POLICIES AND FORMAT

If you decide to resubmit your paper, please ensure that your manuscript complies with our editorial policies and complete and upload the checklist below as a Related Manuscript file type with the revised article:

Editorial Policy Policy requirements (Download the link to your computer as a PDF.)

For your information, you can find some guidance regarding format requirements summarized on the following checklist: (https://www.nature.com/documents/commsj-phys-style-formatting-checklist-article.pdf) and formatting guide (https://www.nature.com/documents/commsj-phys-style-formatting-guide-accept.pdf).

REVIEWER COMMENTS:

Reviewer #1 (Remarks to the Author):

The study is interesting and relevant. The weak part of the paper are the datasets used. In particular the tree cover data MOD44B is not well suited to study savanna tree cover. The authors state that it has been used in this context, but exactly those studies have been highly criticized for the use of this dataset, some studies even state that the results are artifacts from the data. I also spoke with the developers of the MOD44B tree cover dataset who confirmed that it should not be used for savannas. There seems to be a clear binning of the tree cover in Figure 1.

https://onlinelibrary.wiley.com/doi/10.1111/geb.12592

https://onlinelibrary.wiley.com/doi/10.1111/geb.12122

Now I do not want to say that the findings of this study are wrong and the results useless due to flawed data products. However, from my point of view, if the authors wish to publish in this relatively high-impact journal, I would ask for proof that MOD44B can be used in the context of this work, ideally by using independent tree cover measurements, or at least by showing that potential flaws of the dataset do not impact the findings.

-Tree planting in savannas is taken as a framing of this study, but one may distinguish between tree planting that is done for economic purpose, and tree planting for restoration purpose.

-The result section could shortly mention which products are used, in particular tree cover and MCDW. The choice of the datasets is crucial. Generally the results section should help the reader more to understand what has been done without having to scroll to the methods every second sentence. It would also be great to have the figures at the places they are mentioned, so we do not have to permanently scroll up and down.

-I looked at forests in India in Google Earth, and found many areas with closed canopy tree cover, which I would expect to be around 100% cover. Tree cover % in Figure 1 stops however at 80%, is it because of the pixel size of MOD44B, which is actually never mentioned in the paper? It would be crucial to know the spatial resolution of the analyses, also for the figures.

-Figure 2: the tree covered areas of “a” and “b” do not match, is it because areas with less than 1300 MCWD are shown in “a” but not “b”? Was not so clear to me.

-Line 321: “For example, our results of delineation of intermediate zones of tree cover containing savanna vegetation and the low tree cover zone can inform the design of ‘no-go’ areas for tree planting schemes”: Would it be interesting to show a map on this?

Reviewer #2 (Remarks to the Author):

In this study, the relation between tree cover and multiple drivers is studied for India. The drivers include variables such as precipitation, seasonality, fire, herbivory or topography. The study shows that maximum tree cover is constrained by rainfall

seasonality but that observed tree cover can deviate substantially from that maximum due to different drivers. The results are relevant for management and conservation as many Indian savannas have been wrongly classified as degraded forest and identified as suitable for tree planting. The study highlights suitable areas for tree planting. The results are also a relevant contribution to our knowledge of Indian savannas.

Overall, the study is well-conducted and the manuscript is well written. Methods are described in detail. Nonetheless, I have one major comment and a few minor points.

My main concern is related to the separation into different tree cover zones using MCWD. If I understood correctly, this was done only on visual inspection. I would appreciate a more data-driven approach (for example by cluster analyses or piecewise regressions) or at least a better explanation and justification of the thresholds. When looking at Fig 1, there is a MCWD range with few data points at around 500mm that might be better for the separation of the high tree cover zone? Further, tree cover starts saturating at around 750mm, ie a slightly lower value than currently selected for the moderate/mixed zone separation. I think that such changes of the threshold values would not change the main results. Related to the zones, it was not entirely clear to me, why one of the zones is called mixed, does this refer to a mixture of savannas and forest? Or because tree cover varies over a large range (ie is mixed)? But such a variation is also the case in the high tree cover zone. Is it possible to convert these thresholds into MAP units and also highlight the zones in Fig S1?

Minor comments:

L. 99: delete "And" in "And woody biomass.."?

L. 108: maybe "in other regions" instead of "elsewhere"

L 111: maybe "site data" instead of "local information" (also elsewhere)?

L. 114-123: I suggest removing references to figures in this introductory paragraph.

L. 166: "no such ceiling in the mixed tree cover zone". I disagree with this, there is also a clear ceiling at around 70-75% and a linear increase until around 750mm MCWD. Tree cover in the medium zone also increases linearly, why is this denoted a clear ceiling? Given the saturation of tree cover in the mixed zone, I would rather argue that the mixed zone has a ceiling but not the medium zone. Further, such a visual assessment of a ceiling is also related to the thresholds used to define different zones (see first comment).

L. 176: values ... indicate

L. 178: values ... mean

L. 194 and onwards: I suggest to mention that the results refer to the partial effects of the GAMs. Statements such as "Shortfall decreases linearly with..." may suggest that that these are results from scatter plots or models with only one explanatory variable.

L. 211: "we demonstrate that there are four distinct zones" these zones were defined visually but not obtained by analytical methods. I think it is not clear from this analysis, that the distinction used in the analyses is the best choice.

L. 250: "climate change variability and climate change driven...", climate change was not studied, so I suggest to say "climate variability and climate driven...".

The ms highlights the importance of the study for tree planting and no-go areas for planting. The discussion could include explicit recommendation regarding such areas.

L. 361: please add reference for land cover map as well as the spatial resolution of the different data sets.

L. 366: given that 300 points were sampled per MAP bin: does multimodality of tree cover vs MAP occur in the full data set? I also suggest adding the information to Fig 1 and Fig S1 that they show only a subset of the points.

L. 373: "long temporal resolution" does that mean high temporal resolution (ie number of time points in the study period) or the long study period (ie number of years)?

L 382: I suggest adding a definition of MCWD

L. 410: how many data points were available from that data set?

Fig S2: precipitation ranges overlap in the caption high tree cover 0 to 600 and mixed from 580 to 930. Further 1300-1600 is mentioned for panel (a), this is the low tree cover zone. I suggest modifying the colors and not use grey for unburned points in the histogram. Grey is used as background in maps, but the same color does not represent the same points/information in the maps and the histograms.

Fig S6: in the PCA plot, names of herbivores overlap and cannot be read.

Author Rebuttal letter:

Response to Reviewersâ Comments

We would like to emphasize the importance of using site data of endemic savanna species (forbs, graminoids and woody components of the understorey vegetation) in this study. The use of data published in Nerleker et al.1 in our study is important because it directly addresses two of the main comments provided by the reviewers.

First, the MOD44B Vegetation Continuous Fields data product used to characterize tree cover in this study and various other studies2â4 has varying certainty in delineating the savanna biome as noted by Reviewer 1 in Comment 1. A savanna biome is characterized by a continuous C4 herbaceous layer and discontinuous C3 layer of savanna trees as we have explained in the first sentence of the manuscript (lines 44- 46). To the best of our knowledge, no optical satellite product of vegetation neither provides information about the understorey layer nor the type of tree species i.e. forest or savanna tree species, two components needed to describe the savanna biome across the study area. Hence, the use of GPS locations of savanna vegetation (graminoids, forbs and woody components of the understorey vegetation) overcomes the challenge of using a satellite derived tree cover product, as the GPS locations of the site data validates the presence of a savanna biome.

Second, in the initial manuscript submission, we delineated the zones of tree cover by visually inspecting both- the distribution of tree cover against water stress and the tree cover at the GPS locations of the site data of savanna vegetation. Though not data driven, our initial method capitalized on the site data to delineate the zones which might be savanna. In response to Comment 7 by Reviewer 2, we have now used a non-hierarchical clustering algorithm, specifically partitioning around medoids to delineate the different zones. This has yielded similar results but with slightly different thresholds of MCWD of the four zones (see response to Comment 7 for all details). However, we would like to emphasize that the distribution of sites containing savanna vegetation provides additional ecological meaning to the zones.

Reviewer 1
1. The study is interesting and relevant. The weak part of the paper are the datasets used. In particular the tree cover data MOD44B is not well suited to study savanna tree cover. The authors state that it has been used in this context, but exactly those studies have been highly criticized for the use of this dataset, some studies even state that the results are artifacts from the data. I also spoke with the developers of the MOD44B tree cover dataset who confirmed that it should not be used for savannas. There seems to be a clear binning of the tree cover in Figure

https://onlinelibrary.wiley.com/doi/10.1111/geb.12592

https://onlinelibrary.wiley.com/doi/10.1111/geb.12122

Now I do not want to say that the findings of this study are wrong and the results useless due to flawed data products. However, from my point of view, if the authors wish to publish in this relatively high-impact journal, I would ask for proof that MOD44B can be used in the context of this work, ideally by using independent tree cover measurements, or at least by showing that potential flaws of the dataset do not impact the findings.
We acknowledge the relevant point about the utility of the MOD44B product in delineating savanna and forest biomes made by the reviewer. We provide a justification for the use of this product in our study and conclude with the changes we have made in the manuscript.
First, we would like to emphasize that our study does not exclusively examine tree cover in the savanna biome as stated by the reviewer. Instead in our study, we characterize tree cover across all biomes i.e. desert, forests, savannas and grasslands. To support this point, we validate our results of the four different zones against the KÃ¶ppen biome classification which details the potential type of biome based on climatic conditions (lines 142 - 145). Second, we ascertain the existence of the savanna biome in the medium and mixed tree cover zones not only by examining the distribution of tree cover, which is the method employed by many studies that have been criticized as stated by the reviewer. We establish that the medium and mixed tree cover zones are savannas because these zones contain site location data of savanna vegetation (orange diamonds in Fig 1, lines 134 - 135, lines 433 - 436 and in the Supplementary Methods). Hence, in response to the reviewerâs suggestion of validation with independent tree cover measurements, we have already gone one step further by validating the zones with actual savanna vegetation (graminoids, forbs and woody components of the herbaceous layer as per Nerleker et al.1). Even if we were to use independent tree cover measurements, we would have to identify the tree species to be that of a forest or savanna. The distinction of forest and savanna tree species is crucial because savanna trees in India have similar physiognomies to forest trees5 which means that they might have similar tree cover as we have explained in the Introduction (lines 58 - 59), in the Discussion (lines 240 - 243) and in the Methods (lines 431 - 433). Third, considering the above two points, we have written the Results section to focus on tree cover and not biomes directly. We have done so by naming the four zones in terms of tree cover and not as biomes.

We acknowledge the uncertainties associated with this product especially while examining the distribution of biomes. Hence, in lines 145 - 147, we have stated that the use of the site location of savanna vegetation helped us overcome the challenge of uncertainty of the MOD44B VCF product in delineating forests and savannas. Furthermore, we discuss this uncertainty as a caveat in our analyses, suggesting that independent tree cover and tree species information can be used to further validate the medium and mixed tree cover zones to be bistable stable states of forests and savannas in India (lines 247 - 251).

2. Tree planting in savannas is taken as a framing of this study, but one may distinguish between tree planting that is done for economic purpose, and tree planting for restoration purpose.
We agree with the reviewer about this distinction and now have included this point about tree-planting for economic and restoration purposes (lines 54 - 55).

3. The result section could shortly mention which products are used, in particular tree cover and MCDW. The choice of the datasets is crucial. Generally the results section should help the reader more to understand what has been done without having to scroll to the methods every second sentence. It would also be great to have the figures at the places they are mentioned, so we do not have to permanently scroll up and down.
We have now included brief details of the geospatial products for tree cover, the climate variables and the associated spatial resolution in the results section (lines 125 - 129). Furthermore, we have included brief information about the methods throughout the results section. First, we have now included that we have used a non-hierarchical clustering analysis in combination with site data of savanna vegetation to delineate the four zones (in response to Reviewer 2âs comment 7) (lines 134 â 136, lines 430 - 439). Second, we have included that a 95th percentile smooth additive quantile regression model was used to test the relationship between percent tree cover and MCWD and to estimate climatic maximum potential tree cover (lines 172 - 177). Lastly, we state that the partial effects obtained from the generalized additive models of shortfall against the drivers have been analysed to understand the drivers of potential tree cover shortfall (lines 202).

We have followed the formatting guidelines of the journal which require us to include figures and methods towards the end of the manuscript.

4. I looked at forests in India in Google Earth, and found many areas with closed canopy tree cover, which I would expect to be around 100% cover. Tree cover % in Figure 1 stops however at 80%, is it because of the pixel size of MOD44B, which is actually never mentioned in the paper? It would be crucial to know the spatial resolution of the analyses, also for the figures.
Firstly, the maximum annual tree cover in the MOD44B VCF product across the time series of 2000-2020 is 98% in northeastern India for the year 2014, with no areas having 100% tree cover. Secondly, in our study, we have reduced the time series information to extract mean annual tree cover (%). We did so to leverage the 20 years of annual tree cover data available in the MOD44B,

allowing us to capture variations of tree cover through time. Hence, our mean annual tree cover metric does not mean the maximum annual tree cover. For example, the pixel with maximum tree cover of 98% has mean annual tree cover of 82.45%. In our study, the maximum mean annual tree cover is 84.6% with the respective pixel having maximum tree cover of 88%, which explains the upper limit of ~85% shown in Fig 1. Lastly, the confusion is not related to the size of the pixel of the MOD44B product, as Fig 1 is based on a stratified sample of points across the study area (Fig S1) and not pixels.

We acknowledge the confusion expressed by the reviewer. Hence, we have explicitly mentioned that the percent tree cover and consequent metrics such as the potential tree cover shortfall are based on the mean annual tree cover across 2000-2020 in the study area (lines 125 - 126). Furthermore, we have included that the spatial resolution of the geospatial data used is 250m (lines 126 and in caption of Fig 2), while the remaining analyses has been completed based on a stratified sample of points (lines 129, 175 and 383).

5. Figure 2: the tree covered areas of âaâ and âbâ do not match, is it because areas with less than 1300 MCWD are shown in âaâ but not âbâ? Was not so clear to me.
We have now clarified the caption of Fig 2 to explicitly mention that all areas less than 1209 mm MCWD (based on the new non- hierarchical clustering analysis) have been excluded in the map of potential tree cover shortfall i.e. panel (b) to avoid arguments of tree planting in northwestern India, which is primarily desert. We have also clarified this point in the main text (lines 189 - 192).

6. Line 321: âFor example, our results of delineation of intermediate zones of tree cover containing savanna vegetation and the low tree cover zone can inform the design of âno-goâ areas for tree planting schemesâ: Would it be interesting to show a map on this?
We agree with the reviewer that a map of the âno- goâ areas might be useful to show to inform tree-planting strategies. However, the delineation of such areas is a complex issue because we need additional information such as how local people and communities currently use these areas and how would they want to use the areas in the future, when considering tree-planting6,7. Our study does not delve into such important socio-ecological factors. Hence, we think that such a map would extrapolate our results beyond what the results are intended to communicate.
Reviewer 2
7. My main concern is related to the separation into different tree cover zones using MCWD. If I understood correctly, this was done only on visual inspection. I would appreciate a more data-driven approach (for example by cluster analyses or piecewise regressions) or at least a better explanation and justification of the thresholds. When looking at Fig 1, there is a MCWD range with few data points at around 500mm that might be better for the separation of the high tree cover zone? Further, tree cover starts saturating at around 750mm, ie a slightly lower value than currently selected for the moderate/mixed zone separation. I think that such changes of the threshold values would not change the main results. Related to the zones, it was not entirely clear to me, why one of the zones is called mixed, does this refer to a mixture of savannas and forest? Or because tree cover varies over a large range (ie is mixed)? But such a variation is also the case in the high tree cover zone. Is it possible to convert these thresholds into MAP units and also highlight the zones in Fig S1?
Thank you for this valuable suggestion about a data-driven approach to define the thresholds. We have now used a non-hierarchical clustering analysis to delineate the thresholds of the different zones. Specifically, we have used the partitioning around medoids method as this clustering technique is able to handle outliers more efficiently than other clustering techniques8 (lines 134 and lines 424 - 429). We considered 8554 points in Fig 1. Using the percent tree cover and MCWD values of these points we determined that the number of clusters is 4. This is because the total within sum of squares error was the lowest with 4 clusters and because there was negligible drop in total within sum of squares errors subsequently (figure below which has been included as Fig S3). We then extracted the minimum and maximum values of MCWD of each of the 4 clusters to be the thresholds of the zones. The low tree cover zone spanned area below 1209 mm MCWD, the medium tree cover zone with savanna has MCWD range of 845 â 1209 mm, the mixed tree cover zone in the MCWD range of 486 - 845 mm and the high tree cover zone is above 486 mm of MCWD (lines 136 - 141). Furthermore, most of the savanna vegetation still spans the medium and mixed tree cover zones.

Secondly, as we are analysing percent tree cover, all zones are named such that the name characterizes the tree cover. Hence, the mixed tree cover zone is called so because tree cover varies across a large range. While this is true for the high tree cover zone to a degree, there is a significantly greater percent tree cover in the high tree cover zone shown by the high density of points above 50%. To make this point clear, we have now remade Fig 1 with the sampled points shown with a certain degree of transparency (black points), such that the high density of points in the high tree cover zone is clearly visible.

Lastly, we have now included a plot of percent tree cover against MAP with the four zones of tree cover coloured in four different colours as Fig S4 shown below.

Because of the new thresholds of the medium and mixed tree cover zones, the sampled points used in the analyses to understand the drivers of potential tree cover shortfall has reduced from 3106 to 2977. The deviance explained by the generalized additive model of the drivers against potential tree cover shortfall and the associated root mean square errors and explanatory powers of the testing and training data changed negligibly (lines 198 â 201). Of all the drivers, the partial effect of sand fraction on potential tree cover, though still significant, should be interpreted with caution as the confidence intervals are wide (see Fig 3). We have now included this cautionary statement in the Results (lines 203 â 204).

8. L. 99: delete "And" in "And woody biomass.."?
We have deleted the word âAndâ (line 99).

9. L. 108: maybe "in other regions" instead of "elsewhere"
We have replaced âelsewhereâ with âglobally and in other regionsâ (line 108).

10. L 111: maybe "site data" instead of "local information" (also elsewhere)?
We have replaced âlocal informationâ with âsite data of Indian savanna vegetationâ (lines 111 112).

11. L. 114-123: I suggest removing references to figures in this introductory paragraph.
We have deleted the references to figures as per the reviewerâs suggestion (lines 111 - 122).
12. L. 166: "no such ceiling in the mixed tree cover zone". I disagree with this, there is also a clear ceiling at around 70-75% and a linear increase until around 750mm MCWD. Tree cover in the medium zone also increases linearly, why is this denoted a clear ceiling? Given the saturation of tree cover in the mixed zone, I would rather argue that the mixed zone has a ceiling but not the medium zone. Further, such a visual assessment of a ceiling is also related to the thresholds used to define different zones (see first comment).
We clarify that by a clear ceiling, we mean that for a particular MCWD range there is a maximum possible percent tree cover that can be reached in each of the four zones. For example, in the medium tree cover zone, even though the percent tree cover seems to increase linearly with MCWD, majority of the points in this zone has percent tree cover within the linear trajectory. We agree with the reviewer that in the mixed tree cover zone, there is a clear ceiling of 70-75% until around 750 mm MCWD. Hence, we have now rephrased the sentence to âWe hypothesized a clear ceiling for the maximum percent tree cover across the four zones of tree cover, implying the influence of MCWD on the maximum current tree cover possible (Fig 1)â (lines 171 - 172).

Lastly, we have now used a data-driven approach, specifically non-hierarchical clustering analysis to delineate the four zones of tree cover. The low tree cover zone spanned area below 1209 mm MCWD, the medium tree cover zone with savanna has MCWD range of 845 - 1209 mm, the mixed tree cover zone in the MCWD range of 486 - 845 mm and the high tree cover zone is above 486 mm of MCWD (lines 136 - 141).

13. L. 176: values ... indicate
We have corrected the grammar as per the reviewerâs suggestion (line 184).

14. L. 178: values ... mean
We have corrected the grammar as per the reviewerâs suggestion (line 186).

15. L. 194 and onwards: I suggest to mention that the results refer to the partial effects of the GAMs. Statements such as "Shortfall decreases linearly with..." may suggest that that these are results from scatter plots or models with only one explanatory variable.
Thank you for the suggestion. Now, we have stated that partial effects of the drivers on shortfall were analysed using generalized additive models (line 202).

16. L. 211: "we demonstrate that there are four distinct zones" these zones were defined visually but not obtained by analytical methods. I think it is not clear from this analysis, that the distinction used in the analyses is the best choice.
We have now used a non-hierarchical clustering analysis to delineate the thresholds of the different zones. Specifically, we have used the partitioning around medoids method as this clustering technique is able to handle outliers more efficiently than other clustering techniques8 (lines 134 and lines 430 - 435). We considered 8554 points in Fig 1. Using the percent tree cover and MCWD values of these points we determined that the number of clusters is 4. This is because the total within sum of squares error was the lowest with 4 clusters and because there was negligible drop in total within

sum of squares errors subsequently (lines 431 - 434, figure below which has been included as Fig S3). We then extracted the minimum and maximum values of MCWD of each of the 4 clusters to be the thresholds of the zones (lines 434 - 435). The low tree cover zone spanned area below 1209 mm MCWD, the medium tree cover zone with savanna has MCWD range of 845 â 1209 mm, the mixed tree cover zone in the MCWD range of 486 - 845 mm and the high tree cover zone is above 486 mm of MCWD (lines 136 - 141).

17. L. 250: "climate change variability and climate change driven...", climate change was not studied, so I suggest to say "climate variability and climate driven...".
We have deleted âchangeâ as per the reviewerâs suggestion (line 264).

18. The ms highlights the importance of the study for tree planting and no-go areas for planting. The discussion could include explicit recommendation regarding such areas.
We agree with the reviewer that this study would be more useful if we discussed recommendations for management of no-go areas for tree-planting. However, the focus of this study is to estimate the distribution of tree cover and its drivers. Additionally, the focus is to elucidate potential mechanisms that maintain tree-grass dynamics within the Indian savanna biome and consequently biome distribution across India. Hence, we think that the results cannot be extrapolated to provide explicit recommendations of tree-planting. This is because such recommendations for tree-planting is a complex issue involving factors such as local people needs6,7, which we do not examine.

19. L. 361: please add reference for land cover map as well as the spatial resolution of the different data sets.
We have now included the references of land cover maps used with Refs 9,10 and have included the spatial resolution of 100m and 60m and that further analyses was done at 250m spatial resolution after resampling (lines 375 â 379) . We have then directed the reader to the Supplementary Methods for additional details about resampling method and information about coordinate systems.

20. L. 366: given that 300 points were sampled per MAP bin: does multimodality of tree cover vs MAP occur in the full data set? I also suggest adding the information to Fig 1 and Fig S1 that they show only a subset of the points.
There were 9600 points sampled in total across the study area (Fig S1) and we believe there is no discernible pattern of multimodality of the mean tree cover against mean annual precipitation or MCWD for these points, see figure below.
a b
))

We have now included the above figure as Fig S2 in the list of Supplementary Figures. We have also added the information that of the 9600 points, 89.1% i.e. 8554 points were considered in the captions of Fig 1 and the new figure of percent tree cover vs MAP with the four clusters (Fig S4). We have also accordingly made changes in the main text referring to the new Fig S2 and S4 (lines 133, 382, 420, 422 and 384 - 385).

21. L. 373: "long temporal resolution" does that mean high temporal resolution (ie number of time points in the study period) or the long study period (ie number of years)?
We mean long study period i.e. number of years for which annual information of tree cover is available. We have rephrased the sentence to indicate the long study period (lines 388 - 389).

22. L 382: I suggest adding a definition of MCWD
We have defined MCWD as the most negative value of the cumulative difference between precipitation and potential evapotranspiration as per Refs 11,12 (lines 399 - 400).

23. L. 410: how many data points were available from that data set?
Nerleker et al.1 assembled a dataset of 206 locational points of endemic savanna plants (47 families, 34 woody species with non-woody species being dominant- 135 forbs and 37 graminoids) of which we used 98 location points i.e. 47.5% of the available data consisting of locations of graminoids (species n=5 annuals and n=8 perennials), forbs (species n=8 annuals and n=55 perennials) and woody species (n=22) in the study area of our analyses.

We have included a sentence noting that 47.5% of the total dataset has been used in this study (line 436) and furthermore have provided additional information about the complete dataset from Nerleker et al., 2022 in the Supplementary Methods section.

24. Fig S2: precipitation ranges overlap in the caption high tree cover 0 to 600 and mixed from 580 to 930. Further 1300-1600 is mentioned for panel (a), this is the low tree cover zone. I suggest

modifying the colors and not use grey for unburned points in the histogram. Grey is used as background in maps, but the same color does not represent the same points/information in the maps and the histograms.
Thank you for catching the error in the figure caption about the MCWD range for the high tree cover zone; we have now corrected it. Also, we have remade the histogram subplots depicting burned/unburned points such that the part of the histogram showing frequency of unburned points is in white thereby avoiding confusion with the grey areas in the map.

25. Fig S6: in the PCA plot, names of herbivores overlap and cannot be read.
We have remade the figure such that the herbivore labels do not overlap and are legible.

References

1. Nerlekar, A. N. et al. Exponential rise in the discovery of endemic plants underscores the need to conserve the Indian savannas. 1â13 (2022) doi:10.1111/btp.13062.
2. Staver, C., Archibald, S. & Levin, S. The global extent and determinants of savanna and forest as alternative biome states. Science (80-.). 334, 230â232 (2011).
3. Hirota, M., Holmgren, M., van Nes, E. H. & Scheffer, M. Global Resilience of Tropical Forest. Science (80-.). 334, 232â235 (2011).
4. Pletcher, E., Staver, C. & Schwartz, N. B. The environmental drivers of tree cover and forestâ savanna mosaics in Southeast Asia. Ecography (Cop.). (2022) doi:10.1111/ecog.06280.
5. Ratnam, J., Tomlinson, K. W., Rasquinha, D. N. & Sankaran, M. Savannahs of Asia: Antiquity, biogeography, and an uncertain future. Philos. Trans. R. Soc. B Biol. Sci. 371, (2016).
6. Brancalion, P. H. S. & Holl, K. D. Guidance for Successful Tree Planting Initiatives. J. Appl. Ecol. 1â13 (2020) doi:10.1111/1365-2664.13725.
7. Holl, K. D. & Brancalion, P. H. S. Tree planting is not a simple solution. Science (80-.). 368, 580â581 (2020).
8. Ja, H. A k-means clustering algorithm. JR Stat. Soc. Ser. C-Appl. Stat. 28, 100â108 (1979).
9. Roy, P. S. et al. New vegetation type map of India prepared using satellite remote sensing: Comparison with global vegetation maps and utilities. Int. J. Appl. Earth Obs. Geoinf. 39, 142â 159 (2015).
10. Roy, P.S., P. Meiyappan, P.K. Joshi, M.P. Kale, V.K. Srivastav, S.K. Srivasatava, M.D. Behera, A. Roy, Y. Sharma, R.M. Ramachandran, P. Bhavani, A.K. Jain, and Y. V. N. K. Decadal Land Use and Land Cover Classifications across India. (2016) doi:https://doi.org/10.3334/ORNLDAAC/1336.
11. AragÃ£o, L. E. O. C. et al. Spatial patterns and fire response of recent Amazonian droughts. Geophys. Res. Lett. 34, 1â5 (2007).
12. Malhi, Y. et al. The linkages between photosynthesis, productivity, growth and biomass in lowland Amazonian forests. Glob. Chang. Biol. 21, 2283â2295 (2015).

Version 1:

Decision Letter:

** Please ensure you delete the link to your author home page in this e-mail if you wish to forward it to your coauthors **

Dear Dr Gopalakrishna,

Your manuscript titled "The distribution and drivers of tree cover in India" has now been seen by our reviewers, whose comments appear below. In light of their advice we are delighted to say that we are happy, in principle, to publish a suitably revised version in Communications Earth & Environment under the open access CC BY license (Creative Commons Attribution v4.0 International License).

We therefore invite you to edit your manuscript to comply with our format requirements and to maximise the accessibility and therefore the impact of your work.

EDITORIAL REQUESTS:

Please review our specific editorial comments and requests regarding your manuscript in the attached "Editorial Requests Table".

*****Please take care to match our formatting and policy requirements. We will check revised manuscript and return manuscripts that do not comply. Such requests will lead to delays. *****

Please outline your response to each request in the right hand column. Please upload the completed table with your manuscript files as a Related Manuscript file.

If you have any questions or concerns about any of our requests, please do not hesitate to contact me.

SUBMISSION INFORMATION:

In order to accept your paper, we require the files listed at the end of the Editorial Requests Table; the list of required files is also available at https://www.nature.com/documents/commsj-file-checklist.pdf .

OPEN ACCESS:

Communications Earth & Environment is a fully open access journal. Articles are made freely accessible on publication under a CC BY license (Creative Commons Attribution 4.0 International License). This license allows maximum dissemination and re-use of open access materials and is preferred by many research funding bodies.

For further information about article processing charges, open access funding, and advice and support from Nature Research, please visit https://www.nature.com/commsenv/article-processing-charges

At acceptance, you will be provided with instructions for completing this CC BY license on behalf of all authors. This grants us the necessary permissions to publish your paper. Additionally, you will be asked to declare that all required third party permissions have been obtained, and to provide billing information in order to pay the article-processing charge (APC).

Please use the following link to submit the above items:
Link Redacted
** This url links to your confidential home page and associated information about manuscripts you may have submitted or be reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage first **

We hope to hear from you within two weeks; please let us know if you need more time.

Best regards,

Martina Grecequet, PhD
Associate Editor,
Communications Earth & Environment
@CommsEarth

Erika Buscardo
Editorial Board Member
Communications Earth & Environment

REVIEWERS' COMMENTS:

Reviewer #1 (Remarks to the Author):

I'm satisfied with the answers.

Reviewer #2 (Remarks to the Author):

The authors addressed my previous comments and I don’t have additional comments.