THE ROYAL SOCIETY
PUBLISHING

PROCEEDINGS B

Determinants of tree cover in tropical floodplains

Joshua H. Daskin, Filipe Aires and A. Carla Staver

Article citation details
Proc. R. Soc. B **286**: 20191755.
http://dx.doi.org/10.1098/rspb.2019.1755

Review timeline

Original submission:	31 January 2019
1st revised submission:	5 August 2019
2nd revised submission:	26 September 2019
Final acceptance:	4 October 2019

Note: Reports are unedited and appear as submitted by the referee. The review history appears in chronological order.

Review History

RSPB-2019-0275.R0 (Original submission)

Review form: Reviewer 1

Recommendation
Major revision is needed (please make suggestions in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?
Good

General interest: Is the paper of sufficient general interest?
Acceptable

Quality of the paper: Is the overall quality of the paper suitable?
Good

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
Yes

Is it clear?
Yes

Is it adequate?
Yes

Do you have any ethical concerns with this paper?
No

Comments to the Author
Summary:
This manuscript evaluates the environmental controls of remotely-sensed tree cover in tropical floodplains. Previous research in tropical tree cover focused more on terra firme or upland forests. By combining novel high-resolution data sets, the authors try to ask (1) what are the key determinants of tropical tree cover in floodplains and (2) how important is the flooding intensity/variability. Both random forests and linear mixed effect models are used to detect the importance and sensitivity of each environmental variables. Their results suggest that dominant environmental determinants of floodplain tree cover are similar to those for upland forests (water supply, fire). Elevation is also an important determinant with its effect vary significantly across different regions. Flooding has positive effects in dry regions but a negative effect in wet regions. However, the importance of flooding regimes is low compared with other determinants.

The study is interesting in the sense that the authors try to quantify the flooding effects using novel high-resolution data sets and conducted a cross-continental comprehensive statistical analyses. However, I have some concerns as listed as follows:

(1) Overall, what are the special things about tropical floodplains? The manuscript leaves me the impression that floodplain forests are so different from upland/terra firme forests in the tropics (e.g. we can almost use the same set of predictors like aridity, elevation, and fire). In the abstract, the authors stress that more prolonged and frequent flooding can influence tree cover with effects depending on moisture regimes [line 25-27] while in the discussion, they summarized that 'the effects of flood regimes were surprisingly weak' [line 287-289]. At the end of the day, what can we learn from this analysis beyond the recent studies on floodplain tree cover such as Flores et al. 2017, which also pointed out the importance of moisture supply, fire and soil?

(2) I am not so sure how I should interpret the linear slopes from Figure 3 (probably, only the climatic water balance and sand content are quasi-linear). To me, the linear slope is only useful to qualitatively depict the importance of each variable. I was wondering why not just show the partial dependence plots from the random forests results and describe the non-linearity, which can be more useful and novel to the whole community.

(3) Currently, the discussion is organized by different environmental factors. I feel an overarching paragraph to summarize how the environmental determinants in floodplain differ from those in upland forests would be helpful.

See below for minor comments following the order of the text:

Line 29-31: I wonder how much this effect can be, given that the flood duration effect is rather small (Figure 3). Decreasing flood duration from 12 months to 0 months might only induce a few percentage changes in tree cover.

L. 100-102: I appreciate the efforts to combine high-resolution spatial data. My question is that do all data products have an original resolution of 90m or higher? To me, it makes more sense to aggregate (do the analysis at the coarsest resolution) instead of downscale (unless we know pretty sure how to represent fine-scale heterogeneity). At least, the soil data and MODIS fire data are at 250. How was downscaling performed? Why not just conduct the analysis at 250m?

L. 134-136: It is OK to use this self-defined metric for variability although the metric would be the same between the scenario of a flooding every alternate year and that of continuous flooding for the first half and no flooding for the second half... I want to comment a little it on the computational efficiency. Even the Colwell's predictability or other entropy-based seasonality indices might not be so computationally expensive if the codes use some simple optimization for matrix manipulation. For example, just use the numpy package in python can significantly reduce computational time for spatial analysis from my own experience.

L. 268-270: 'However, this effect saturated in the wettest regions'. Is this really unexpected? The previous analysis also shows the saturated effect of water balance in upland tropical forests.

Review form: Reviewer 2

Recommendation
Accept with minor revision (please list in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?
Excellent

General interest: Is the paper of sufficient general interest?
Excellent

Quality of the paper: Is the overall quality of the paper suitable?
Good

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
Yes

Is it clear?
N/A

Is it adequate?
N/A

Do you have any ethical concerns with this paper?
No

Comments to the Author
This is an important contribution, and I am unaware of any similar study that has attempted to examine drivers of flood plain tree cover at such a large geographic scale. This is a nice complement to global studies of tropical tree cover that have ignored flooding as a contributing factor, and I expect it to be widely cited. Most of my comments are minor, but I have a couple of concerns that require attention.
Humans are not mentioned in this study at all. Isn't agriculture important in the Nile flood plain, for example? Is direct human cutting of forests important in any of the regions? The authors need to address whether direct human influence is of sufficient scale to influence the relationships, and if so how this may influence conclusions.
The authors find fire to be a strong predictor of tree cover, and interprets its causality as being entirely unidirectional (i.e. frequent fire reduces tree cover and this effect is stronger in wetter regions). I would argue that causality in the opposite direction dominates the relationship (areas with low tree cover are more flammable, and this effect is stronger in wet areas). This being said, I believe that fire probably does have a stronger impacts in wetter areas, but it will come across as disingenuous if you simply assume the direction of causality and do not acknowledge the complexity here.
Unless I missed it, the authors did not mention that it may be particularly challenging for trees to cope with alternating flooding and water deficit. There seem to be relatively few species that can cope with this, such as Melaleuca in Australia, which interestingly has become invasive in the Everglades. Some palms seem particularly good at this as well. There may be good reason to expect alternating anoxia and drought to be a particularly difficult situation to deal with because, for example, coping with waterlogged soil generally requires shallow roots, while coping with water deficit generally requires deep roots. I cannot think of a specific reference that deals with this specific constraint, so I would not insist that the authors discuss this, but it does seem to provide context for many of the results.
Line 64. Varzea is spelled with an accent over the first 'a'. Also, I would specify 'varzea forest' here, since the term varzea also includes grassy areas.
Line 160-163. Check this explanation. Doesn't a greater MSE indicate a poorer fit and lower variable importance?
Line 185. Remove hyphen in 'water-balance'.
Line 270. I don't understand how table 1 shows this saturation in wetter regions. Figure 3 does a better job of convincing me of this, but also suggests a more likely explanation for the saturation, i.e. tree cover saturates at high water balance because it is impossible to have more than 100% tree cover. Since tree cover inherently saturates at 100% it seems unreasonable to seek a mechanism to explain this, and I suggest not highlighting it by devoting the first paragraph to this point.

Line 344-354. This paragraph was a little confusing. I think it would help to make it clear early in the paragraph that, in upland savanna, sandy soils are associated with high tree cover. Many readers will assume the opposite, which makes later statements confusing. Also, clarify why would including forests would change the result (line 353).
Line 373. Change ‘shrub’ to ‘tree’? Both are true, but this study focuses on trees.

Decision letter (RSPB-2019-0275.R0)

05-Jun-2019

Dear Dr Daskin,

We have now received two referees' reports on your manuscript RSPB-2019-0275 entitled "Determinants of tree cover in tropical floodplains". I apologise for the time this has taken as we had trouble finding referees.

The manuscript has, in its current form, been rejected for publication in Proceedings B. This action has been taken on the advice of referees, who have recommended that substantial revisions are necessary. With this in mind we would be happy to consider a resubmission, provided the comments of the referees are fully addressed. However please note that this is not a provisional acceptance.

The resubmission will be treated as a new manuscript. However, we will approach the same reviewers if they are available and it is deemed appropriate to do so by the Editor. Please note that resubmissions must be submitted within six months of the date of this email. In exceptional circumstances, extensions may be possible if agreed with the Editorial Office. Manuscripts submitted after this date will be automatically rejected.

Please find below the comments made by the referees, not including confidential reports to the Editor, which I hope you will find useful. If you do choose to resubmit your manuscript, please upload the following:

1) A ‘response to referees’ document including details of how you have responded to the comments, and the adjustments you have made.
2) A clean copy of the manuscript and one with 'tracked changes' indicating your 'response to referees' comments document.
3) Line numbers in your main document.

To upload a resubmitted manuscript, log into http://mc.manuscriptcentral.com/prsb and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Resubmission." Please be sure to indicate in your cover letter that it is a resubmission, and supply the previous reference number.

Sincerely,
Professor Loeske Kruuk
mailto: proceedingsb@royalsociety.org

Associate Editor
Board Member: 1
Comments to Author:
We have now received two reviews of this paper on the drivers of tree cover in tropical floodplains. Floodplain forests are quite a complex and also understudied ecosystem. This study is interesting in that it both deals with the complexity of the system while maintaining a large geographic scope, which gives the analysis the potential to be interesting to a wide audience. However, the current manuscript does not express its results in a way that communicates to a broad audience. Two main areas for improvement:

1. as pointed out by reviewer 1: the manuscript needs highlight better the uniqueness of floodplain ecosystems (e.g. are the topographic effects in flood plains different from those in upland forests)

2. as pointed out by reviewer 2: the effect of humans needs to clearer and more fully explored both in the analysis and the text. Part of the rationale for publishing at Proc B has to be the conservation importance of tropical floodplains and the manuscript needs to address what humans are doing in these ecosystems to affect tree cover.

The reviewers also have many other specific comments to improve the manuscript moving forward.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s)
Summary:
This manuscript evaluates the environmental controls of remotely-sensed tree cover in tropical floodplains. Previous research in tropical tree cover focused more on terra firme or upland forests. By combining novel high-resolution data sets, the authors try to ask (1) what are the key determinants of tropical tree cover in floodplains and (2) how important is the flooding intensity/variability. Both random forests and linear mixed effect models are used to detect the importance and sensitivity of each environmental variables. Their results suggest that dominant environmental determinants of floodplain tree cover are similar to those for upland forests (water supply, fire). Elevation is also an important determinant with its effect vary significantly across different regions. Flooding has positive effects in dry regions but a negative effect in wet regions. However, the importance of flooding regimes is low compared with other determinants.

The study is interesting in the sense that the authors try to quantify the flooding effects using novel high-resolution data sets and conducted a cross-continental comprehensive statistical analyses. However, I have some concerns as listed as follows:

(1) Overall, what are the special things about tropical floodplains? The manuscript leaves me the impression that floodplain forests are so different from upland/terra firme forests in the tropics (e.g. we can almost use the same set of predictors like aridity, elevation, and fire). In the abstract, the authors stress that more prolonged and frequent flooding can influence tree cover with effects depending on moisture regimes [line 25-27] while in the discussion, they summarized that 'the effects of flood regimes were surprisingly weak' [line 287-289]. At the end of the day, what can we learn from this analysis beyond the recent studies on floodplain tree cover such as Flores et al. 2017, which also pointed out the importance of moisture supply, fire and soil?

(2) I am not so sure how I should interpret the linear slopes from Figure 3 (probably, only the climatic water balance and sand content are quasi-linear). To me, the linear slope is only useful to qualitatively depict the importance of each variable. I was wondering why not just show the partial dependence plots from the random forests results and describe the non-linearity, which can be more useful and novel to the whole community.

(3) Currently, the discussion is organized by different environmental factors. I feel an overarching paragraph to summarize how the environmental determinants in floodplain differ from those in upland forests would be helpful.

See below for minor comments following the order of the text:

Line 29-31: I wonder how much this effect can be, given that the flood duration effect is rather small (Figure 3). Decreasing flood duration from 12 months to 0 months might only induce a few percentage changes in tree cover.

L. 100-102: I appreciate the efforts to combine high-resolution spatial data. My question is that do all data products have an original resolution of 90m or higher? To me, it makes more sense to aggregate (do the analysis at the coarsest resolution) instead of downscale (unless we know pretty sure how to represent fine-scale heterogeneity). At least, the soil data and MODIS fire data are at 250. How was downscaling performed? Why not just conduct the analysis at 250m?

L. 134-136: It is OK to use this self-defined metric for variability although the metric would be the same between the scenario of a flooding every alternate year and that of continuous flooding for the first half and no flooding for the second half... I want to comment a little it on the computational efficiency. Even the Colwell's predictability or other entropy-based seasonality indices might not be so computationally expensive if the codes use some simple optimization for matrix manipulation. For example, just use the numpy package in python can significantly reduce computational time for spatial analysis from my own experience.

L. 268-270: 'However, this effect saturated in the wettest regions'. Is this really unexpected? The previous analysis also shows the saturated effect of water balance in upland tropical forests.

Referee: 2

Comments to the Author(s)
This is an important contribution, and I am unaware of any similar study that has attempted to examine drivers of flood plain tree cover at such a large geographic scale. This is a nice complement to global studies of tropical tree cover that have ignored flooding as a contributing factor, and I expect it to be widely cited. Most of my comments are minor, but I have a couple of concerns that require attention.
Humans are not mentioned in this study at all. Isn’t agriculture important in the Nile flood plain, for example? Is direct human cutting of forests important in any of the regions? The authors need to address whether direct human influence is of sufficient scale to influence the relationships, and if so how this may influence conclusions.
The authors find fire to be a strong predictor of tree cover, and interprets its causality as being entirely unidirectional (i.e. frequent fire reduces tree cover and this effect is stronger in wetter regions). I would argue that causality in the opposite direction dominates the relationship (areas with low tree cover are more flammable, and this effect is stronger in wet areas). This being said, I believe that fire probably does have a stronger impacts in wetter areas, but it will come across as

disingenuous if you simply assume the direction of causality and do not acknowledge the complexity here.
Unless I missed it, the authors did not mention that it may be particularly challenging for trees to cope with alternating flooding and water deficit. There seem to be relatively few species that can cope with this, such as Melaleuca in Australia, which interestingly has become invasive in the Everglades. Some palms seem particularly good at this as well. There may be good reason to expect alternating anoxia and drought to be a particularly difficult situation to deal with because, for example, coping with waterlogged soil generally requires shallow roots, while coping with water deficit generally requires deep roots. I cannot think of a specific reference that deals with this specific constraint, so I would not insist that the authors discuss this, but it does seem to provide context for many of the results.
Line 64. Varzea is spelled with an accent over the first 'a'. Also, I would specify 'varzea forest' here, since the term varzea also includes grassy areas.
Line 160-163. Check this explanation. Doesn't a greater MSE indicate a poorer fit and lower variable importance?
Line 185. Remove hyphen in 'water-balance'.
Line 270. I don't understand how table 1 shows this saturation in wetter regions. Figure 3 does a better job of convincing me of this, but also suggests a more likely explanation for the saturation, i.e. tree cover saturates at high water balance because it is impossible to have more than 100% tree cover. Since tree cover inherently saturates at 100% it seems unreasonable to seek a mechanism to explain this, and I suggest not highlighting it by devoting the first paragraph to this point.
Line 344-354. This paragraph was a little confusing. I think it would help to make it clear early in the paragraph that, in upland savanna, sandy soils are associated with high tree cover. Many readers will assume the opposite, which makes later statements confusing. Also, clarify why would including forests would change the result (line 353).
Line 373. Change 'shrub' to 'tree'? Both are true, but this study focuses on trees.

Author's Response to Decision Letter for (RSPB-2019-0275.R0)

See Appendix A.

RSPB-2019-1755.R0

Review form: Reviewer 3 (Zhenzhong Zeng)

Recommendation
Accept with minor revision (please list in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?
Good

General interest: Is the paper of sufficient general interest?
Excellent

Quality of the paper: Is the overall quality of the paper suitable?
Excellent

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
Yes

Is it clear?
Yes

Is it adequate?
Yes

Do you have any ethical concerns with this paper?
No

Comments to the Author
I now read through the manuscript, supplementary information, and the rebuttal letter. It is a well-done research and the authors have addressed most of the concerns raised by the referees. I recommend accept after minor revisions.
First, water balance is computed as rainfall minus potential evapotranspiration; but scientifically, it should be rainfall minus actual evapotranspiration. Potential evapotranspiration does not represent actual evapotranspiration. Please use actual evapotranspiration or explain the reason.
Second, the percentage tree cover from Hansen et al., 2013 is for the year of 2000 rather than 2010 as mentioned by the authors (Line 112). If the authors used percentage tree cover for 2000 minus forest loss during 2000-2010, they may neglect the forest gain during 2000-2010. But Hansen et al 2013 had not provided forest gain during 2000-2010.
Third, I find it difficult to understand the method without reading "Delineation of study regions" in Supplementary Information. I recommend moving this material to the main text (combined to Study regions and data extraction in Methods).
Lines 218-220, I do not observe Northern Nile in Table S2.
Lines 225, 231 and more, there is no Table S3 in Supplementary Information.
Lines 434-436, it is unclear how the authors get the number of tree cover loss caused by dams. Which dams and how they calculate?
Figure 3(b), it is difficult to understand why model-predicted tree cover increases with burned times for dry climate (red line).

Decision letter (RSPB-2019-1755.R0)

11-Sep-2019

Dear Dr Daskin:

Your manuscript has now been peer reviewed and the review has been assessed by an Associate Editor. The reviewer's comments and the comments from the Associate Editor are included at the end of this email for your reference. As you will see, the reviewer has raised some concerns with your manuscript and we would like to invite you to revise your manuscript to address them.

We do not allow multiple rounds of revision so we urge you to make every effort to fully address all of the comments at this stage. If deemed necessary by the Associate Editor, your manuscript will be sent back to one or more of the original reviewers for assessment. If the original reviewers are not available we may invite new reviewers. Please note that we cannot guarantee eventual acceptance of your manuscript at this stage.

To submit your revision please log into http://mc.manuscriptcentral.com/prsb and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions", click on "Create a Revision". Your manuscript number has been appended to denote a revision.

When submitting your revision please upload a file under "Response to Referees" in the "File Upload" section. This should document, point by point, how you have responded to the reviewers' and Editors' comments, and the adjustments you have made to the manuscript. We require a copy of the manuscript with revisions made since the previous version marked as 'tracked changes' to be included in the 'response to referees' document.

Your main manuscript should be submitted as a text file (doc, txt, rtf or tex), not a PDF. Your figures should be submitted as separate files and not included within the main manuscript file.

When revising your manuscript you should also ensure that it adheres to our editorial policies (https://royalsociety.org/journals/ethics-policies/). You should pay particular attention to the following:

Research ethics:
If your study contains research on humans please ensure that you detail in the methods section whether you obtained ethical approval from your local research ethics committee and gained informed consent to participate from each of the participants.

Use of animals and field studies:
If your study uses animals please include details in the methods section of any approval and licences given to carry out the study and include full details of how animal welfare standards were ensured. Field studies should be conducted in accordance with local legislation; please include details of the appropriate permission and licences that you obtained to carry out the field work.

Data accessibility and data citation:
It is a condition of publication that you make available the data and research materials supporting the results in the article. Datasets should be deposited in an appropriate publicly available repository and details of the associated accession number, link or DOI to the datasets must be included in the Data Accessibility section of the article

(https://royalsociety.org/journals/ethics-policies/data-sharing-mining/). Reference(s) to datasets should also be included in the reference list of the article with DOIs (where available).

In order to ensure effective and robust dissemination and appropriate credit to authors the dataset(s) used should also be fully cited and listed in the references.

If you wish to submit your data to Dryad (http://datadryad.org/) and have not already done so you can submit your data via this link http://datadryad.org/submit?journalID=RSPB&manu=(Document not available), which will take you to your unique entry in the Dryad repository.

If you have already submitted your data to dryad you can make any necessary revisions to your dataset by following the above link.

For more information please see our open data policy http://royalsocietypublishing.org/data-sharing.

Electronic supplementary material:
All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI. Please try to submit all supplementary material as a single file.

Online supplementary material will also carry the title and description provided during submission, so please ensure these are accurate and informative. Note that the Royal Society will not edit or typeset supplementary material and it will be hosted as provided. Please ensure that the supplementary material includes the paper details (authors, title, journal name, article DOI). Your article DOI will be 10.1098/rspb.[paper ID in form xxxx.xxxx e.g. 10.1098/rspb.2016.0049].

Please submit a copy of your revised paper within three weeks. If we do not hear from you within this time your manuscript will be rejected. If you are unable to meet this deadline please let us know as soon as possible, as we may be able to grant a short extension.

Thank you for submitting your manuscript to Proceedings B; we look forward to receiving your revision. If you have any questions at all, please do not hesitate to get in touch.

Best wishes,
Professor Loeske Kruuk
mailto: proceedingsb@royalsociety.org

Associate Editor
Comments to Author:
I agree with the reviewer that the authors have addressed the major issues raised in the previous round of review. There are a few outstanding issues to be dealt with especially related to AET versus PET and the Hansen loss/gain from specific years. I look forward to these revisions.

Reviewer(s)' Comments to Author:

Referee: 3

Comments to the Author(s).
I now read through the manuscript, supplementary information, and the rebuttal letter. It is a well-done research and the authors have addressed most of the concerns raised by the referees. I recommend accept after minor revisions.
First, water balance is computed as rainfall minus potential evapotranspiration; but scientifically, it should be rainfall minus actual evapotranspiration. Potential evapotranspiration does not represent actual evapotranspiration. Please use actual evapotranspiration or explain the reason.
Second, the percentage tree cover from Hansen et al., 2013 is for the year of 2000 rather than 2010 as mentioned by the authors (Line 112). If the authors used percentage tree cover for 2000 minus forest loss during 2000-2010, they may neglect the forest gain during 2000-2010. But Hansen et al 2013 had not provided forest gain during 2000-2010.
Third, I find it difficult to understand the method without reading "Delineation of study regions" in Supplementary Information. I recommend moving this material to the main text (combined to Study regions and data extraction in Methods).
Lines 218-220, I do not observe Northern Nile in Table S2.
Lines 225, 231 and more, there is no Table S3 in Supplementary Information.
Lines 434-436, it is unclear how the authors get the number of tree cover loss caused by dams. Which dams and how they calculate?
Figure 3(b), it is difficult to understand why model-predicted tree cover increases with burned times for dry climate (red line).

Author's Response to Decision Letter for (RSPB-2019-1755.R0)

See Appendix B.

Decision letter (RSPB-2019-1755.R1)

04-Oct-2019

Dear Dr Daskin

I am pleased to inform you that your manuscript entitled "Determinants of tree cover in tropical floodplains" has been accepted for publication in Proceedings B.

You can expect to receive a proof of your article from our Production office in due course, please check your spam filter if you do not receive it. PLEASE NOTE: you will be given the exact page length of your paper which may be different from the estimation from Editorial and you may be asked to reduce your paper if it goes over the 10 page limit.

If you are likely to be away from e-mail contact please let us know. Due to rapid publication and an extremely tight schedule, if comments are not received, we may publish the paper as it stands.

If you have any queries regarding the production of your final article or the publication date please contact procb_proofs@royalsociety.org

Open Access
You are invited to opt for Open Access, making your freely available to all as soon as it is ready for publication under a CCBY licence. Our article processing charge for Open Access is £1700. Corresponding authors from member institutions (http://royalsocietypublishing.org/site/librarians/allmembers.xhtml) receive a 25% discount to these charges. For more information please visit http://royalsocietypublishing.org/open-access.

Your article has been estimated as being 10 pages long. Our Production Office will be able to confirm the exact length at proof stage.

Paper charges
An e-mail request for payment of any related charges will be sent out after proof stage (within approximately 2-6 weeks). The preferred payment method is by credit card; however, other payment options are available

Electronic supplementary material:
All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI.

Thank you for your fine contribution. On behalf of the Editors of the Proceedings B, we look forward to your continued contributions to the Journal.

Sincerely,

Professor Loeske Kruuk
Editor, Proceedings B
mailto: proceedingsb@royalsociety.org

Associate Editor:
Board Member
Comments to Author:
Thank you for your thoughtful response to the reviewers. I feel confident that this manuscript will make a major contribution to the literature in this area.

Appendix A

Joshua H. Daskin, Ph.D.
Gaylord S. Donnelley Postdoctoral Environmental Fellow
Yale University
Osborn Memorial Laboratories
165 Prospect St.
New Haven, CT 06511
(*e*) joshua.daskin@yale.edu
(*p*) 914-960-2270

August 3, 2019

Dear Dr. Kruuk,

Thank you for your letter of June 5 inviting a resubmission of our manuscript RSPB-2019-0275 entitled "**Determinants of tree cover in tropical floodplains**". We very much appreciated the constructive and careful feedback from you and the two reviewers.

We have now considered and replied to each editor and reviewer comment and present these responses below in blue. In particular, we have tried to better present the differences and similarities between the results of our floodplain-specific analysis and those from studies of upland tree cover. We have also given additional space to describing the applied importance of our study for the conservation of floodplain ecosystems and moved much of the original table 1 to the supplement to accommodate the 10-page limit. Line numbers in blue refer to the tracked-changes version of the revised manuscript. We have also included a clean version of the new manuscript, as requested.

We hope that our revision will now be judged suitable for publication in *Proceedings B*, although we of course remain willing to consider any additional changes deemed necessary by the editors.

Sincerely,
Joshua H. Daskin, Ph.D.

on behalf of
Felipe Aires, Ph.D. and A. Carla Staver, Ph.D.

Associate Editor
Board Member: 1
Comments to Author:
We have now received two reviews of this paper on the drivers of tree cover in tropical floodplains. Floodplain forests are quite a complex and also understudied ecosystem. This study is interesting in that it both deals with the complexity of the system while maintaining a large geographic scope, which gives the analysis the potential to be interesting to a wide audience.
We appreciate this compliment. The ability to investigate ecological characteristics—here especially tree cover and flood regimes—of multiple ecosystems at medium-to-high resolution allows exactly this kind of cross-continental comparison without abstracting away too much detail to be informative within complex ecosystems. Indeed, one motivation for this study was the recent availability of relatively fine (spatial and temporal) scale flood regime data from the GIEMS-D3 dataset (Aires et al. 2017).

However, the current manuscript does not express its results in a way that communicates to a broad audience. Two main areas for improvement:

1. as pointed out by reviewer 1: the manuscript needs highlight better the uniqueness of floodplain ecosystems (e.g. are the topographic effects in flood plains different from those in upland forests)
Certainly we agree it is important to differentiate between the present study of tropical floodplain tree cover determinants and those of tree cover in upland ecosystems. Given the concern from both the editor and reviewer one, we have made substantial efforts to ensure this comes through better in the revised version of our manuscript.

We have completely rewritten the abstract and now include explicit statements comparing the drivers of floodplain and upland tree cover:

> "*…we show that floodplain tree cover increases with climatic water balance (= rainfall - potential evapotranspiration) and decreases with frequent fire, as in uplands. Additionally, in floodplains, higher elevations and steeper slopes have higher tree cover, likely because these areas escape the worst inundation*"

Indeed, topographic effects are strong in most of the floodplain regions we studied, likely because higher elevations and steeper slopes avoid long-term water logging during flood events. To our knowledge, elevation and slope do not have consistent effects on tree cover in upland ecosystems.

Additionally, and as in our original manuscript (original lines 345–354), we also describe the clear differences in how tree cover is related to soil composition in floodplains versus uplands. This is now on lines 395–411 of the discussion.

Finally, although not the strongest predictors, we did find that flood regimes impacted tree cover. Of course, this does not occur in uplands.

As suggested by Reviewer 1, we have also added a paragraph to the start of the discussion section (lines 278–284) which outlines our findings and draws these same contrasts between floodplains and uplands.

2. as pointed out by reviewer 2: the effect of humans needs to clearer and more fully explored both in the analysis and the text. Part of the rationale for publishing at Proc B has to be the conservation importance of tropical floodplains and the manuscript needs to address what humans are doing in these ecosystems to affect tree cover.

Without a doubt, we are concerned with the conservation of floodplain ecosystems for their biodiversity and for the ecosystem services they provide. And, floodplains are certainly under threat from agricultural and residential development, resource harvesting (logging, fishing, mining), nutrient deposition, and climate change. These human impacts interact with ecological processes, which makes a fundamental understanding of the ecological and hydrological drivers of floodplain tree cover in natural ecosystems essential. To this end, as described in the original submission (original line 95 and original Supplementary Information lines 38 and 58–68), we removed all deforested, agricultural (crops and livestock), and urban areas. To be sure this is clear, we have now added detail to the main text methods section on lines 106–107).

Despite focusing primarily on the basic ecological aspects of floodplains, because both the editor and Reviewer 2 expressed interest in a more human-centric angle to paper, we have amended our discussion to more fully describe the potential impacts of dams (particularly their impoundments) and to mention the importance of basic ecological studies for informing the protection of wetlands in the face of ongoing human encroachment.

The reviewers also have many other specific comments to improve the manuscript moving forward.

Referee: 1
Comments to the Author(s)
Summary:
This manuscript evaluates the environmental controls of remotely-sensed tree cover in tropical floodplains. Previous research in tropical tree cover focused more on terra firme or upland forests. By combining novel high-resolution data sets, the authors try to ask (1) what are the key determinants of tropical tree cover in floodplains and (2) how important is the flooding intensity/variability. Both random forests and linear mixed effect models are used to detect the importance and sensitivity of each environmental variables. Their results suggest that dominant environmental determinants of floodplain tree cover are similar to those for upland forests (water supply, fire). Elevation is also an important determinant with its effect vary significantly across different regions. Flooding has positive effects in dry regions but a negative effect in wet regions. However, the importance of flooding regimes is low compared with other determinants.

We thank the reviewer for this excellent summary of our work. In addition to his/her summary of our findings, we would like to draw attention to the result that water balance not only had direct effects on floodplain tree cover (which is similar to the effects in uplands) and interacted with flood duration, but also interacted with topography and fire. We are not aware of an elevational effect or an interaction between elevation and climatic water balance in uplands. While the flood regime variables were relatively weak predictors, we have edited parts of the and discussion (lines 320–340) to better convey that elevational effects very likely reflect flood impacts over longer time scales than captured by the GIEMS-D3 data.

The study is interesting in the sense that the authors try to quantify the flooding effects using novel high-resolution data sets and conducted a cross-continental comprehensive statistical analyses.

Again, we appreciate the reviewer's kind words regarding our approach. As mentioned in our reply to the editor above, it was indeed one of our goals to utilize the recently improved fine scale flood regime data made available in the GIEMS-D3 product to make cross-continental comparisons at a high enough resolution to retain within-ecosystem relevance.

However, I have some concerns as listed as follows:

(1) Overall, what are the special things about tropical floodplains? The manuscript leaves me the impression that floodplain forests are so different from upland/terra firme forests in the tropics (e.g. we can almost use the same set of predictors like aridity, elevation, and fire). In the abstract, the authors stress that more prolonged and frequent flooding can influence tree cover with effects depending on moisture regimes [line 25-27] while in the discussion, they summarized that 'the effects of flood regimes were surprisingly weak' [line 287-289]. At the end of the day, what can we learn from this analysis beyond the recent studies on floodplain tree cover such as Flores et al. 2017, which also pointed out the importance of moisture supply, fire and soil?

Our manuscript provides several new and important insights about both the determinants of floodplain tree cover, and the differences between upland and floodplain tree cover correlates. We elaborate on aspects of this in our reply to the editor's first comment (see above). Here we elaborate further on the differences between upland and floodplain ecosystems, and answer specific questions from this reviewer's comments, e.g., re: the Flores et al. 2017 paper.

First, it is interesting that flood regimes have little effect on tree cover. There is a large literature on the physiological mechanisms by which individual trees cope with flooding and anoxia [e.g., 1–4]. Analogously, we know a lot about the mechanisms by which trees can cope with drought, *and* we know that arid places have fewer trees [5–7]. However, we did not know before this study whether or how flood regimes shape floodplain tree cover. We found that the answer is that flooding doesn't affect tree cover nearly as strongly as precipitation, fire, or elevation. This already represents a step forward; to us, it is surprising that flood regimes do not have particularly strong effects on floodplain tree cover, given the potentially lethal effects of inundation and anoxia on individual trees.

Our analyses also build on the Flores et al. 2017 paper. Their analysis is focused entirely on the Amazon basin. We include the Amazon, but also 12 other tropical and sub-tropical floodplain ecosystems. We also include several predictors (elevation, distance to nearest river and large river, and all our flood regime metrics) that were not considered by Flores et al. (or any other cross-ecosystem studies of floodplain tree cover which we are aware of). While perhaps intuitive, our demonstration that tree cover is higher in higher-elevation locations within floodplains is novel, and helps capture some of the likely impacts of long-term (centennial to millennial) flood regimes that 15 years of GIEMS-D3 data do not capture. (The reviewer mentions 'elevation' here as one of the factors that our results share with those from studies of upland tree cover; however, we are not aware of studies showing that elevation is a major determinant of tree cover in the tropics, excepting the effect of very high elevation in montane regions which create the alpine tree line phenomenon [e.g., Bader and Ruitjen, 2008 *J. Biogeography*)] As mentioned above we have also edited part of the discussion (lines 320–340) to better reflect that the strong effects of elevation very likely impacts of flooding over these longer time scales. Flores et al. also do not consider the decadal scale climate trends we include (flooding and water balance); rather, their paper was focused on specific drought events—climatic extremes, albeit events that may become more frequent.

Our study is also the first ecological study to utilize the GIEMS-D3 data (Aires et al. 2017 *J. Hydrometeorology*), a high spatial resolution temporally varying inundation dataset that is less sensitive to the presence of vegetation than other similar datasets (Aires et al. 2018 *Remote Sens Environ*).

Finally, this reviewer's comment (3) and Reviewer 2's question about the effect of soil composition is related to the present question about what is novel in our study. Reviewer 1 recommended a new opening paragraph to the discussion to better summarize the differences between upland and floodplain tree cover determinants. We have now written such a paragraph (lines 278–284), which we hope will help satisfy the reviewer. In response to Reviewer 2, we clarified the discussion paragraph about soil effects (lines 395–411), which focuses primarily on the differences between floodplains and uplands.

(2) I am not so sure how I should interpret the linear slopes from Figure 3 (probably, only the climatic water balance and sand content are quasi-linear). To me, the linear slope is only useful to qualitatively depict the importance of each variable. I was wondering why not just show the partial dependence plots from the random forests results and describe the non-linearity, which can be more useful and novel to the whole community.
We recognize that the reviewer's suggestion to "just show the partial dependence plots" is the most common approach to presenting random forest regression results. However, we disagree somewhat about the strength of inference possible from partial dependence plots relative to the linear regressions, especially for a study (like ours) seeking to explain the effect of predictors on a response, rather than to predict the value of future observations from given predictor values. It is well established that a strength of machine learning techniques (including random forests) is prediction, whereas classical statistical modeling (like regression) is more suited to quantitative explanatory analysis [8–10].

In our original methods section, we explained our decision to use linear models to determine directional effects:

> "Despite its strengths for identifying important predictors and dealing with collinearities, non-normality, and higher-order interactions, it can be difficult to interpret the directional effects of weaker and non-monotonic predictors from random forest regression. Often the direction of effects are visualized using partial dependence plots [11], but interpretation of these plots is subjective. We therefore used linear models—again one global and one for each region—to determine the direction, strength, and significance of predictors' effects on tree cover."

Partial dependence plots make it difficult to visualize two key components of any multiple regression analysis. First, although random forests and other decision tree-based methods handle interactions between predictors very well when forecasting the expected value of future observations, partial dependence plots do not illustrate these interactions (i.e., they are two-dimensional). As an example, we present below (Fig. R1) the partial dependence plot (from the random forest regression) and the conditional regression plot (from the linear regression) for the effects of soil sand content on Everglades floodplain tree cover data. As in the full manuscript, the green, purple, and red lines depict the interaction of soil sand content and climatic water balance. They show the regression results for the 90th, 50th, and 10th percentiles of water balance. From the figure, it is clear that information regarding the interaction between climatic water balance and soil sand content is lost in the partial dependence plot.

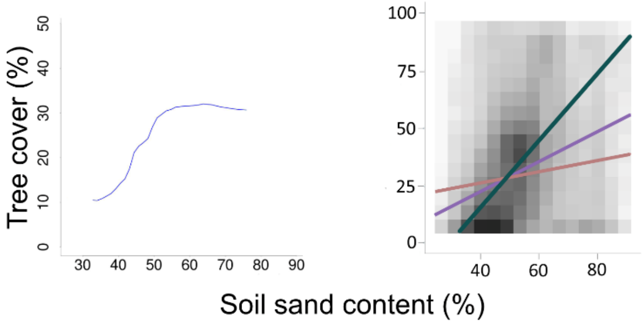

Figure R1. Partial dependence plot (left) and conditional regression plot (right) illustrating the effect of soil sand content on tree cover in Florida Everglades floodplains. The partial dependence plot is derived from the random forest regression and shows how the model-predicted value of tree cover changes across levels of sand content, correcting for the effects of all other predictors. The conditional regression plot is derived from OLS regression and shows model-predicted tree cover for locations at the 10th, 50th, and 90th percentile of climatic water balance (red, purple, and green lines, respectively; i.e., showing the interaction between water balance and soil composition). Other predictors are held constant at their medians. 95% confidence intervals are plotted around the regression lines but are almost always thinner than the line's width. The plot background on the right shows the relative density of tree cover observations in bivariate space.

Second, a *major* limitation to interpretation of partial dependence plots, is that computation of confidence intervals around the plotted lines (and around random forest model results generally) remains an active area of research [12–14]. Without confidence intervals, it is difficult to determine whether the displayed non-linearities (which the reviewer asked about) and non-monotonic curves are real or artifacts of the data. For instance, in Fig. R2 we show the partial dependence plot for climatic water balance in the Beni ecosystem.

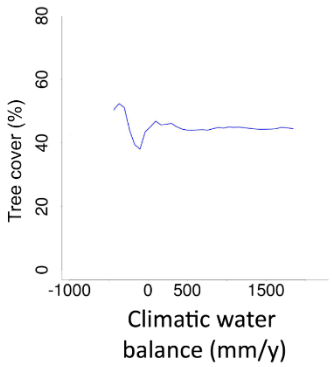

Figure R1. Partial dependence plot illustrating the effect of climatic water balance on tree cover in the Beni floodplains. The partial dependence plot is derived from the random forest regression and shows how the model-predicted value of tree cover changes across levels of climatic water balance, correcting for the effects of all other predictors.

Which of the inflection points illustrate real non-linearity and non-monotonicity? Without confidence intervals, it is not possible to say. We have chosen the more parsimonious linear models for inference regarding directional effects, after using random forests to determine variable importance.

We are also uncertain as to why the reviewer suggests that "probably, only the climatic water balance and sand content are quasi-linear," and why he/she feels that linear regression results are "qualitative."

For all these reasons we do prefer to keep our linear regression results in the manuscript, rather than to switch partial dependence plots.

(3) Currently, the discussion is organized by different environmental factors. I feel an overarching paragraph to summarize how the environmental determinants in floodplain differ from those in upland forests would be helpful.
As described in our reply to this reviewer's comment (1) and in our reply to the editor's first comment, we have now written such a paragraph (lines 278–284), which we hope will help satisfy the reviewer.

See below for minor comments following the order of the text:

Line 29-31: I wonder how much this effect can be, given that the flood duration effect is rather small (Figure 3). Decreasing flood duration from 12 months to 0 months might only induce a few percentage changes in tree cover.
The reviewer is referring to our statement at the end of the abstract about the potential impacts of dam construction, and we do not dispute that this is a fair point. It is perhaps not best to end the abstract with a statement about the applied importance of one of the weaker effects we found [though we do note that is not a negligible effect: moving from 12 to 0 months flooded in a fairly wet ecosystem (top bar in Fig. 3's flood duration panel)] would yield an approximately 15% increase in tree cover. In any event, we have completely rewritten the abstract and in place of the original statement, which was exclusively about dam impacts due to changed flood regimes, we now include one that covers dams as well as other relevant anthropogenic effects on floodplains and their tree cover: *"While outright wetland conversion proceeds globally, floodplain tree cover faces other anthropogenic threats, like changing fire frequency and dam construction, especially in wet climates."*

L. 100-102: I appreciate the efforts to combine high-resolution spatial data. My question is that do all data products have an original resolution of 90m or higher? To me, it makes more sense to aggregate (do the analysis at the coarsest resolution) instead of downscale (unless we know pretty sure how to represent fine-scale heterogeneity). At least, the soil data and MODIS fire data are at 250. How was downscaling performed? Why not just conduct the analysis at 250m?
Yes, the reviewer is correct that the MODIS fire data and soil composition (% sand) data were sourced at coarser resolutions (500m and 250m) than the other spatial data we employed. For clarity, we give below a summary of the resolutions of each included dataset, then answer the reviewer's question about the scale of our analysis.

Summary of original spatial resolutions:

Data	Resolution
Tree cover from Hansen et al.	30m
Flood regime from GIEMS-D3	90m
Climatic water balance from WorldClim	90m
Elevation from NASA Shuttle Radar Tomography	90m
Soil composition from Soilgrids	250m
MODIS burned area	500m
Distance to nearest river and large river from Hydrosheds	Vector

As described in our original methods section (now on lines 109–111), we used the 90m resolution of the GIEMS-D3 data as the spatial grain of our analysis. This means that when extracting tree cover and covariates, regardless of the original resolution of a dataset, we always used the value of cell in which the centroid of the relevant GIEMS-D3 cell fell. This means that only one of 9 tree cover cells falling within the larger GIEMS-D3 cells was used, and that values from each (coarser) soil and fire cell contributed to up to 9 and 25 observations in our dataset. In practice, these values tended to be much lower because contiguous 90m grid cells were not all included in the dataset, either because they were not all floodplains or because they were anthropogenically altered (urban, livestock, cropland, or deforested).

We acknowledge that this setup does lead to a certain level of pseudoreplication in our dataset where a single fire frequency or soil composition value is used for more than one observation. However, for three reasons we feel this is still the best approach:

(1) First, the simplest reason, is that we have very large sample sizes (tens-of-thousands to millions per floodplain region; Table S1) and so are not concerned about introducing spurious, marginal effects as falsely significant.

(2) While there is sometimes a convention in GIS studies to upscale all data to the coarsest resolution, we feel this is not warranted in every case, and specifically not in our case. Part of the motivation for this study was the ability to utilize the new GIEMS-D3 product which was deliberately downscaled to create a temporally variable inundation dataset at higher spatial resolution than previously available global products. As both reviewer 1 and the associate editor have noted, part of the novelty of this study is the ability to maintain relevance of the analyses within ecosystems while making cross-continental contrasts.

As a comparison, we give the example of a fine scale study of e.g., the drivers of deforestation using 90m pixels and a range of predictors (say human population density, accessibility, elevation, slope, and national GDP). Certainly, the goal is to understand local-scale deforestation, and all predictors save for national GDP might be available at 90m resolution. But we would not coarsen the grain of all data to the national level; rather, we would accept some level of pseudoreplication in GDP to avoid losing the value of high resolution predictors.
In our case, the level of pseudoreplication of soil and fire is lower than that in many studies that combine local data with regional or national ones [e.g., 15,16].

(3) Finally, if we were to coarsen the grain of the GIEMS-D3 data to match the 500m fire frequency data, we would lose the clean, binary nature (flooded or not) of the dataset and be left with a proportion of each 500m grid cell that was inundated each month. This would require calculating flood regime metrics (frequency, duration, variance in duration, and seasonal variability of flooding), based on a the less intuitive proportional area flooded within a 500m pixel, not the current and simpler flood status—yes or no—of a single 90m grid cell.

L. 134-136: It is OK to use this self-defined metric for variability although the metric would be the same between the scenario of a flooding every alternate year and that of continuous flooding for the first half and no flooding for the second half... I want to comment a little it on the computational efficiency. Even the Colwell's predictability or other entropy-based seasonality indices might not be so computationally expensive if the codes use some simple optimization for matrix manipulation. For example, just use the numpy package in python can significantly reduce computational time for spatial analysis from my own experience.
Yes, we acknowledge our metric for variability of flood seasonality does not take into account the ordering of flooded and not-flooded years. Neither do any of Colwell's predictability, constancy, and

contingency metrics, although that has little bearing on our paper. Reviewer 2 raised a similar question about fluctuating drought and flooding, and we have added a sentence in the discussion (lines 392–394) acknowledging the temporal sequence of flood and drought plays a role in regulating floodplain tree community dynamics.

While interesting, assessing the effects of the temporal sequence of flooded and non-flooded years would require a separate study, and indeed is one we will consider for future exploration. Measuring subtle behavior in the yearly sequences is not an easy task, requiring the right diagnostic(s) that would capture periodicities at the correct temporal scale to explain the most variance in tree cover.

Finally, yes, there probably are more efficient ways to calculate the original Colwell metrics than what our R code does. However, our own metric for variability of flood seasonality is near-perfectly correlated with Colwell predictability, as stated in the original manuscript for the Gorongosa and Everglades regions (r = -0.992 and -0.985, respectively; now line 148). We have also now confirmed that for the third smallest region, the Okavango Delta, this correlation is also near-perfect (r = -0.972).

L. 268-270: 'However, this effect saturated in the wettest regions'. Is this really unexpected? The previous analysis also shows the saturated effect of water balance in upland tropical forests.
We did not intend to imply that the saturating effect was particularly surprising, although with this feedback from the reviewer, we can see how this section may have been interpreted that way. We have reworded the paragraph to better reflect the similarities and differences between our findings and those from the literature on upland tree cover.

Finally, we note that this question is similar to one from Reviewer 2 who asked whether the saturation of the climatic water balance effect in wetter ecosystems might be accounted for by tree cover approaching its maximum, 100%. We note in our reply to him/her below that there is actually quite a bit of space for tree cover to increase in most of the wetter ecosystems studied. In addition, if the saturating effect were due to tree cover approaching its upper limit, we would expect other effects to saturate, as well, which we do not observe (e.g., elevation effects are *stronger* in wetter ecosystems).

Referee: 2

Comments to the Author(s)
This is an important contribution, and I am unaware of any similar study that has attempted to examine drivers of flood plain tree cover at such a large geographic scale. This is a nice complement to global studies of tropical tree cover that have ignored flooding as a contributing factor, and I expect it to be widely cited. Most of my comments are minor, but I have a couple of concerns that require attention.
We are glad to hear the reviewer agrees that this is a worthwhile contribution to the literature on tree cover and biome determinants.

Humans are not mentioned in this study at all. Isn't agriculture important in the Nile flood plain, for example? Is direct human cutting of forests important in any of the regions? The authors need to address whether direct human influence is of sufficient scale to influence the relationships, and if so how this may influence conclusions.
Because the primary goal of our study was to characterize the natural drivers of vegetation patterns in natural ecosystems, we did not consider areas that were heavily impacted by humans. In the original

version of the supplementary information, we described the procedures used to remove croplands, areas with high livestock densities, urban areas, and deforested areas. In the main text, we originally only wrote that we studied areas of "natural vegetation." To clarify that our study does not include regions with intensive human modification, we have added a short statement to this effect at the start of the main text methods section (lines 106–107) and slightly amended the caption to the map in Figure 1.

That said, the reviewer is correct that there are intensive human impacts in many of our study regions, including agriculture in the Nile floodplain, deforestation in the Amazon basin, and dam and other water control structures in the Everglades. While the reviewer wrote that we did not mention humans "at all," our original abstract, introduction, and discussion mentioned dam construction, climate change, and human modification of fire regimes as important disturbances whose impacts are informed by our study.

As stated above in a reply to Review 1, we also believe that focusing on the natural processes is more interesting. Deforestation, e.g., removes tree regardless of whether it occurs in a floodplain or in an upland ecosystem.

The authors find fire to be a strong predictor of tree cover, and interprets its causality as being entirely unidirectional (i.e. frequent fire reduces tree cover and this effect is stronger in wetter regions). I would argue that causality in the opposite direction dominates the relationship (areas with low tree cover are more flammable, and this effect is stronger in wet areas). This being said, I believe that fire probably does have a stronger impacts in wetter areas, but it will come across as disingenuous if you simply assume the direction of causality and do not acknowledge the complexity here.
This is a very fair criticism. Whereas in the introduction, we discussed fire and its *association* with lower tree cover, our original discussion did suggest unidirectional causality. We are aware of the fire-vegetation feedbacks that are well-characterized in upland ecosystems, and agree with the reviewer that these should occur in floodplains, too. As such, we have edited our discussion of fire. On lines 298–299 we now write that "frequent fire *was more strongly associated with lower tree cover* in wetter climates" where we had previously written that "fire *decreases* tree cover more strongly in wetter climates." In addition, at the end of that paragraph we added:

> *"Fire-vegetation feedbacks are well-known from upland savannas [5], meaning that not only do fires create ecosystems more dominated by herbaceous plants, but also that more herbaceous ecosystems tend to have higher fuel loads, promoting fire frequency and intensity. The same feedbacks are likely in floodplains.*" (Lines 306–310)

Unless I missed it, the authors did not mention that it may be particularly challenging for trees to cope with alternating flooding and water deficit. There seem to be relatively few species that can cope with this, such as Melaleuca in Australia, which interestingly has become invasive in the Everglades. Some palms seem particularly good at this as well. There may be good reason to expect alternating anoxia and drought to be a particularly difficult situation to deal with because, for example, coping with waterlogged soil generally requires shallow roots, while coping with water deficit generally requires deep roots. I cannot think of a specific reference that deals with this specific constraint, so I would not insist that the authors discuss this, but it does seem to provide context for many of the results.
This is a good point. Absolutely, even floodplains can be water-limited at times. This question is also similar to Reviewer 1's question about alternating flood and drought. As mentioned in our reply to their query, we have added a sentence in the discussion (along with 2 relevant references) acknowledging

that the temporal fluctuations between drought and flood can affect floodplain tree recruitment and survival.

We also now present in our reply to Reviewer 1's related comment a supplementary analysis of the correlation between floodplain tree cover and the number of fluctuations between years with and without flooding.

Line 64. Varzea is spelled with an accent over the first 'a'. Also, I would specify 'varzea forest' here, since the term varzea also includes grassy areas.
Done and fixed!

Line 160-163. Check this explanation. Doesn't a greater MSE indicate a poorer fit and lower variable importance?
In fact, the original wording is correct, but the reviewer's comment reveals that our explanation was unclear. This procedure is standard in random forest regression, and it is described in the two references given. To quantify variable importance in the random forest models, we compared predicted values using "correct," original predictor variables and a set of permuted predictors (mismatching the predictors to the tree cover values). The procedure was done for each predictor individually. When permuting more important predictors, model-predicted tree cover will be farther from the actual tree cover values (increasing mean squared error) than when permuting less important predictors. We have edited the wording (lines 170–177) to try to ensure clarity on this procedure. This section now reads:

> *Specifically, variable importance is assessed by permuting predictor values from a subset of the data held out of the model training set, then comparing the fit of model-predicted tree cover to actual tree cover values using the randomly permuted versus original predictor values. We completed this procedure for each predictor and computed the average increase in mean squared error of predicted versus actual tree cover across regression trees [48,51]. Larger values indicate a greater variable importance, because they show that permuting a given predictor's values reduces fit to a greater degree.*

Line 185. Remove hyphen in 'water-balance'.
Thanks. Done!

Line 270. I don't understand how table 1 shows this saturation in wetter regions. Figure 3 does a better job of convincing me of this, but also suggests a more likely explanation for the saturation, i.e. tree cover saturates at high water balance because it is impossible to have more than 100% tree cover. Since tree cover inherently saturates at 100% it seems unreasonable to seek a mechanism to explain this, and I suggest not highlighting it by devoting the first paragraph to this point.
The call out to Table 1 is because three of the four wettest regions (Llanos, Tonle Sap, and Beni) have negative relationships between tree cover and water balance. In contrast, all of the seven driest regions have positive relationships. We have slightly modified the text to make this clearer (line 285–286). We also now point to figures S3–S13, which show the effect of climatic water balance is weaker (or even negative) in the wetter ecosystems.

Regarding the possibility that saturation of effects in wetter ecosystems is just because tree cover is approaching 100% (its upper limit), we present two lines of evidence that suggest this is not the case. First, as shown in our original Fig. S1 (pasted below) there is quite a bit of room for upward shifts in tree cover in all the wet regions except the Amazon (e.g., Llanos, Tonle Sap, Beni, Everglades).

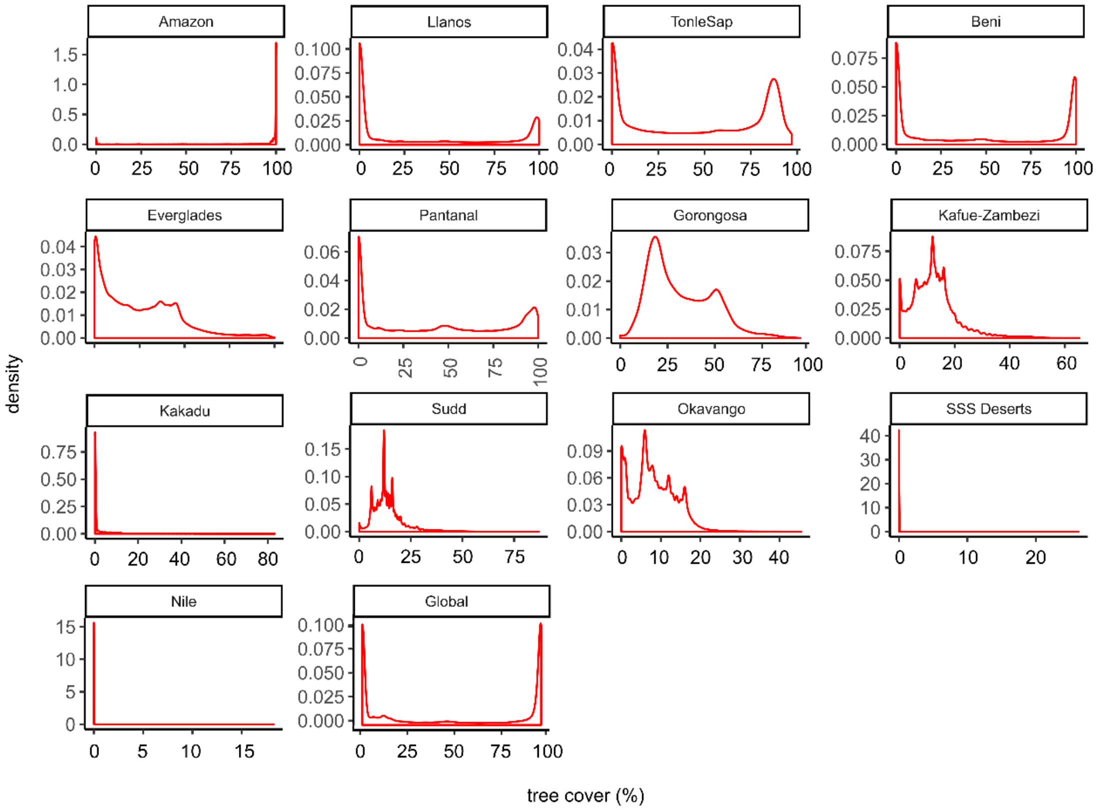

Fig. S1. Density plots showing the frequency distribution of tree cover across the 13 studied floodplain regions and the overall distribution for all regions combined (global). Regions are ordered from climatically wettest to driest.

Second, if the saturation of the climatic water balance effect were simply due to tree cover approaching 100%, we would expect similar saturation of the effects of other predictors with positive effects on tree cover. In the case of elevation, we actually see a stronger effect in wetter ecosystems (Fig. 3, top right & magnitude of parameters in the elevation row of Table 1).

Line 344-354. This paragraph was a little confusing. I think it would help to make it clear early in the paragraph that, in upland savanna, sandy soils are associated with high tree cover. Many readers will assume the opposite, which makes later statements confusing. Also, clarify why would including forests would change the result (line 353).

In line with the reviewer's helpful suggestions, we have now clarified the reasoning in this section. We restate the known effect of sandier soils in upland savannas, then explain that forests tend to occur in wetter areas than grassland and savanna, and that we therefore we do not expect sandier soil to confer the same water-competition-mediated benefits for trees as in more arid ecosystems. Further, the near absence of grass in forests means that the potential for root depth differentiation between trees and grasses is possibly irrelevant there. Thus, sandier soil, which allows deep percolation of water is only advantageous for trees in the more arid ecosystems and including forests in our analysis may explain why we don't find the same relationship between soil composition and tree cover as in upland studies of grasslands and savannas. The full paragraph text (from lines 395–411) is pasted below:

"Higher soil sand content was associated with lower tree cover at the global scale, although this effect was idiosyncratic within individual floodplain ecosystems. This contrasts with its effects in

upland savannas, where sandier soils tend to increase tree cover by promoting water infiltration and therefore root niche separation of trees and grasses. Infiltration depth may be less important in floodplains because wetland water tables are generally shallower, precluding the possibility of root depth differentiation between grasses and trees. Rather, our result that sandier soil decreased tree cover suggests that negative effects of high sand content on plants due to reduced nutrient availability may predominate in floodplains. In addition, whereas previous studies of upland ecosystems have focused on savannas and grasslands, we also included forested floodplains, which are less likely to be water-limited and which have little-if-any grass, again reducing the expected influence of any mechanism reliant on competition for water."

Line 373. Change 'shrub' to 'tree'? Both are true, but this study focuses on trees.
Yes, we don't disagree that it may be better not to use the term "shrub" here. We originally used this term because the phenomenon of woody plants increasing in many grassland ecosystems globally is often referred to as "shrub encroachment." Given that the reviewer is correct that our study is more directly regarding trees, we have opted for the somewhat intermediate (and still commonly used) term, "woody encroachment," which we feel easily includes trees, without abandoning the clear parallels to true shrub encroachment.

Appendix B

Dear Dr. Kruuk,

Thank you for writing September 11, 2019 regarding our manuscript *Determinants of tree cover in tropical floodplains* (ID RSPB-2019-1755). We were glad to see that you and the reviewers believe we have addressed most comments in our last resubmission and we welcome the additional few comments from Referee 3, which we found constructive. We have now replied to each of the reviewer's new concerns and made corresponding edits to the paper. Our replies and edits are detailed below with line numbers that refer to the tracked-changes version of our revised manuscript.

Along with the tracked changes version of the manuscript, we have attached a clean version, separate figure files, and the supplementary information as a separate document. We have additionally confirmed that our manuscript is estimated to fit into 10 printed pages using the journal's Page Estimator tool. We hope you will find our edits are sufficient for acceptance of the manuscript and will look out to hear back from you again soon.

Sincerely,
Joshua Daskin, Ph.D.

On behalf of A. Carla Staver, Ph.D. and Felipe Aires, Ph.D.

Referee: 3

Comments to the Author(s).
I now read through the manuscript, supplementary information, and the rebuttal letter. It is a well-done research and the authors have addressed most of the concerns raised by the referees. I recommend accept after minor revisions.
We appreciate the reviewer's kind assessment of our research and our last round of edits. Below we give detailed replies to the reviewer's latest round of comments.

First, water balance is computed as rainfall minus potential evapotranspiration; but scientifically, it should be rainfall minus actual evapotranspiration. Potential evapotranspiration does not represent actual evapotranspiration. Please use actual evapotranspiration or explain the reason.
This is an interesting comment, and one we are glad the reviewer has asked us to address. Potential evapotranspiration (PET) is a measure of the evaporative demand of the local climate, heavily influenced by solar irradiation and wind. In our experience, using PET is the standard for ecologists and climatologists determining moisture availability. For example, the U.S. National Oceanographic and Atmospheric Administration calculates the oft-used Palmer Drought Severity Index using precipitation and PET (not actual evapotranspiration, AET; https://www.ncdc.noaa.gov/monitoring-references/dyk/potential-evapotranspiration) and another common drought index, the Standardized Precipitation and Evapotranspiration Index (Vicente-Serrano *et al.* 2013) does the same. Similarly, we originally cited two papers in our supplementary materials that used the same formulation we did to compute climatic water balance (precipitation minus PET; Lehman et al. 2011 & 2014, they called this quantity "effective rainfall"). In addition, a wide range of recent studies in ecology, ecohydrology, and climatology use the same characterization of moisture availability. A selected subset of such references is given here:

Dietrich O, Kaiser T. 2017 Impact of groundwater regimes on water balance components of a site with a shallow water table. *Ecohydrology* **10**, e1867.

Konings AG, Williams AP, Gentine P. 2017 Sensitivity of grassland productivity to aridity controlled by stomatal and xylem regulation. *Nature Geoscience* **10**, 284.

Levesque M, Andreu-Hayles L, Pederson N. 2017 Water availability drives gas exchange and growth of trees in northeastern US, not elevated CO 2 and reduced acid deposition. *Scientific reports* **7**, 46158.

Mitchell PJ *et al.* 2016 An ecoclimatic framework for evaluating the resilience of vegetation to water deficit. *Global Change Biology* **22**, 1677–1689.

Vicente-Serrano SM *et al.* 2013 Response of vegetation to drought time-scales across global land biomes. *PNAS* **110**, 52–57. (doi:10.1073/pnas.1207068110)

Vitousek PM, Paulus EL, Chadwick OA. 2019 Nitrogen dynamics along a climate gradient on geologically old substrate, Kaua'i, Hawai'i. *Oecologia* **189**, 211–219. (doi:10.1007/s00442-018-4285-1)

Of course, the fact that PET is commonly employed does not alone make it a better option than AET. Rather, the static nature of the PET compared to AET (which varies year-to-year with precipitation and other weather events) is more conducive to our analysis of proportional tree cover across the world's tropical floodplains. Had we used an average of several years' AET, our calculation of climatic water balance could have been biased by irregularly high- or low-AET years. PET provides a consistent and static measure of climatic availability (or demand) across the globe. This, combined with our desire to follow the widely employed standard, led us to use PET. We therefore have not changed our calculation of climatic water balance.

Second, the percentage tree cover from Hansen et al., 2013 is for the year of 2000 rather than 2010 as mentioned by the authors (Line 112). If the authors used percentage tree cover for 2000 minus forest loss during 2000-2010, they may neglect the forest gain during 2000-2010. But Hansen et al 2013 had not provided forest gain during 2000-2010.
The reviewer is correct that the main site for distribution of Hansen et al.'s tree cover, deforestation, and forest gain data does not include a 2010 tree cover product, only one with a year-2000 tree cover estimate. However, the tree cover values we used in this study are in fact for the year 2010 and are from the Hansen et al. 2013 study. These data are available here: https://www.glad.umd.edu/dataset/global-2010-tree-cover-30-m. This product is produced from Landsat data at 30m resolution (as for the data distributed at the primary Global Forest Change website) and gives a pixel-wise estimate of tree cover (1-100%) in integer values. We have updated the citation (#45 in the main text) and now include the dataset URL to better-direct readers to the appropriate dataset. The citation now reads as below:

Hansen MC et al. 2013 Data from: High-resolution global maps of 21st-century forest cover change. *Science* 342, 850–853. https://www.glad.umd.edu/dataset/global-2010-tree-cover-30-m.

Third, I find it difficult to understand the method without reading "Delineation of study regions" in Supplementary Information. I recommend moving this material to the main text (combined to Study regions and data extraction in Methods).
As suggested, we moved most of this section into the main text, now on lines 100-127. Some details were left in the supplement to keep the manuscript length below the required 10-page limit.

Lines 218-220, I do not observe Northern Nile in Table S2.
The results for this region were present in Table S2, however we have updated the column header for this study region in Table S2 from "Nile" to "Northern Nile" and apologize if this difference in naming may have confused the reviewer.

Lines 225, 231 and more, there is no Table S3 in Supplementary Information.
Table S3 is found on the final two pages of the Supplementary Information. We do notice that on the

cover page of the supplementary information we had incorrectly listed the contents as including only two supplementary tables, though. We have now fixed this error.

Lines 434-436, it is unclear how the authors get the number of tree cover loss caused by dams. Which dams and how they calculate?
We estimated the likely decrease in tree cover for a location that goes from 0 to 12 months flooded—as may happen in the case of a dam impoundment—from the slope and intercept of the regression line on the conditional regression plot for flood duration in the climatically wettest regions (Fig. 3, green line). This line shows tree cover decreasing from ~90% to ~75%. We have updated the text to clarify this.

Figure 3(b), it is difficult to understand why model-predicted tree cover increases with burned times for dry climate (red line).
We thank the reviewer for this good question; indeed, it is an intriguing result. Given that the positive effect is rather small—a <5% increase in tree cover in the climatically driest ecosystems along the gradient from never burned to >10 times burned—we believe this is indicative of the strongly fire-adapted nature of savanna and grassland plants and not of an actual positive effect of fire on tree cover, *per se*. We alluded to this in our discussion section, although we focused on the maladaptation of wet-climate trees to fire, since those are the ecosystems where the fire frequency effect was greatest:

> We found that frequent fire was more strongly associated with lower tree cover in wetter climates, mirroring upland ecosystems where fire also has the biggest effects on woody plant abundance at higher-rainfall sites [2]. Trees tend to be poorly adapted to fire in moist tropical forests [33] where fire is relatively rare [53], meaning trees are more likely to die when fires do occur there. In addition, wetter climates encourage higher fuel loads and fire intensity in savannas [67]. Indeed, fire frequency had stronger negative effects on tree cover in the Pantanal and Gorongosa, floodplain savannas with relatively wet climates compared to the Okavango and Sudd, which are relatively dry floodplain savannas and where fire frequency had little impact on tree cover (Figs. S6, S10–S12).

We have now added a sentence to emphasize that dry-climate trees are well-adapted to frequent fire (lines 311–313): "In contrast, many trees in more arid climates survive regular, low-intensity fires [19,33], accounting for the relative lack of an effect of fire in climatically drier floodplains (Fig. 3)." We also removed the reference in the Results section to a slight positive effect of fire in dry ecosystems, opting instead to describe it as a null effect, which better describes our finding (lines 258–259).