

Reviewer 1

We thank reviewer 1 for the detailed and thorough comments. Our replies to the comments are inserted below *in blue colour*.

Reviewer 1

Review: Improving the LPJmL4-SPITFIRE vegetation-fire model for South America using satellite data

General comments

Process-based global fire models are widely considered critical components of dynamic global vegetation models, with certain biomes—especially tropical and subtropical savannas and grasslands—being strongly regulated by fire disturbance. However, many such fire models have been developed based on parameterizations from extratropical biomes. In this manuscript, Drücke et al. use an automated technique to reparameterize a global fire model to improve its performance in the Brazilian Caatinga and Cerrado biomes with regard to both burned area and biomass. The authors perform this optimization using actual runs of the vegetation model rather than in some kind of offline mode—something that has only rarely been done before for fire models, but which could be a valuable component of the global vegetation-fire modeling toolbox. This, combined with the fact that the authors describe their methods thoroughly and walk through the results in a logical manner, lead me to recommend that this manuscript be accepted for publication pending minor revisions.

We thank the Reviewer for this positive feedback.

Specific comments

•P2 L13: The authors describe the Caatinga as fire-prone, but the referenced map (Fig. 1) does not provide much support for that assertion. The authors should clarify in the text what they mean.

We agree that the term fire-prone is not strictly applicable for the Caatinga region. Without human influence on vegetation composition and fire, the Caatinga biome is not prone to fire. Fire risk has increased with human influence. We wanted to stress the fact, that fire occurrence of the Caatinga in the prior SPITFIRE implementation was too large, because dead and live fuel was overestimated in the model. It is one of the achievements of this study to reduce this model error. We now make a better distinction between fire-ecological conditions of the biome and what is captured in the model. We therefore changed the wording accordingly to only describe the Cerrado as fire prone in P2 L12-13:

“This study focuses on the fire behavior in central-northern South America and especially on the Brazilian biomes Caatinga and Cerrado, which is the most fire-prone region in South America (Fig. 1)”

•P6 L1: “their effectiveness to ignite a fire is 0.04” is unclear. Better would be something like, “4% of cloud-to-ground strikes can start a fire.”

We followed the Reviewer’s suggestion and express these numbers now in percent in P6 L1:

“assuming that 20 % of the flashes reach the ground and 4 % of cloud-to-ground strikes can start a fire.”

•P6 L9:

As far as I can tell, this is the first time the parameter named p_d " has been used in regard to SPITFIRE. I suggest using some other symbol, as this p_d " could be easily confused with P_d (population density) from Thonicke et al. (2010).

Thanks for pointing to this inconsistency in naming our model variables. We have corrected it and changed it to p_t , where the t stands now for time and not for the duration (d).

"per day" is misleading; SPITFIRE as described in Thonicke et al. (2010) does not allow for multi-day fires, and thus this is simply the maximum fire duration.

Indeed, SPITFIRE does not directly allow multi-day fires. The maximum fire duration is 240 minutes. Since the model time-step is however one day this time duration limits the fire to 240 minutes for each time-step. SPITFIRE does not model individual fires, but assumes the burning conditions for all fires ignited on the same day to be similar. If burning conditions are comparable the next day, new ignitions would be computed resulting in comparable daily area burnt. Hence, to our opinion the wording in this case is correct.

•P11 L1–2, P21 L21–27: Because LPJmL does not allow fire on managed lands, the authors exclude cropland burning from the observed data in their comparisons. This is reasonable, but ignores the fact that a fair amount of Cerrado is actually used as pasture, primarily in the southern part of the region (Sano et al., 2010; Parente et al., 2017). I don't think this makes a huge difference in the context of this manuscript, because (a) only a few of the 40 sampled grid cells were from the southern Cerrado, and (b) the main takeaway from this paper should be the use of the optimization algorithm, rather than the exact parameter values it gives. However, the authors should (briefly) address this issue in the text.

We acknowledge the fact that excluding cropland in the optimization process is not optimal. Unfortunately we did not have a dataset, which also excludes pastures. This leads to a slightly wrong burnt area in the optimization process. We already addressed this issue in P24 L26–29:

"We do, however, not account for the proportion of grassland, which is used as pastures for e.g. cattle ranching. Since in SPITFIRE fire is not enabled on managed grassland, our results show a slightly smaller fire amplitude throughout South America than could be expected with pastures included and hence also compared to the GFED4 evaluation data set."

We changed in the above paragraph the term "grassland" to "pasture" and included a sentence about this issue in P11 L32–34:

"As LPJmL does not simulate fire on managed lands, we excluded burnt area on cropland classes from model-data comparisons. Due to lack of data we however did not account for the proportion of pastures."

•Sect. 2.4 and/ or Sect. 3.2: For the benefit of other researchers interested in using this or a similar optimization algorithm, it would be helpful to know various pieces of info about the process. How many model runs were required? How long did they each take? How was the decision made to halt the optimization—was it manual, or did the algorithm reach a stop condition? If the latter, what was/were the stop condition(s)? Etc. This level of technical detail is more than appropriate for GMD.

We agree that some more technical details are important to the GMD reader and we therefore

added the following paragraph in Sect. 2.4 (P11 L5-17):

“In genetic optimization algorithms, each model parameter is called an individual with a corresponding fitness, which represents the cost of the model against the observations. At the beginning of the optimization process, the GENOUD algorithm creates a generation of individuals based on random sampling of parameter sets within the prescribed parameter ranges. After the calculation of the cost of all individuals of the first generation, a next generation is generated by cloning the best individuals, by mutating the genes or by crossing different individuals (Mebane and Sekhon, 2011). This results after some generations in a set of individuals with highest fitness, i.e. parameter sets with minimized cost. To find an optimum parameter set we were also using the BFGS gradient search algorithm (named after the authors Broyden, 1970; Fletcher, 1970; Goldfarb, 1970; Shanno, 1970) within the GENOUD algorithm. An optimized parameter set of the BFGS algorithm is used as individual in the next generation. We were applying the GENOUD algorithm with 20 generations and a population size of 800 individuals per generation, which corresponds to 16000 single model runs. We decided on this amount of iterations, because the cost kept almost constant in the last iterations and the parameter values did not change to the 6th digit, beyond which changes are not really relevant for model applications. During the optimization we ran the model parallel in for each gridcell (40 grid cells and CPU's, 3.2GHz) and had a total optimization time of ca. 24 hours.”

•Sect. 2.6: The authors do a good job describing how to interpret values of the NMSE, but they should do the same for the Willmott coefficient of agreement. What are the possible values? What are “milestone” values (e.g., for NMSE, 0 vs. 1 vs. >1)?

Thanks for pointing to this missing information. The information on the meaning of the the different values for the Willmott coefficient was indeed missing. We have now added the following sentence in P12 L17-20:

“The Willmott coefficient is a squared index, where a value of 1 stands for perfect agreement between simulated and modelled runs and gets smaller for worse agreements with a minimum of 0. Unlike the coefficient of determination, the Willmott coefficient is additionally sensitive to biases between simulations and observations.”

•P13 Fig. 5: It would be helpful to use the same Y-axis for all subplots in the right column (subplots b, d, f), as was done for the left column.

We also thought about this issue but had decided before to not use the same Y-axis for all subplots in the right column, because the values differ by an order of magnitude. Hence, the lines in Fig. 5 f would be hardly visible. We agree with the reviewer that different scales in this case are very unusual and we apply now a logarithmic Y-axis for the Caatinga. This enables us to use the same Y-axis for all three model versions. A non-logarithmic version remains in the Appendix. We now updated Fig. 5 and added the following statement to the description of the Figure:

„Note the logarithmic scale for the Caatinga, which was applied in order to account for the large differences between the different model versions (for a non-logarithmic version see Fig. A6).“

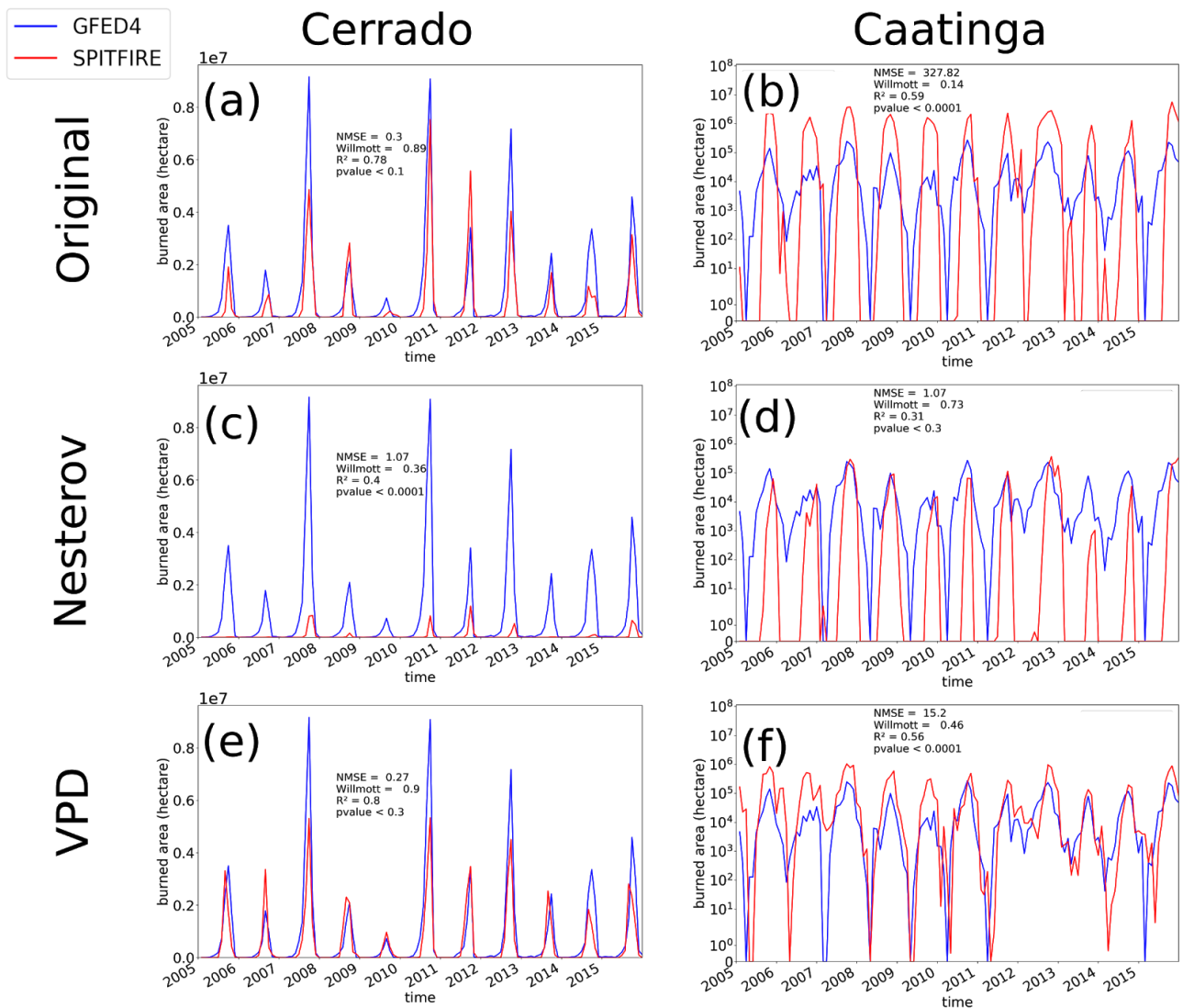


Fig. 5. Time-series of monthly burnt area from 2005 - 2015 simulated by SPITFIRE (red lines) compared to GFED4 evaluation data (blue lines) for: (a) The Cerrado region, using NI_orig. (b) The Caatinga region, using the NI_orig. (c) The Cerrado region, using NI_optim. (d) The Caatinga region, using NI_optim. (e) The Cerrado region, using VPD_optim. (f) The Caatinga region, using VPD_optim. Note the logarithmic scale for the Caatinga, which was applied in order to account for the large differences between the different model versions (for a non logarithmic version see Fig. A6).

•P13 Fig. 5 and P16 Fig. 6: Nesterov and VPD rows should be swapped, since in the rest of the paper the Nesterov Index is usually discussed first.

We swapped the Nesterov and the VPD rows, as suggested by the Reviewer.

•P16 Fig. 6: The use of lines here is confusing, since that usually implies some kind of change over time. The authors should seriously consider using a bar graph here instead.

We thank the reviewer for the suggestion to change Fig. 6.

We completely remade Fig. 6 by excluding the lines and changing the X- and Y-axis. We also now show each parameter individually. To our opinion the use of a bar plot would limit the information about the exact uncertainty for each parameter for the different PFTs. We hope that we have addressed the reviewer's concern adequately with this new figure.

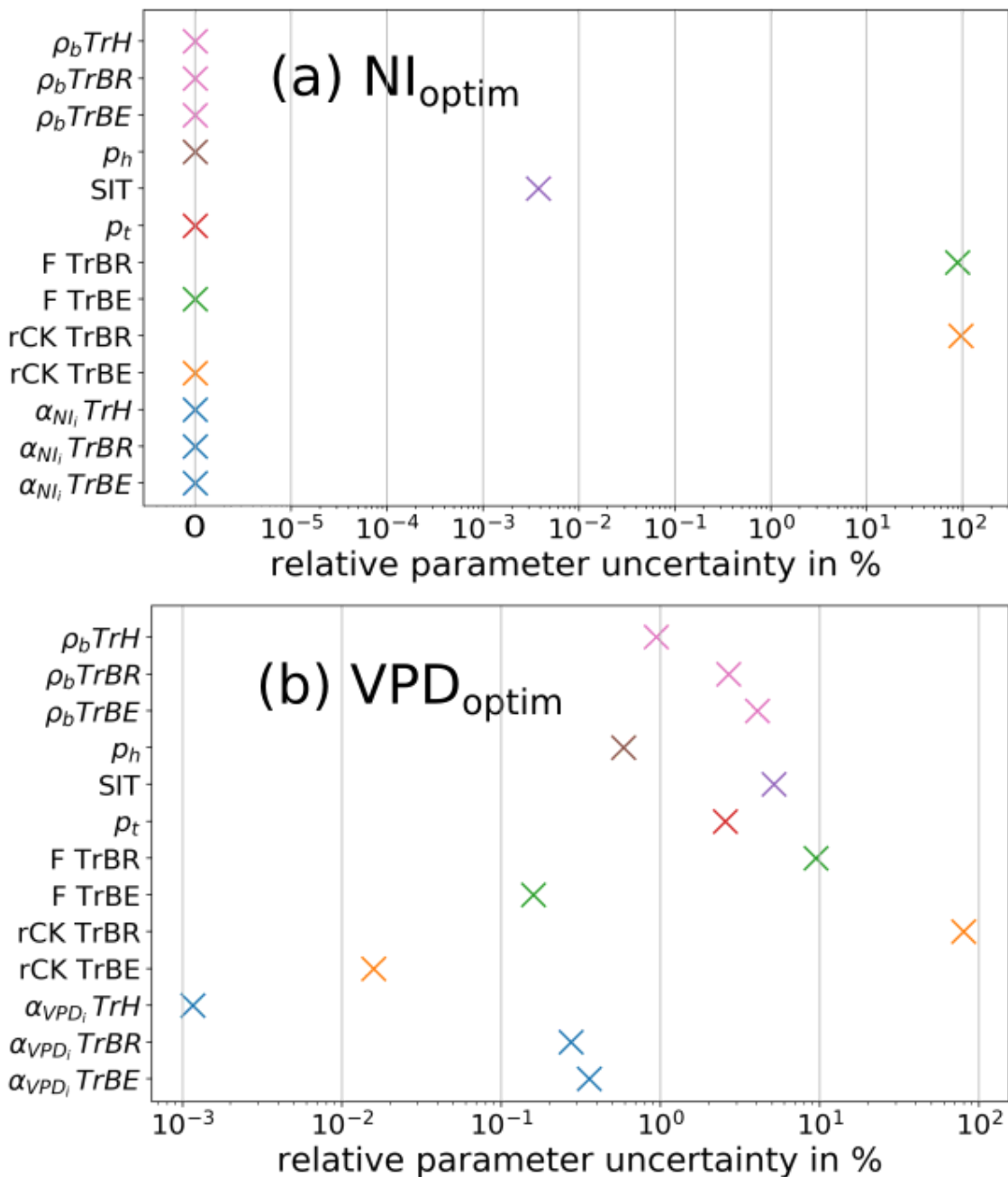


Fig. 6. Relative uncertainty of model parameters after optimization for (a) NI_{optim} and (b) VPD_{optim} . The relative uncertainty is the ratio of the uncertainty after the optimization (range of all parameter sets with low cost, below the 0.05 quantile) divided by the uncertainty before the optimization (range of the parameters for the optimization). Low and high values of relative uncertainty indicate strongly and weakly constrained parameters, respectively. SIT denotes the surface intensity threshold

•P17 L3–4: “The model optimization scheme focuses on fire parameter [sic], hence the model performance can only improve in fire-prone biomes, i.e. not in, e.g., wet tropical forest where fire is absent.” This is not strictly true. Model performance could improve in wet tropical forest if the initial parameterization (a) performed badly there

with regard to burned area (i.e., simulated almost any fire at all) and (b) underestimated biomass. It just so happens that neither of these conditions are met by the initial LPJmL configuration. This may seem like a minor quibble, but it could mislead other researchers interested in applying this or a similar optimization algorithm to their own models. It is important to be clear that optimizing a fire model can improve performance with regard to vegetation parameters not necessarily where fire is frequent, but rather where fire is modelled poorly.

We agree that this statement was a bit misleading. We wanted to emphasize that we did not get an improvement in AGB and FPC in areas not affected by fire, both in the model and in the evaluation data, and hence it does not contribute to the improvement due to the optimization process. We now clarify this issue by changing the quoted sentence into (P19 L5-10):

“The model optimization scheme focuses on fire parameters, hence the model performance for AGB can only improve in areas, where the fire occurrence has been modelled poorly and the vegetation-fire interactions have improved due to the optimization process. For example in the center of the Amazon rainforest almost no fire is found in the evaluation data nor is simulated. Hence no improvement of burned area as well as AGB can be achieved. On the other hand, in regions where the modelling error of burnt area is now reduced, this can also improve simulated AGB, hence vegetation-fire interactions.”

•P20 L7–8: Presumably the authors are making this assertion based on the fact that the indicated region is modelled as ~50% tropical evergreen, but Fig. 8 does not appear to say anything about the tropical raingreen PFT—just evergreen. The authors should clarify this.

We agree with the reviewer that readers not familiar with DGVMs or specifically LPJmL should be provided with such information for clarity. We new added the maps of the FPC of tropical raingreen and herbacious C4 PFTs to the supplement (Figure A1 and A2) and added a short description to the result section now for all 3 tropical PFTs in P20 L5-8:

‘Also for the TrBR and TrH PFT distributions the optimization lead to an improved performance using the VPD_optim in the Caatinga and Cerrado, whereas the PFT distribution in the Amazon remained similar to the prior PFT distribution. In the NI_optim version, parameter optimization only slightly reduced TrBR cover showing a worse performance compared to VPD_optim. However, herbaceous cover changed only slightly in all optimization experiments (Fig. A1 and A2).’

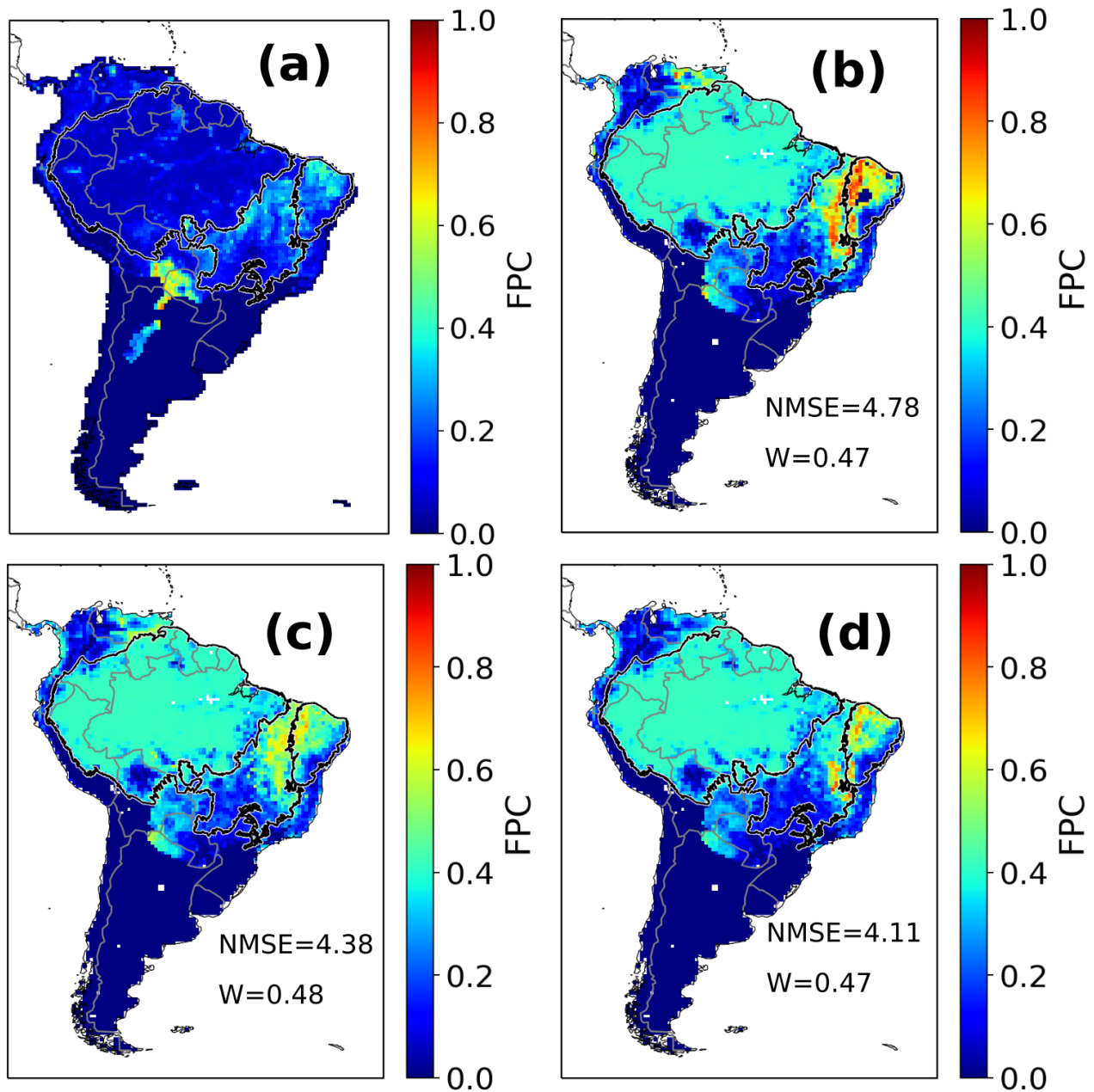


Fig. A1. Annual FPC cover by tropical broadleaved raingreen PFT over a mean from 2005-2015 as fraction per cell. (a) ESA-CCI evaluation data (b) Simulated FPC by LPJmL4-SPITFIRE using the NI_orig version (c) Simulated FPC by LPJmL4-SPITFIRE using the VPD_optim version (d) Simulated FPC by LPJmL4-SPITFIRE using the NI_optim version

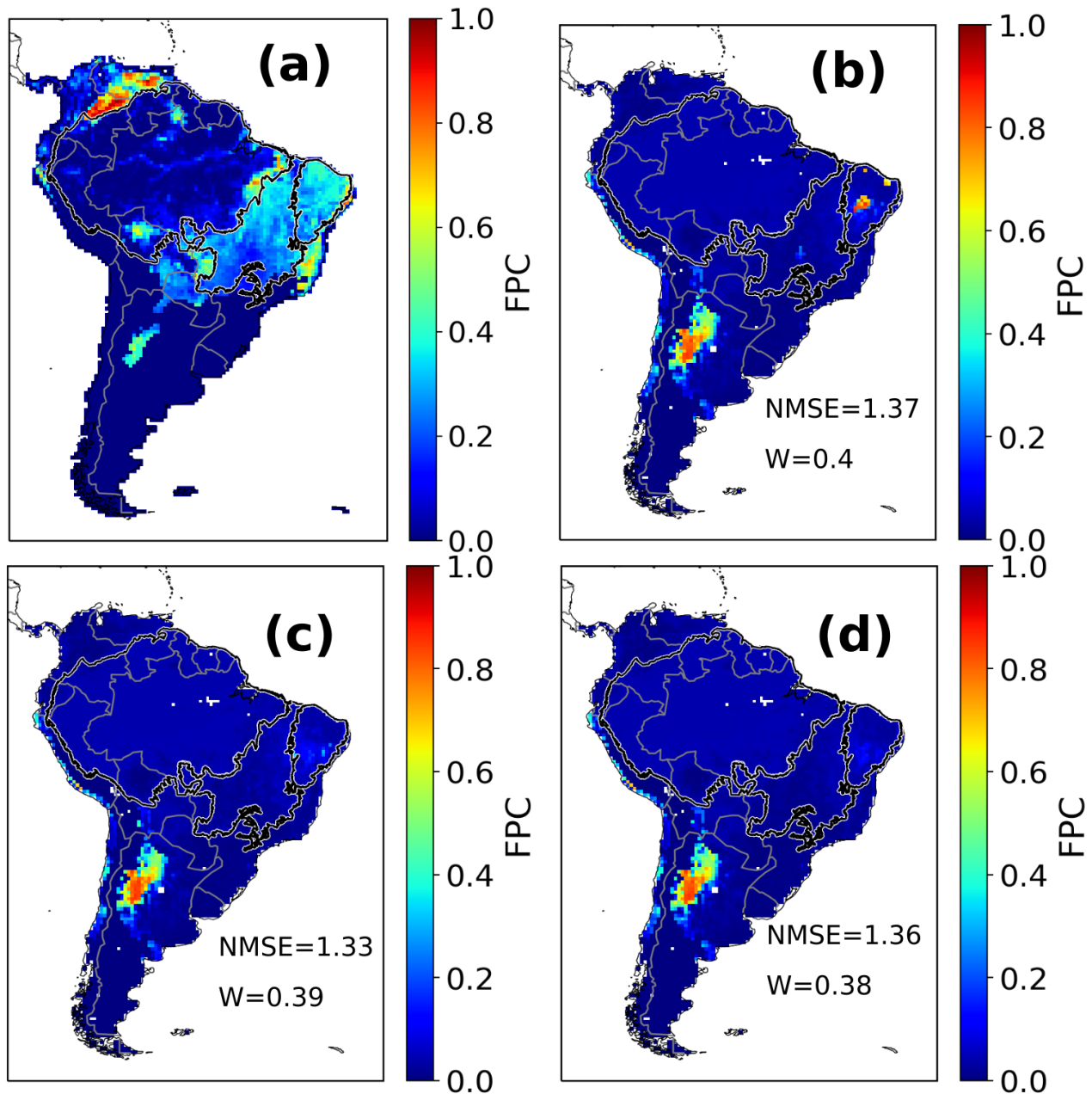


Fig. A2. Annual FPC cover by tropical herbaceous PFT over a mean from 2005-2015 as fraction per cell. (a) ESA-CCI evaluation data (b) Simulated FPC by LPJmL4-SPITFIRE using the NI_orig version (c) Simulated FPC by LPJmL4-SPITFIRE using the VPD_optim version (d) Simulated FPC by LPJmL4-SPITFIRE using the NI_optim version

•P21 L23–27:

How do the authors reach the conclusion that including fire on managed land would increase “fire amplitude” (this phrase should be reworked, by the way) and improve interannual variability? Why might it not also (or instead) improve annual mean?

Citations should be added regarding the real-life use of fire on managed lands (e.g., Laris, 2002). Citations should be added regarding the simulation of fire on managed lands (e.g., Pfeiffer et al., 2013; Rabin et al., 2018).

We agree that this statement is relatively vague. We meant to say that we would have a slightly larger burnt area in regions, which also include pastures. The larger burnt area would cause a

larger annual total (or mean) burnt area and also likely a larger seasonal amplitude (if there are no fires during the wet season). Also the overall model performance would improve and not just the interannual variability. However, we expect this effect to be relatively small, since only a few grid cells are covered by a substantial fraction of pasture land in the study area. . We thank the reviewer for the valuable suggestions, including citations, and rewrote this section as following (P24 L27-33):

“Since in SPITFIRE fire is not enabled on pasture, our results show a slightly smaller burnt area throughout South America than could be expected with managed land included and hence also compared to the GFED4 evaluation data set. This effect is however small, because pasture lands cover a substantial fraction only in very few grid cells (e.g. southern Cerrado; Parente et al., 2017). Fire on managed land is generally difficult to predict in a DGVM because the reason and timing of using fire depends less on climatic factors but mostly on social and political decisions which can vary between countries, regions and localities. We expect further improvement of model performance especially in regions of large land-use areas with fires on pastures included (e.g. Rabin et al., 2018; Pfeiffer et al., 2013).”

•P22 L1–2: This is incorrect; Rabin et al. (2018) did indeed optimize FINAL.1 within a dynamic global vegetation model (LM3). Also, the authors should (at least briefly) discuss the pros and cons of their method relative to the one used by Rabin et al. (2018); this would be valuable for other researchers interested in optimization methods.

We thank the reviewer for noting this error. Rabin et al. (2018) did indeed optimize FINAL.1 within a DGVM. We now refer to this study and explain the major differences to the approach used in our study. As suggested we added a paragraph about the pros and cons of our method relative to the one us used by Rabin et al. (2018) in the discussion (P23 L27 - P24 L8):

“Systematic optimizations within a model-data integration setup of fire models which are embedded in a DGVM are still very rare. Previously, Rabin et al. (2018) optimized the fire model FINAL.1 within the land-surface model LM3. Our study differs from Rabin et al. (2018) in the conceptual design of the vegetation-fire models and the optimization process. While LM3 has been run on a 2° longitude by 2.5° latitude, running LPJmL at 0.5° by 0.5° grid cell resolution allows us to account for spatial differences in climate, vegetation and fire interaction. While FINAL.1 is a process-based model, many calculations (e.g. the fire spread routine) are done by multiplying the important factors and fitting the resulting values to observational data. SPITFIRE tries to model the important fire variables by simulating the underlying processes, and by taking the influence of climate and the different fire ignitions into account. An advantage of FINAL.1 is the inclusion of agricultural fires based on a statistical approach. Whereas Rabin et al. (2018) used a local search algorithm (Levenberg-Marquardt algorithm) to optimize their model, we used a global search algorithm (genetic optimization). Local search algorithms depend on the chosen initial parameter sets and might eventually end up in a local optimum. A genetic optimization algorithm allows to explore the full parameter space and hence gives a higher chance to find the global optimum. However, local search algorithms require less iterations than global search algorithms (300 in Rabin et al. (2018) vs. 16000 in our study). Forkel et al. (2014) tested the optimization of LPJmL with different optimization algorithms and found that it was not feasible to optimize LPJmL with a local search algorithm. Rabin et al. (2018) ran the model during the optimization process only for the period of 1991-2009, whereas in our optimization setup we made complete model runs including 5000 years of spinup in LPJmL in order to get a model equilibrium for each tested parameter combination.”

•P22 L14–15: The authors address availability of the model code, which presumably refers to LPJmL. But what about the genetic optimization code?

We will publish the code of LPJmLmdi along with the model code on the github page of LPJmL: <https://github.com/PIK-LPJmL/LPJmLmdi>.

Technical corrections

- P9 L5: “form” should be “from”
- P10 L22: “dependend” should be “dependent”
- P11 L19: “simulations” should be “simulation”
- P11 L25: “Caating” should be “Caatinga”
- P17 Table 3: “Evergreem” should be “Evergreen”
- P17 L3: “parameter” should be “parameters”
- P18 L12: “significante” should be “significance”
- P18 L13: “particular” should be “particularly”
- P19 L6: “particular” should be “particularly”
- P21 L12: “seperate” should be “separate”

We thank the reviewer for the detailed and focused review of our manuscript. We have applied all technical corrections as suggested by the reviewer.

Reviewer 2

We thank Reviewer 2 for the detailed and thorough comments. Our replies to the comments are inserted below in blue colour.

Review of

“Improving the LPJmL4-SPITFIRE vegetation-fire model for South America using satellite data”

Drüke et al. GMD-2019-92

The paper utilises a genetic optimization algorithm and a revised fire danger index to improve the representation of burnt area and biomass in the LPJmL4-SPITFIRE model compared to satellite-derived datasets, optimised against those same datasets. The authors also benchmarked the fractional cover of one PFT and claimed improvements to PFT distribution and temporal dynamics both (inter-annual variability and seasonal patterns). They also advocate the use of such methods for improving fire-vegetation models in general. Investigating alternatives to the Nesterov Index in SPITFIRE (and other global fire models) and using optimisation algorithms to develop DGVMs and fire models are laudable aims and this work makes useful contributions in these directions. Simultaneously using both burnt area and biomass observations to constrain the model parameters, and the application of rigorous benchmarking metrics are also to be commended. Many parameters in SPITFIRE are very poorly constrained, so this is a promising approach to improve the model.

However, I do have substantial reservations regarding the presentation and, to some extent, the methodology, which I believe need to be addressed prior to publication. I first list my main concerns, and then a series of comments to the text. I feel confident that these concerns can be addressed in a revised version of the manuscript, perhaps with some additional analysis.

We thank the Reviewer for this feedback and we hope that the revised manuscript through its improved analysis and by clarification of the methods will remove the reviewer’s concern .

Main concerns

1. Whilst the optimisation procedure produces very reasonable results in the case of the VPD FDI, the Nesterov Index results are not so clear cut and cast some doubt on the efficacy of the method. Yes, the summary metrics for spatial BA do get better (at least the NSME does, the Willmott coefficient goes down, which I assume means worsening agreement?), the temporal metrics do improve drastically for Caatinga but worsen for the Cerrado, and the biomass results are basically unchanged. So that is a mixed bag. But, most critically, a visual inspection of the BA produced shows a massive reduction in fire and almost complete spatial mis-match compared to the observations, not the preferred behaviour of a fire model! There is much to discuss here which is missing from the manuscript. Benchmarking/optimising burnt area is hard due to the large amount of zero values and then high peaks, and so getting a fire peak wrong by one or two gridcells is heavily penalised. Thus an optimisation will tend towards a conservative ‘no fire strategy’. This appears to be what is happening here, but is not discussed. This obviously raises

questions about whether or not BA can effectively be used in such a context when it produces results which objectively (in terms of metrics) are perhaps better, but somewhat subjectively may not actually produce a more useful model.

We thank the Reviewer for pointing out that we need to explain and discuss in more detail how the optimization was done and discuss better the implication of spatial mismatch. Indeed, the reviewer is right that optimizing a model against burned area is challenging because #1 of the skewed statistical distribution of burned area and #2 because temporal or spatial mismatches in simulated burning can cause large model-data errors. We added a paragraph to the Discussion in P22 L19-33 to clarify and explain this issue:

Generally, optimizing a model against burned area is challenging because 1) of the skewed statistical distribution of burned area and 2) because temporal or spatial mismatches in simulated burning can cause large model-data errors. These issues can be avoided with the choice of an appropriate cost function. For example, squared-error metrics tend to underestimate the variance of burned area in comparison to, e.g., the Kling-Gupta efficiency as it has been shown in the optimization of an empirical model for burned area (see Table A3 in Forkel et al. 2017). Here, the optimum parameter set for the Nesterov index-based model resulted in almost no fires across South America. Thereby the optimization algorithm tries to decrease the model error by tending towards a conservative ‘no fire strategy’ for all biomes. This result nicely demonstrates the need to evaluate model optimization results against spatially and temporally independent data and independent variables (Keenan et al. 2011).

The Nesterov index is not able to capture fire variability within the Caatinga as well as the Cerrado at the same time. This shows that the difference in the PFT distribution between these two biomes is not adequately modelled by LPJmL or just using PFT dependent scaling factors did not sufficiently improve the model performance when using the Nesterov index. On the other hand, using the VPD fire danger index reduced the model error for burned area in both biomes, by improving the modelled performance for the Caatinga and maintaining the good performance of the Cerrado region. Since improved performance of the fire model mainly had minor effect on improving FPC of the tropical PFTs, the presented optimization scheme has to go along with process-based improvements in both, in the fire and in the vegetation modules of LPJmL.

Furthermore thanks to the reviewers comment we noted a small error in Figure 5: The R^2 of 5a (original model version for the Cerrado) has been wrongly written as 0.87. The real value, as was correctly written in the text of the first manuscript versionaper, is however 0.78. Hence we have a slight improvement by the VPD version in all three metrics. We are very sorry for the confusion caused by this spelling error.

We think that by fixing this formal error and by providing additional analysis in the revised version of the manuscript (see the answers to the other concerns raised by the reviewer below) the improvements in the performance of modelled burnt area in South America are now demonstrated better. We hope to have addressed the major concerns of the reviewer adequately.

2. The optimisation to both BA and biomass is definitely a good idea, and as far as I can tell combining the two KGE metrics is reasonable. However, as part of the paper is to demonstrate this approach, I think there must be more discussion and analysis of this method. In particular, can the authors disentangle the relative constraints of each dataset in the method? I think this is important information for such a method. If all else fails, perhaps simply running the optimisation for BA and biomass individually would be an option.

The main focus of this paper was an optimization just focussed on fire parameters. Hence, the change in biomass is relatively small and heavily depending on changes in the burnt area. We included the biomass in the optimization to make sure that the biomass would not be impacted by fire-effect processes to avoid the model performance getting worse. To show the small improvement in the cost of the biomass, compared to the cost of the burnt area, we have now included in the Appendix a comparison of the cost reduction during the optimization process which is also shown here.

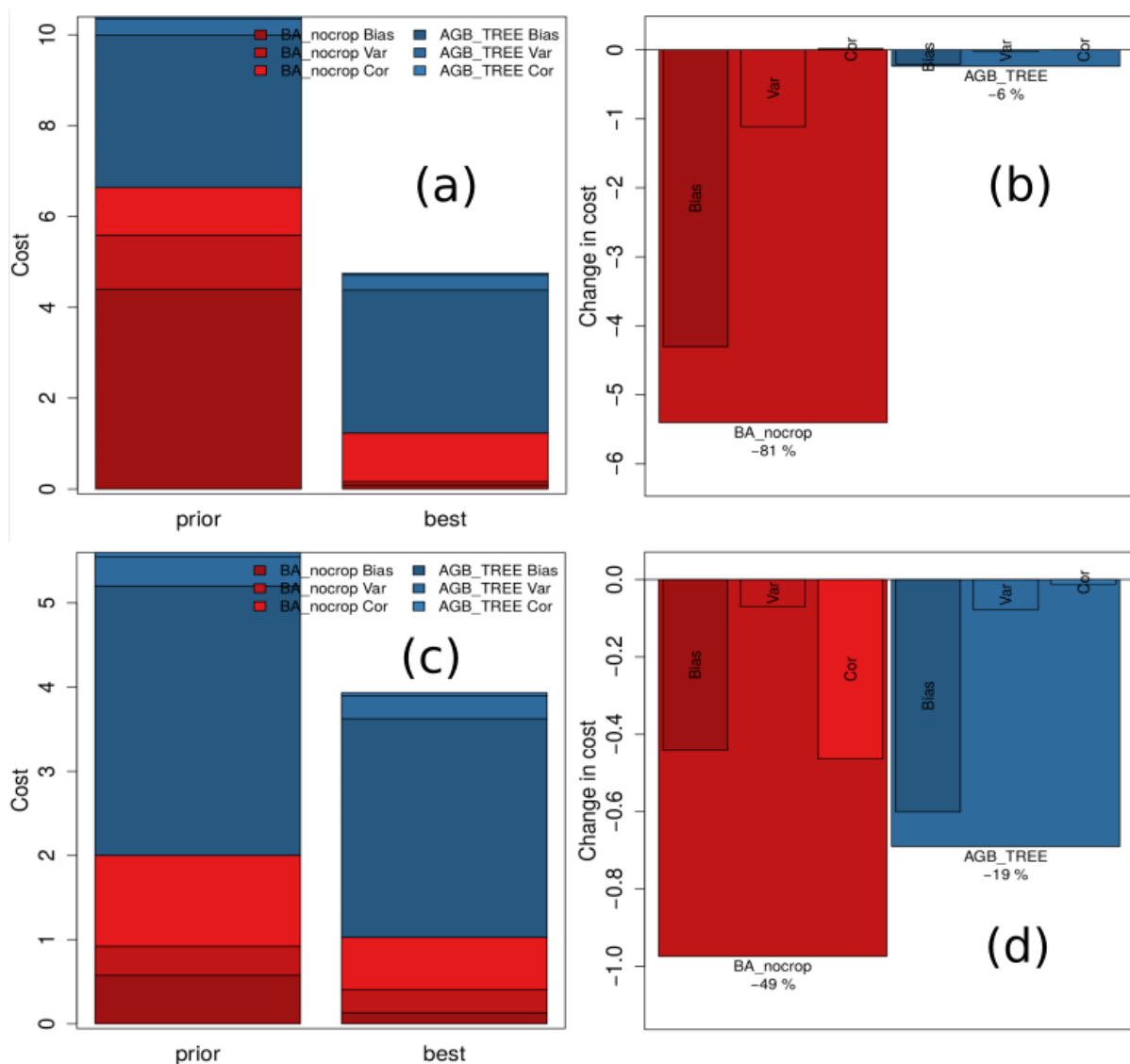


Fig A4: Cost reduction of the burnt area and the biomass during the optimization process, by showing the various components of the cost that are related to model-data bias, variance ratio and correlation. The cost for burnt area for NI_optim decreased by ca. 81%, whereas the cost of the biomass only decreases by ca. 6% (a and b). For VPD_optim the cost decreased by ca. 48% for burnt area and about 19% for the biomass (c and d). Hence the impact of the optimization process on burnt area is much larger due to the focus on fire parameters.

We added a comment on this and the reference to the appendix in the Results in P18 L3-6:

The decrease in the model error (cost) due to the optimization process has been mainly due to improvement in the burnt area. While for the NI_optim the cost of the burnt area dataset improved by 81%, the cost of the biomass dataset improved just by 6%. In case of the VPD_optim the cost of the burnt area dataset improved by 49%, whereas the biomass dataset improved by 19% (Fig .A5).

Furthermore we added to the Discussion in P21 L27-30:

Due to the focus on fire related parameter, the cost of the burnt area dataset decreases much more than the cost of the biomass dataset (Fig. A5). Hence we only get a substantial improvement in model performance in semi-arid, fire-prone biomes, where vegetation dynamics and fire are strongly coupled.

3. No specific information on how the gridcells used in the optimisation were selected. It seems to have been done just by 'picking some'. By the authors' own admission this may bias the optimisation. Could they justify their choice a little better? Furthermore, could it be possible to run with random gridcells every time? Or gridcells close to the meteorological stations used in the preparation of the climate data? A more concrete method for select the gridcells, or at least a clearer justification, is required.

We thank the reviewer for noting this lack of explanation of how the grid cells were selected. In this study we selected the grid cells manually. We justified our choice better by adding the following paragraph to the methods in P10 L22-31.

The optimization was performed for 40 grid-cells in South America to represent a variety of fire regimes (Fig. 2). We selected the grid cells manually to cover active fire regions (either in the model or in the evaluation data), specifically in the Cerrado and Caatinga. We selected a high density of grid cells in the Caatinga region to improve the very poor model performance in this region. To make sure that the model performance in the Caatinga and Cerrado was not achieved at the cost of a poor performance in other areas, we also additionally selected some cells in areas where initial fire modeling gave good results, as well as in areas where minimal or no fire occurs (central Brazilian Amazon). After inspection of the results, minor adjustments were made and the selection of the grid cells was modified to account for

neglected regions (which showed worsening of the model performance). These initial analyses actually demonstrate that the choice of grid cells is important for the model optimization and requires the development of a more thorough selection method in future model optimization applications.

A random sampling of grid cells during the optimization might potentially result in more robust model parameters as similar methods are successfully used in several machine learning approaches to make more robust predictions (e.g. bagging in random forest regressions). However, such a bagging of training points (or grid cells) within an optimization of a DGVM is currently computationally not feasible because it would require to run the optimization algorithm several (hundreds) times. Sampling random grid cells within a single optimization run will likely not result in a parameter optimum because the cost would change in each iteration which is however not related to model parameters but to the sampling of the grid cells.

We hope to have explained better why and how we manually selected our grid cells.

4. There is no discussion of what the optimised parameters mean in terms of process understanding or what the newly introduced 'alpha' for the VPD FDI really means. Many of the existing parameters move very little (perhaps a little surprising but also perhaps reassuringly), but the rCKs are very interesting. For NI optim these converge to very similar values and move away strongly from their initial values. Having similar crown kill probability for raingreen and evergreen trees flies in the face of the assumptions in SPITFIRE so far. But for VPD optim the story is somewhat different, with rCH for TrBE remaining very high, but rCK for TrBR also increasing. Please discuss these results, including some ecological context.

And regarding the new 'alpha', what does this really mean? The very different value for TrBE compared to TrBR and TrH definitely deserves some discussion as it appears to be integrating some new factor into the FDI which the NI does not include and is not adequately represented in the other SPITFIRE PFT-specific parameters. Some discussion, even if it is a little speculative, is necessary here. In general I can see no problem in tuning process-based models with 'black box' optimisation procedures and somewhat unphysical variables, but there must be at least some attempt to interpret and relate the results back to the processes.

We thank the reviewer for this important note. While we already explain some of our parameter results in section 3.2 in terms of process understanding we now added some ecological interpretation of the VPD_optim results in the Discussion in P22 L12-17.

The mortality risk of TrBE for VPD_optim remains close to the prior value of 1, confirming previous assumptions about its high fire sensitivity. Whereas the rCK for TrBR increased to 0.48, close to the upper boundary of the optimization, meaning that a mortality risk of 50% when the full crown is scorched and a 7% mortality risk when 50% of the crown is scorched, which makes the TrBR less resistant against crown damage than before. Due to this changes the overestimation of biomass in the original model for the Cerrado/Caatinga region decreased (see Fig. 7).

Moreover, we agree that the new alpha values deserve some further discussions and we now added a paragraph to the Discussion section in P22 L2-12:

'The fire danger index scaling factors (α_{NI_i} and α_{VPD_i}) convert the quantified fire risk (NI or VPD) into the actual fire danger (FDI). Both scaling factors thus set the magnitude of the fire danger for the different PFTs. Hence they impact directly the fire spread, burnt area and the number of fires as well as indirectly fire mortality. These very important parameters vary significantly for the different PFTs. TrH has the smallest scaling factor in case of both FDIs, which leads to a lower fire danger compared to the other PFTs. This indicates a prior overestimation of the fire danger of grass in tropical South America, as grasslands are generally parametrized to have a low fire resistance and moisture content and can hence burn very easily. This overestimation, compared to tree PFTs has been decreased by the optimization. In case of the VPD also the TrBR is scaled by a much smaller factor than the TrBE, which leads to a lower fire danger index. This is due to the fact, that the TrBR is dominant in dry and fire-prone regions, which experience frequent fires. Here the burnt area was often overestimated by SPITFIRE (e.g. Caatinga or eastern Cerrado) and is now decreased. On the other hand, a larger FDI for the TrBE allows more fire in wetter regions at the edge between the Cerrado and the Amazon rainforests, where TrBE is more dominant.'

5. Again, relating to the process-understanding, plots of the fire intensity resulting from the methods should be shown (possibly in an appendix if necessary). The “fuel moisture -> combustion completeness -> fire intensity -> mortality” link is a critical pathway in these results, it should be discussed explicitly but is not.

We thank the reviewer for this this comment. Unfortunately, the simulation of fire spread and fire behaviour in SPITFIRE does not have the link “fuel moisture -> combustion completeness -> fire intensity -> mortality”. Fuel consumption depends on fuel moisture in SPITFIRE, hence it is not a fixed parameter like consumption completeness, which is often used in other fire models. Surface fire intensity thus depends on the consumed fuel, fire spread and wind speed. The surface fire intensity is then used twofold: 1) to check if it is too low to support a spreading fire. If this is the case, number of fires, burnt area and all fire effects are set to zero. 2) Surface fire intensity is used to quantify flame length to quantify if the flame could scorch the canopy. If this is the case, fire mortality from crown scorch is quantified and the biomass of the dead trees is distributed to the dead fuel classes (see eqs. 5-7). Please note, that the model does not simulate active crown fires. In our opinion surface fire intensity is temporarily highly variable which makes it difficult to plot it into a map and interpret its influence on fire behaviour. We hope that we could clarify the role of surface fire intensity on linking fuel moisture and tree mortality with this explanation. We regard the influence of fire spread and fire danger index as the more important variables in SPITFIRE compared to the fire intensity, because they impact fuel consumption, fire intensity, and tree mortality. With the support of Fig. 3 we discussed, why and how a changing fire danger index has an impact on burnt area. To make this point clearer we added the following text to the manuscript describing Fig. 3 in P8 L 22-30:

,The general behavior of the two indices as modelled by LPJmL in dependence of relative humidity and temperature is shown in Fig. 3. The Nesterov index shows a strong but very localized maximum for high temperatures and a small humidity. Hence a spreading fire is only possible in a very small climate range (here ca. from 25° Celsius and a relative humidity smaller than 0.5). The VPD on the other hand shows a less pronounced maximum but a medium fire danger also for wetter and colder regions. The slope of towards lower VPD values is also smaller compared to the Nesterov index. Especially in regions with temperatures colder than 20°C and relative humidity smaller than ca. 0.6 a fire is still possible. This might increase the area in which fires can occur compared to the Nesterov index, which could be an important improvement, enabling SPITFIRE to simulate more fire in wetter and colder regions. The calculated VPD and NI values shown in Fig. 3 are based on a LPJmL-SPITFIRE run, and thus the influence of vegetation distribution on both fire danger indices.'

6. There is no benchmarking of the PFTs that we expect to be effected by fire! The inclusion of TrBE PFT FPC is great, but what about TrBR and TrH? These should be at least plotted, and ideally benchmarked. If the ESA CCI dataset does not have useful classifications in this regard, at least MODIS VCF MOD44B Tree-Nontree-Bare would provide some reference data for the Caatinga and Cerrado.

We thank the reviewer for pointing out that this information is also important for the reader to get a complete picture of the spatial distribution of all tropical PFTs in the study region. As suggested by the Reviewer we added the benchmarking of TrBR and TrH to the Appendix and changed the text accordingly in the Results (20 L5-6) (see answer to question 27 and A1 and A2). The maps of the PFT distributions were derived from the ESA CCI land cover map V2.0.7 (Li et al., 2018; Forkel et al., 2014).

Specific comments to the text

Abstract

7. 'partly poor'

-rephrase

We changed the sentence into:

'However, most fire-enabled DGVMs have problems in capturing the magnitude, spatial patterns, and temporal dynamics of burnt area as observed by satellites.'

8. 'as a starting point'

– rephrase, this is not the first work to improve fire in DGVMs

This might have been a misunderstanding. The term 'as a starting point' refers to the

improvements related to improving the SPITFIRE model. To clarify we excluded ,as a starting point‘.

9. ‘improves simulation of ... plan functional type’
– is that really demonstrated?

It is demonstrated that the distribution of the PFTs improve for the fire-prone regions Caatinga and Cerrado. As suggested by the Reviewer we provide more details and clarifications also about the other PFTs (see answers to questions 27 and 30). To clarify we add here “distribution”:

‘improves simulation of ... the spatial distribution of plant functional types’

Introduction

10. P2 ‘recent decline in global burnt area’
– now contested (indeed by one of the authors)

We thank the Reviewer for noting this and changed the sentence into:

‘Despite a tendency for globally declining burnt area (Andela et al., 2017; Forkel et al. 2019), more frequent and [...]’

11. P2 ‘Especially in South America, tropical forests, woodlands and other ecosystems are vulnerable to increasing fire danger and land use change’
– reference?

As suggested by the Reviewer we added a reference:

Cochrane, M., & Laurance, W. (2008)

[<https://bioone.org/journals/AMBIO-A-Journal-of-the-Human-Environment/volume-37/issue-7/0044-7447-37.7.522/Synergisms-among-Fire-Land-Use-and-Climate-Change-in-the/10.1579/0044-7447-37.7.522.full>]

Material and Methods

12. P6 ‘SPITFIRE further includes a surface intensity threshold’
– please state this threshold here. I realise that this is in the Table 2 but the units are not given.

We added the threshold as suggested by the Reviewer. The parameter is the fraction of burnt area per gridcell, hence has no unit. We added the threshold in P6 L28-29:

‘SPITFIRE further includes a surface intensity threshold (10^{-6} , fraction burnt area per grid cell), which describes the threshold of the possible area burnt[...]’

13. P7 ‘The fire danger index is scaled by a PFT-dependent constant, α_i , over the number of PFTs n (Thonicke et al., 2010)’

- in the original Thonicke et al. 2010 implementation, the α varied over fuel classes (ie. 1hr, 10hr, 100hr, 1000hr and live grass fuels) not PFTs. Please explain and justify this change. Also, there no mention of live grass fuels. Are they parameterised as in original SPITFIRE?

We thank the Reviewer for noting this. We indeed forgot to state in the paper that the fire danger index calculation of SPITFIRE had changed with the publication of LPJmL4 (Schaphoff et al. 2018) compared to the original version in Thonicke et al. (2010). The scaling over the relative moisture content of the 1-h, 10-h and 100-h fuel classes did no longer allow a stable modelling performance in LPJmL4.0 and had therefore been replaced as an average, PFT dependent parameter (see Eq. (63) in Schaphoff et al. 2018). The calculation of the moisture content for the live fuel consumption remained the same in LPJmL4.0 as well as in this study as described in Thonicke et al., (2010), the original SPITFIRE implementation. We added the following paragraph to the methods in P7 L20-24:

,The resulting fire danger index has been calculated as in Schaphoff et al. 2018a (slightly different compared to Thonicke et al. 2010) by taking into account the NI as measure for weather conditions and a PFT dependent scaling factor α_{NI_i} :

$$FDI_{NI} = \max \left(0, 1 - \frac{1}{m_e} \exp \left(- \frac{\sum \alpha_{NI_i} \cdot NI}{n} \right) \right)$$

where n is the number of PFTs and m_e the moisture of extinction, which is a PFT-dependent parameter and is weighted over the litter amount.'

14. P8 'and a monthly mean for R to avoid unrealistic high flammability fluctuations in time steps with isolated events of very low R '

– can the authors justify this further? I know it is stated in the Pechnoy and Shindell paper, but it is not immediately clear why flammability fluctuations due to rainfall events should be 'unrealistic'. Perhaps with their experience with this method, the authors can provide a more convincing argument.

We justified this further by adding the following paragraph in P8 L7-11:

,The soil is a natural buffer for drought periods and heavy rainfall events. In the Nesterov index this was taken into account by the cumulative nature of this index. Since the VPD-based fire danger index is not cumulative, this buffering effect is taken into account by taking the monthly mean of the precipitation. In doing so we avoid unrealistic high flammability fluctuations in time steps with isolated events of very low or very high precipitation (R).'

15. P8 'Hence, we scaled the VPD up with a PFT-dependent scaling factor α_i '

– since this has a very different physical meaning than the α_i above, I strongly suggest using a different symbol.

We followed the reviewers suggestion and changed both alphas: The scaling factor for the

Nesterov index is now called α_{Ni} and the scaling factor for the VPD α_{VPDi} . (See also Question 13)

16. P8 'The general behaviour of the two indices as modelled by LPJmL in dependence of relative humidity and temperature is shown in Fig. 3'

– Fig 3 is a nice plot, but please explain in a little more detail how the panels are comparable, as in how was the effect of vegetation taken in to account in the lower panel for VPD FDI?

We thank the reviewer for the useful comment and added more details on the Figure in P8 L22-30, see suggested text on our answer to main concern 5.

17. P8 'We regridded and aggregated the data set to the LPJmL resolution of $0.5^\circ \times 0.5^\circ$ and to a daily time step'

– normally climate data is the limiting factor when it comes to spatial resolution in DGVMs. Is there any reason that the authors chose to aggregate this rather than use 0.25° degree? Especially when the evaluation data sets are available at 0.25° degree or finer. It seems like throwing away information.

The reviewer raised an interesting point. We decided to keep the LPJmL spatial resolution at $0.5^\circ \times 0.5^\circ$ because it is also the standard resolution of other DGVMs in ISIMIP and Trendy which makes our results comparable to these models. Furthermore a gridcell size of $0.25^\circ \times 0.25^\circ$ would slow down the computation time of global model runs by a factor of 4. For the optimization we also would have to use four times the amount of tested grid cells to keep the ratio of tested grid cells to untested grid cells. In our opinion, the loss of information by taking a $0.5^\circ \times 0.5^\circ$ instead of a $0.25^\circ \times 0.25^\circ$ gridcell size is rather small for our large scale modelling approach. We hope that the reviewer can agree to our argumentation.

18. P10 'The optimization was performed for 40 grid-cells in South America to represent a variety of fire regimes (Fig. 2). Most of them were selected in active fire regions, especially in the Cerrado and Caatinga. In addition a few pixels with no or almost no fire occurrence (e.g. central Brazilian Amazon) were chosen.'

– this is a rather vague description of what may be a very important choice in the optimisation procedure! See my main concern above. Please give more details in the logic here.

We thank the reviewer for noting this lack of explanation on how the grid cells were selected. We addressed this problem in the answer to main concern 3 and added a paragraph to the methods in P10 L22-31 (see answer to main concern 3).

19. P10 –Despite being important the FDI NI and the Rothermal equations, and being poorly

constrained, moisture of extinction was not mentioned as a possible parameter for optimisation. Could the authors discuss this?

We agree that the moisture of extinction could be also a possible parameter for the optimization. We decided not to use this parameter in order to make the optimization for both fire danger indices more comparable. Without the moisture of extinction, both fire danger indices are optimized with one PFT-dependent scaling factor. Another optimized parameter in the fire danger routine would give the Nesterov index more weight in the optimization compared to the VPD.

20. P11 'NMSE'

– can the authors justify their choice of NMSE over NME?

Generally the NMSE and the NME should show very similar results with the NMSE being the squared error term. We added a paragraph to justify better our choice of the NMSE in P12 L11-14:

'We chose the NMSE to represent and compare the model errors, as it has a squared error term, which puts a stronger emphasis on large deviations between simulations and observations as compared to a linear term, and due to its normalization it is comparable across different parameters. Especially for fire simulations we have a relatively large deviation between simulations and observations.'

21. P11 'Willmott coefficient'

– please explain its range and meaning, as is done for NMSE.

It is true that we forgot to give an overview about the different values for the Willmott coefficient. We added the following sentence in P12 L17-20:

'The Willmott coefficient is a squared index, where a value of 1 stands for perfect agreement between simulated and modelled runs and gets smaller for worse agreements with a minimum of 0. Unlike the coefficient of determination, the Willmott coefficient is additionally sensitive to biases between simulations and observations.'

Results

22. P12 'mainly by shifting much of the simulated burnt area from the sparsely vegetated Caatinga towards the Cerrado region'

– this is true to some extent, but it also much is moved into Amazonia in regions where very little fire is observed in reality. In order to back up this statement, a table with the burnt area in each region for each simulation should be provided. The overestimation of fire in Amazonia should be discussed in the Discussion section.

It is true, that much of the fire is also moved into Amazonia regions. With the Nesterov index fire was strongly underestimated in the Amazonia region, while the optimized VPD fixes this

underestimation. Fires are present at the edges of the Amazon (both in model and observation, see Fig. 4), where tree density is lower and deforestation takes place. In the closed continuous forest area towards the center of the Amazon almost no fire is observed and also not simulated. We initially did not add the Amazonia region to Fig. 5 and Table 1, because the focus of this paper lies on the Cerrado and Caatinga region. But we agree that a comparison between all three regions is useful and added the Amazonia region to Table 1 and added a Figure of the timeseries of the Amazon to the Appendix. Furthermore we added a paragraph to the Results in P14 L4-11:

'Also the Amazonia region mostly improved by using the VPD_optim version (Tab. 1 and Fig. A3). The R^2 and the Willmott coefficient improved, while the NMSE increased slightly. With the Nesterov index fire was strongly underestimated in the Amazon region, while the optimized VPD fixes this underestimation. The fire is only modelled (and also observed, see Fig. 4) at the edges to the Amazon, where wood density is lower and deforestation takes place. In the closed continuous forest area towards the center of the Amazon almost no fire is observed and neither modelled. The total burnt area increased from 0.7 million ha to 4.8 million ha (for VPD_optim) , which is now a bit overestimated to the observed burnt area of 3.4 million ha. Using the NI_optim all error metrics as well as the total burnt area decreased.'

and to the Discussion section in P23 L8-10:

'The poorly modelled PFT distribution also is responsible for the overestimation of the burnt area in the Amazon region. Because of the too large fraction of TrBE in the Cerrado/Caatinga region the scaling factor for this PFT is relatively high. This leads in turn to an overestimation in the Amazon region, where the fraction of the TrBE is larger.'

Region	NMSE	Willmott	Sum
Spatial - South America			
GFED			$1.9 \cdot 10^7$
NI _{orig}	1.80	0.27	$1.4 \cdot 10^7$
VPD _{optim}	0.82	0.56	$1.6 \cdot 10^7$
NI _{optim}	1.09	0.08	$0.2 \cdot 10^7$
Temporal - Cerrado			
GFED			$9.2 \cdot 10^6$
NI _{orig}	0.30	0.89	$5.2 \cdot 10^6$
VPD _{optim}	0.27	0.90	$6.4 \cdot 10^6$
NI _{optim}	1.07	0.36	$0.6 \cdot 10^6$
Temporal - Caatinga			
GFED			$0.4 \cdot 10^6$
NI _{orig}	327.82	0.14	$6.0 \cdot 10^6$
VPD _{optim}	15.2	0.46	$1.6 \cdot 10^6$
NI _{optim}	1.07	0.73	$0.3 \cdot 10^6$
Temporal - Amazonia			
GFED			$3.4 \cdot 10^6$
NI _{orig}	0.83	0.56	$0.7 \cdot 10^6$
VPD _{optim}	0.93	0.83	$4.8 \cdot 10^6$
NI _{optim}	1.22	0.32	$0.02 \cdot 10^6$

Tab. 1: Comparison of the results in terms of NMSE, the Willmott coefficient of agreement and the sum (in ha per year) between NI_{orig}, VPD_{optim}, NI_{optim} and the GFED evaluation data

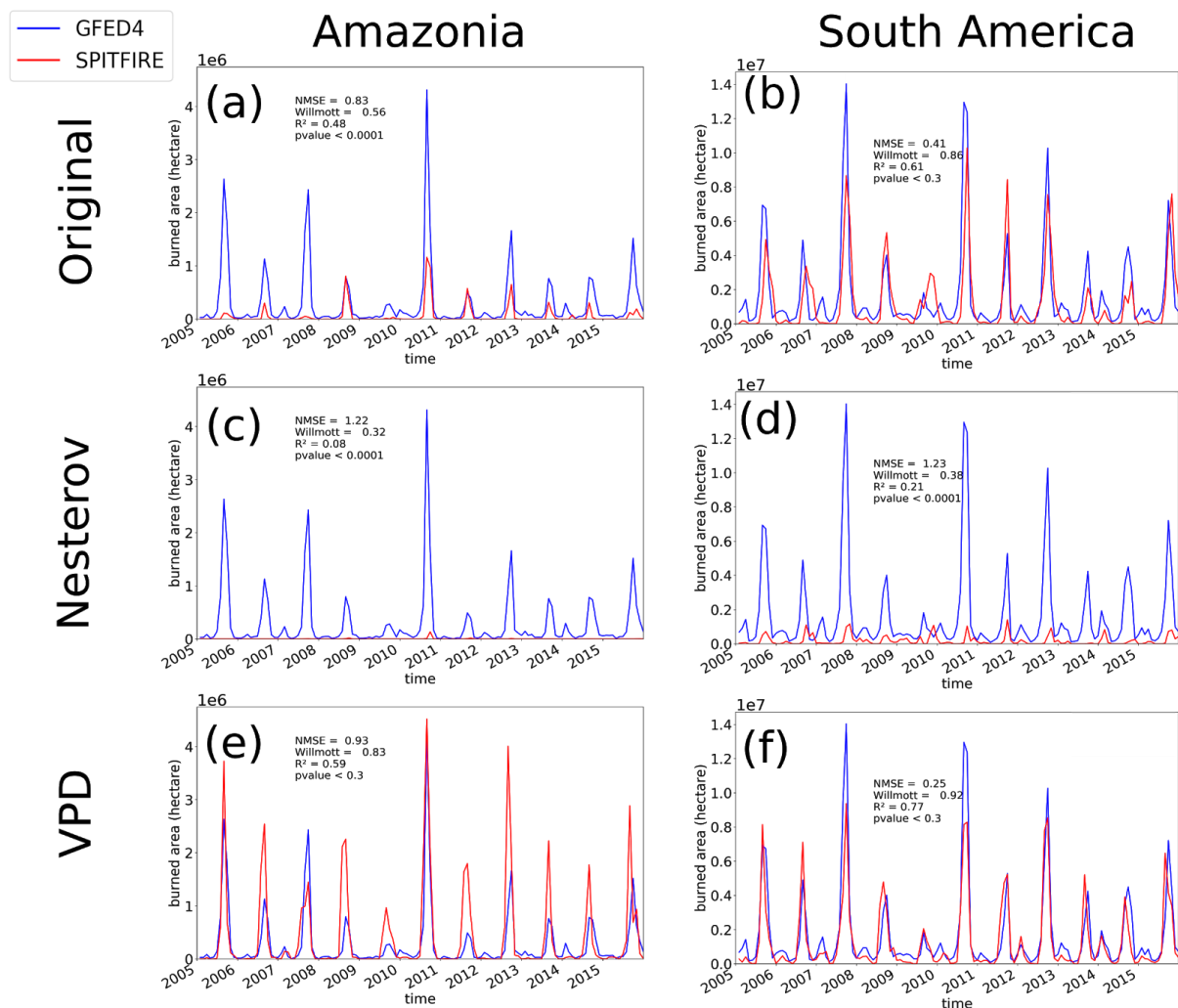


Fig. A6. Time-series of monthly burnt area from 2005 - 2015 simulated by SPITFIRE (red lines) compared to GFED4 evaluation data (blue lines) for: (a) The Amazonia region, using NI_orig. (b) Total South America, using the NI_orig. (c) The Amazonia region, using NI_optim. (d) Total South America, using NI_optim. (e) The Amazonia region, using VPD_optim. (f) Total South America, using VPD_optim.

23. P13 Figure 5

- There is some fire in the Amazonia region, both in the data and in the simulations. Therefore, this region should be included in Figure 5 and Table 1, and discussed.

We now added the Amazonia region to Table 1 and the timeseries of the region to the Appendix. Please see more discussion to this point in question 22.

24. P13 'Here, the TrBE showed the largest value (22.41), ca. 20 times as large as the TrBR (1.21) and TrH (1.13) (Tab. 2)'

– there is no discussion of what this actually means in the Discussions section, please include an interpretation.

We agree that the new alpha values deserve some further discussions and added a paragraph to the Discussion section in P22 L2-12 (see main concern 4).

25. P17 ‘... but also here we got an even larger improvement, when only the fire-prone regions Cerrado or Caatinga are considered (Tab. 3)’

- Caatinga results are not shown in Tab. 3, although I think they should be. Possibly also results for Amazonia (see above)

We focused in this Table on showing the larger improvements of fire in the fire prone Cerrado (where the fire optimization has a large impact), in comparison to the only small improvements in total South America. We did not show other regions and PFTs, because this would enlarge the Table/Figures and the Result/Discussion part and go beyond the scope of this paper. We see the usefulness of further regions and PFTs and now added the Caatinga to Table 3. We decided to not add the Amazon region in this context, because the main message of this table is, that in the fire-prone regions (Cerrado and Caatinga) the improvements in model performance are much larger compared tototal SA (which has several areas without fire).

We however now included Amazonia in Tab 1 and Fig. A3 (see answers to questions 22 and 23). We hope that this additional information now describes our results better.

Region	NMSE	Willmott
AGB		
SA (NI _{orig})	0.97	0.83
SA (VPD _{optim})	0.91	0.84
SA (NI _{optim})	0.99	0.83
Cerrado (NI _{orig})	15.06	0.25
Cerrado (VPD _{optim})	12.36	0.28
Cerrado (NI _{optim})	16.06	0.24
Caatinga (NI _{orig})	11.93	0.32
Caatinga (VPD _{optim})	8.57	0.36
Caatinga (NI _{optim})	10.44	0.33
FPC - Evergreen (TrBE)		
SA (NI _{orig})	0.42	0.82
SA (VPD _{optim})	0.41	0.82
SA (NI _{optim})	0.43	0.81
Cerrado (NI _{orig})	1.04	0.60
Cerrado (VPD _{optim})	0.70	0.64
Cerrado (NI _{optim})	1.40	0.55
Caatinga (NI _{orig})	1.73	0.40
Caatinga (VPD _{optim})	1.54	0.29
Caatinga (NI _{optim})	2.05	0.44

Tab. 3. Comparison of the results for AGB and the TrBE PFT cover in terms of NMSE and the Willmott coefficient of agreement between NI_{orig}, VPD_{optim} and NI_{optim} in South America (SA), in the Cerrado and in the Caatinga.

26. P 18 Figure 7

– difference plots are great and I can see the logic behind including the difference relative to the original model version to show improvements (as you have done) but please show the absolute values too (as in Figure 4).

We decided to not show the absolute values, because it is not possible to see any difference between the original model versions by eye. Hence, a difference plot is the best way to show the changes between the different model versions. We added the figure for the absolute values to the Appendix (Fig. A5).

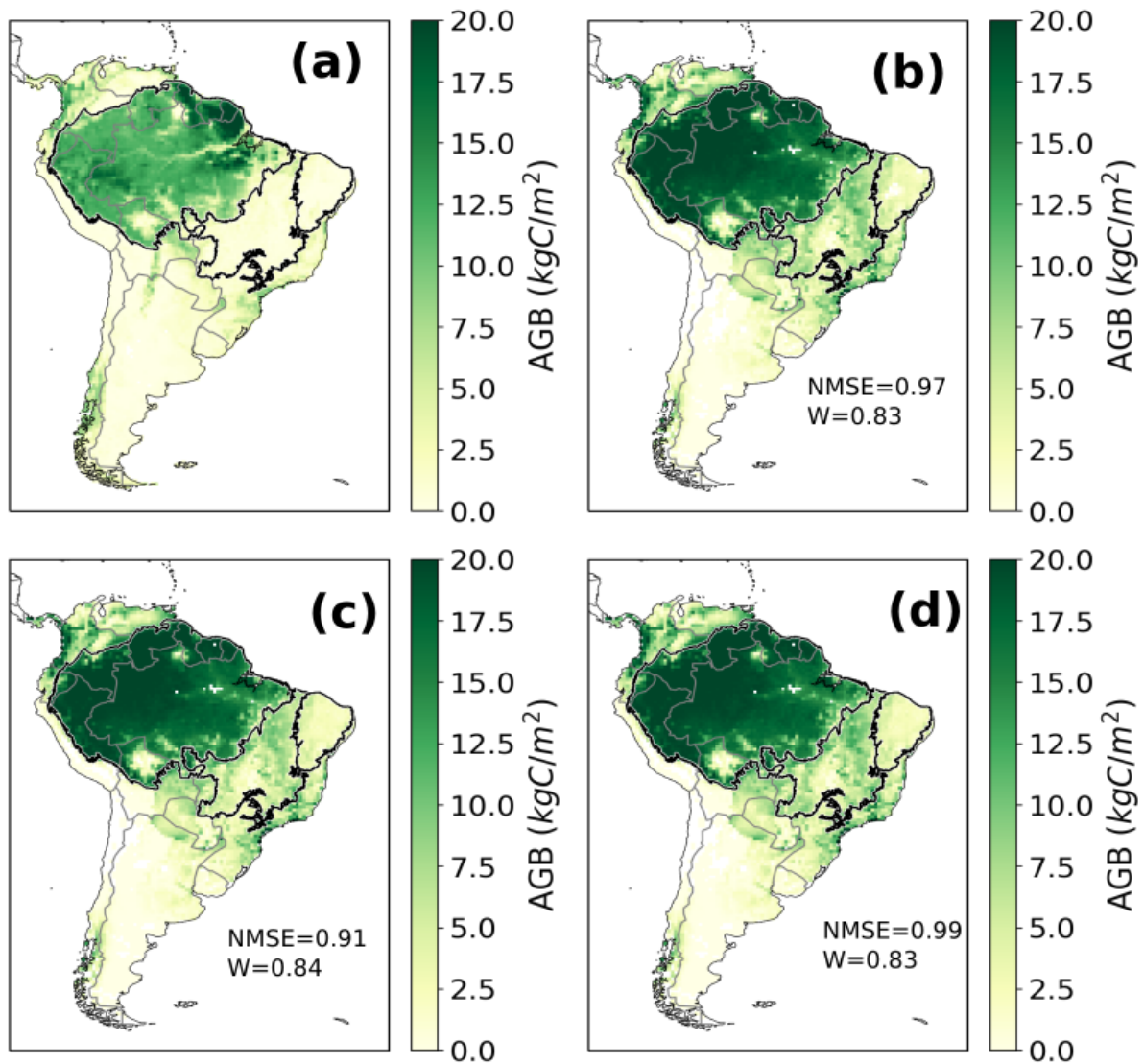


Fig. A5. Annual above ground biomass (AGB) of trees over a mean from 2005-2015 in kgC/m^2 . (a) Avitabile evaluation data. (b) Simulated AGB by LPJmL4-SPITFIRE in the NI_orig version. (c) Simulated AGB by LPJmL4-SPITFIRE in the VPD_optim version. NI\$_orig\$. (d) Simulated AGB by LPJmL4-SPITFIRE in the NI_optim version.

Discussion

27. P19 'Another result of the optimizing procedure, using FDI_VPD , was the improvement of the PFT distribution..'

– I am not sure that statement is justified given the very small improvement in TrBE and no demonstrated improvement in the other PFTs.

We agree that the improvement of the PFT distribution in total South America is relatively small. But since we only optimize fire parameters, we can only improve areas, where fire was misrepresented in SPITFIRE. Here an optimized burnt area can have a feedback on the PFT distribution. As demonstrated in the fire-prone Cerrado the improvement of PFT distribution is relatively large (the NMSE is halved compared to the original model version). Since we have in most parts of northern South America only two tree PFTs (TrBE and TrBR) and a very small amount of grass, we did just show the TrBE, assuming that the TrBR improves accordingly. Now we show (and discuss) the distribution of the other PFTs in the supplement.

It is the aim of our paper to show the improvement of burnt area due to the new fire danger index and the optimization. Hence, it was important that the improved fire representation does not decrease the performance of the PFT distribution. For the fire-prone regions, with large fire-vegetation feedbacks, we even find an improved PFT distribution.

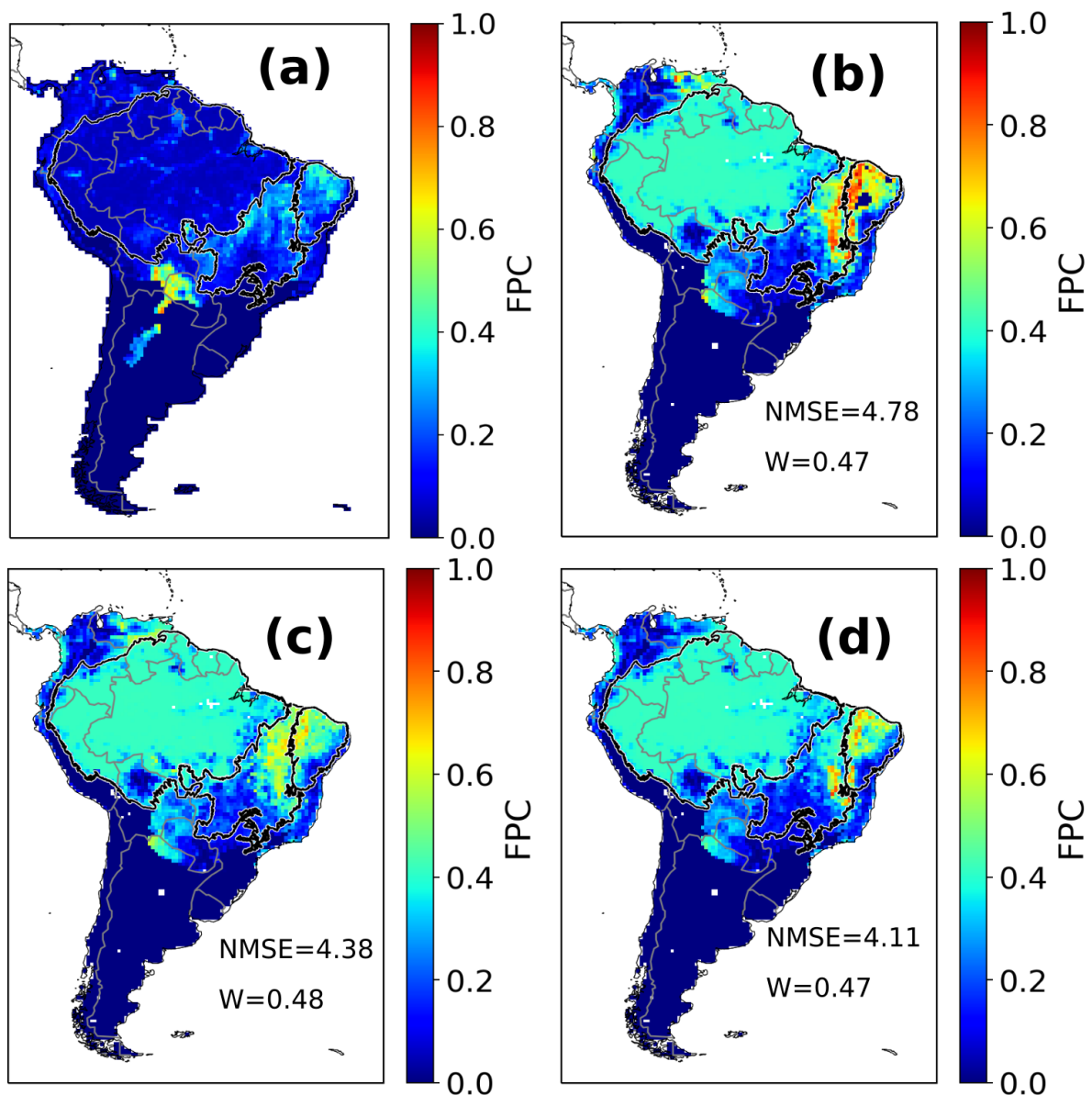


Fig A1. Annual FPC cover by tropical broadleaved raingreen PFT over a mean from

2005-2015 as fraction per cell. (a) ESA-CCI evaluation data (b) Simulated FPC by LPJmL4-SPITFIRE using the NI orig version (c) Simulated FPC by LPJmL4-SPITFIRE using the VPD optim version (d) Simulated FPC by LPJmL4-SPITFIRE using the NI optim version

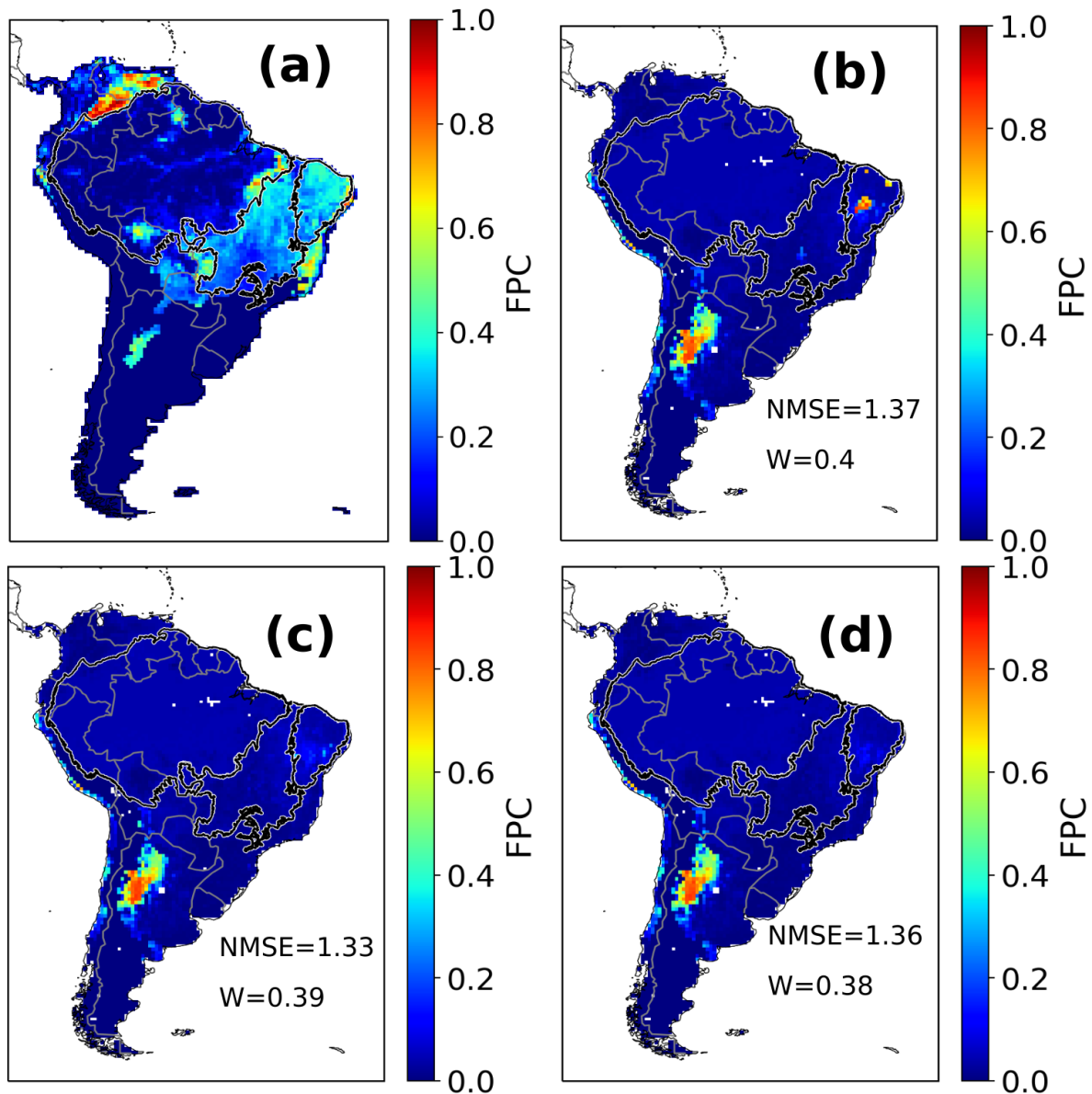


Fig. A2. Annual FPC cover by tropical herbaceous PFT over a mean from 2005-2015 as fraction per cell. (a) ESA-CCI evaluation data (b) Simulated FPC by LPJmL4-SPITFIRE using the NI orig version (c) Simulated FPC by LPJmL4-SPITFIRE using the VPD optim version (d) Simulated FPC by LPJmL4-SPITFIRE using the NI optim version

We added a short description to the result section in P20 L5-6:

Also for the the TrBR and TrH PFT distributions we got an improved performance using the VPD_optim and mostly a worse performance using the NI_optim version (Fig. A1 and A2).

28. P19 'it emphasizes that three parameter sets determining PFT distribution'
– what three parameter sets? You mean three PFTs? Or something else? Please clarify.

We agree with the Reviewer that this sentence is not clear. We rewrote it as following in P21 L25-27:

'Hence, it emphasizes that we need to include further parameters in the optimization, which impact directly the PFT distribution, biomass and fire to obtain a significant improvement in the spatial and temporal distribution of both, vegetation and fire. However, this study focused solely on the parameters within the SPITFIRE module.'

29. P20 'Limitations during the optimization process'
– this heading is somewhat confusing and maybe should better be 'Limitations of the optimization process'

We thank the Reviewer for this suggestion and changed the title of the subsection accordingly.

30. P20 'As shown in Fig. 8, the modelled PFT coverage showed an equal distribution of tropical raingreen and evergreen PFTs throughout wide parts of central-northern South America'

– Fig. 8 shows no such thing, it only shows the FPC of the evergreen PFT. Of course, it may simply be that the caption is incorrect somehow, but otherwise the distribution of the raingreen PFT must be shown to demonstrate this.

We now show the fraction of the raingreen and herbaceous PFT in the Appendix (see answer to question 27). We added this information to the text in P23 L2-3:

'As shown in Fig. 8, A1 and A2, the modelled PFT coverage showed an equal distribution of tropical raingreen and evergreen PFTs throughout wide parts of central-northern South America'

31. P20 'By choosing a large amount of optimization cells in the, by NI orig , strongly overestimated Caatinga region, the burned area decreased there significantly after the optimization'

– this (slightly confusing statement) would appear to indicate that the authors acknowledge that their results depend heavily on the choice of grid cells for the optimisation (see above)

We thank the Reviewer for noting that we could discuss this point better. We addressed the cell selection method in the answer to main concern 3. Some further issues with the optimization procedure for the different fire danger indices and regions were discussed in the answer to main concern 1. As stated in the answer to main concern 3 we changed a part of the Discussion in P10 L22-31.

32. P20 'In the Cerrado and especially the Caatinga, however, trees suffer from water stress in the dry season and should shed their leaves to avoid mortality related to drought or growth efficiency. The resulting dominance of the TrBR PFT has a very different effect on fire spread and is more fire-tolerant (different fuel characteristics and resulting fire intensity), thus has a lower fire-related mortality.'

– whilst this a reasonable enough statement (in fact pretty much inherent in the construction of DGVMs and SPITFIRE) it is hard to see what it has to do with the limitations of the optimisations process.

The aim of this statement was to emphasize the importance of a correct PFT distribution. If we have ca. 50% raingreen PFTs in the Amazon (which should rather establish in the drier Cerrado and Caatinga region) their specific traits pose a problem to the optimization procedure. But we agree that this sentence could be omitted. Hence, we deleted this sentence and changed the paragraph (P23 L5-8) into:

'In the tropical rainforest the TrBR proportion is overestimated, which leads to problems in the optimization procedure, since TrBR has very different effects on fire spread and is more fire-tolerant (different fuel characteristics and resulting fire intensity). This leads to a lower fire-related mortality, which fits better to the drier and fire prone savanna-like regions (e.g. Cerrado).'

33. P21 – 'Nonetheless, we were able to improve the interannual variability and hence, the model performance during extreme years for the Cerrado and Caatinga regions (e.g. for 2007/2008, Fig. 5). The optimized SPITFIRE is now able to model accurately the climate dependent seasonal and interannual variability as well as the spatial extent of fire on natural land throughout the fire-prone woodlands of South America.'

– yes and no. In the Cerrado the results from Fig 5. are not significantly different between VPD and Original, and whilst the results are better in the Caatinga for VPD, most of this comes down to the overall normalisation, it is hard to see if VPD really catches between IAV and seasonal dynamics. In fact, the R^2 (which is insensitive to the normalisation) actually gets worse going from Original to VPD. So these statements need much more nuance. And a plot of the normalised time series (equivalent to Fig 5., at least for the Caatinga) might be a more effective way showing improvements in IAV and seasonal dynamics.

We thank the Reviewer for noting that we could discuss more about the raised concern. We addressed this issue in the answer to main concern 1. Furthermore, thanks to the reviewer's comment we noted a small error in Figure 5: The R^2 of 5a (original model version for the Cerrado) has been wrongly written as 0.87. The real value, which was correctly written in the original text of the previous manuscript version, is however 0.78. Hence we have a slight improvement by the VPD version in all three metrics. We are very sorry for the confusion this spelling error has caused.

In the revised manuscript we now also show a Figure of the Amazonia region and for total South America in the Appendix (see Fig. R3). In total South America all the metrics improve significantly which indicates an improved IAV. While we have a large improvement in the Caatinga and in total South America, the performance and IAV for Amazonia and the Cerrado region improved slightly. The spatial extent gets better for the whole area as shown in Fig 4.

We also included now the p-value, which is mostly smaller using the VPD, indicating a more significant correlation. To make our statement clearer we rewrote the paragraph into (P23 L21-25):

'Nonetheless, we were able to improve the interannual variability and hence, the model performance to a great extent for the Caatinga and slightly for the Cerrado and Amazon regions (Fig. 5 and A3). The Cerrado already had a very good modelling performance before the optimization process, which now only slightly improved. The performance of the interannual and seasonal variability of burnt area for total South America improved substantially (Fig. A3). The optimized SPITFIRE is now better able to simulate accurately the climate dependent seasonal and interannual variability as well as the spatial extent of fire on natural land throughout the fire-prone woodlands of South America.'

We also agree with the raised concern about a more effective way to show improvements in IAV and seasonal dynamics for the Caatinga. We tried different normalization approaches, which did not lead to a better visualization. In the end we decided to show a logarithmic scale for the Caatinga in order to take into account the large differences between observation and the different model versions. A version of the plots with a non logarithmic scale remains in the Appendix.

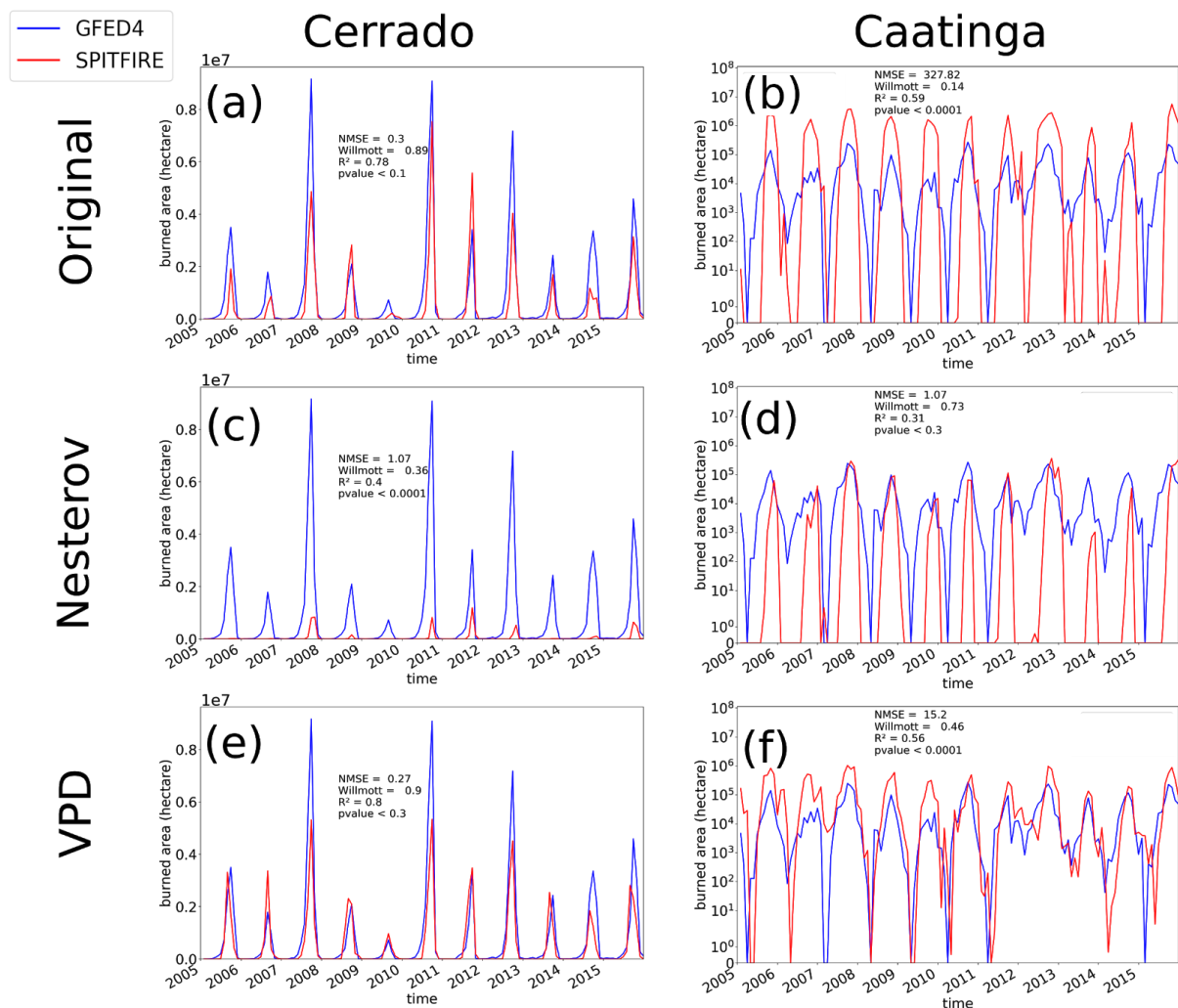


Fig. 5. Timeseries of monthly burnt area from 2005 - 2015 simulated by SPITFIRE (red lines) compared to GFED4 evaluation data (blue lines) for: (a) The Cerrado region, using NI_orig. (b) The Caatinga region, using the NI_orig. (c) The Cerrado region, using NI_optim. (d) The Caatinga region, using NI_optim. (e) The Cerrado region, using VPD_optim. (f) The Caatinga region, using VPD_optim. Note the logarithmic scale for the Caatinga, which was applied in order to account for the large differences between the different model versions (for a non logarithmic version see Fig. A6).

34. P21 entire section titled 'Outlook

- the way ahead in improving fire modules in DGVMs' – this text does not really fit the title. Much of it refers specifically SPITFIRE or LPJml, specifically their current limitations. Please reconsider/revise/re-title this section.

We thank the Reviewer and followed the suggestion and retitled this section. The name is now: *Limitations of fire modeLling in LPJmL4-SPITFIRE*.

35. P21 The statements 'it would be possible to use an even more comprehensive fire

danger index (e.g. Canadian Fire Weather Index; Wagner et al., 1987) or different fire danger indices for different biomes' and 'In a global modelling approach, however, we need to find one fire danger index' seem to contradict each other, please resolve!

It is true that both statements contradict each other. We wanted to express, that it would be ideal to have different fire danger indices for different biomes, because of the different fire dynamics. However, in our global modelling approach it is necessary to have just one fire danger index. The reason for this is, e.g., the computational effort or the exact definitions of biomes, which are also changing within one model run and hence do not fit to the LPJmL model logic. We have now removed the first sentence to avoid the confusion and changed the next sentence slightly in P24 L17-19:

'Fire models embedded in DGVMs should build on a FDI which is complex enough to account for various fire dynamics, while its parameterization should be simple enough to be accurately applied on a global scale. While the VPD is more complex and takes into account more climatic input as the Nesterov index, it is still relatively easy to implement in a global fire model.'

Conclusions

36. P21 'We have demonstrated a major improvement of the fire representation within LPJmL4-SPITFIRE by implementing a new fire danger index and applying a model-data integration setup to optimize fire-related parameters.'

- whilst there are tangible improvements, they are only tested and in the Caatinga and Cerrado, the region for which the optimisation was done (which you do mention in the next sentence). I would suggest toning this down slightly.

We are confident that the revised manuscript now shows better the improvements in performance of modelled burnt area in South America. As suggested by the reviewer we added the study region to this sentence.

'We have significantly improved the fire representation within LPJmL4-SPITFIRE, applied for South America, by implementing a new fire danger index and applying a model-data integration setup to optimize fire-related parameters.'

37. P21 'We improved the seasonal and interannual variability'

– I have yet to be convinced of this, especially as the R^2 for the time series are not improved with VPD. And I am not sure how to interpret the Willmott coefficient as this is not described.

We addressed this issue in the answer to main concern 1 and question 33 and describe now the Willmott coefficient (see answer to question 21).

We hope to have convinced the reviewer that the revised manuscript now shows better the improvements in performance of modelled burnt area in South America.

In the revised manuscript we now also show a Figure of the time-series of total South

America and the Amazon region in the Appendix, which show the improvements in model performance and IAV of the burnt area when using the optimized VPD. Hence a large improvement in the Caatinga and total South America and a slightly better performance in the Amazon and Cerrado region leads to a total improvement of IAV. Regarding the modelled R^2 please see our response to question 33.

38. P21 'A realistic representation of fire is also crucial for fire-vegetation-climate feedbacks and is hence necessary for DGVMs coupled within and comprehensive Earth system model.'
– I think you can drop that sentence, as it attempts to summarise and justify fire modelling in general rather than this work. The penultimate sentence is fine to end with.

We thank the Reviewer for this suggestion and dropped the last sentence.

We thank the reviewer for the detailed and focused review of our manuscript.

Improving the LPJmL4-SPITFIRE vegetation-fire model for South America using satellite data

Markus Drüke^{1,2}, Matthias Forkel³, Werner von Bloh¹, Boris Sakschewski¹, Manoel Cardoso⁴, Mercedes Bustamante⁵, Jürgen Kurths^{1,2}, and Kirsten Thonicke¹

¹Potsdam Institute for Climate Impact Research, Telegraphenberg A 31, Potsdam, Germany

²Humboldt Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Deutschland

³TU Wien, Department of Geodesy and Geoinformation, Gusshausstr. 27-29, 1040 Vienna, Austria

⁴Instituto Nacional de Pesquisas Espaciais, Av. dos Astronautas, 1.758 - Jardim da Granja, São José dos Campos - SP, 12227-010, Brazil

⁵Instituto de Ciências Biológicas, Universidade de Brasília, Campus Universitário Darcy Ribeiro - Asa Norte, 70910-900 Brasília, Brazil

Correspondence: Markus Drüke (drueke@pik-potsdam.de)

Abstract. Vegetation fires influence global vegetation distribution, ecosystem functioning, and global carbon cycling. Specifically in South America, changes in fire occurrence together with land use change accelerate ecosystem fragmentation and increase the vulnerability of tropical forests and savannas to climate change. Dynamic Global Vegetation Models (DGVMs) are valuable tools to estimate the effects of fire on ecosystem functioning and carbon cycling under future climate changes.

5 However, most fire-enabled DGVMs have ~~partly poor performances~~ problems in capturing the magnitude, spatial patterns, and temporal dynamics of burnt area as observed by satellites. As fire is controlled by the interplay of weather conditions, vegetation properties and human activities, fire modules in DGVMs can be improved in various aspects. ~~As a starting point, we here~~ In this study we focus on improving the controls of climate and hence fuel moisture content on fire danger in the LPJmL4-SPITFIRE DGVM in South America and especially for the Brazilian fire-prone biomes Caatinga and Cerrado. We

10 therefore test two alternative model formulations (standard Nesterov index and a newly implemented water vapor pressure deficit) for climate effects on fire danger within a formal model-data integration setup where we estimate model parameters against satellite data sets of burnt area (GFED4) and above ground biomass of trees. Our results show that the optimized model improves the representation of spatial patterns and the seasonal to inter-annual dynamics of burnt area especially in the Cerrado/Caatinga region. In addition, the model improves the simulation of above-ground biomass and the spatial distribution of

15 plant functional types (PFTs). We obtained the best results by using the water vapor pressure deficit (VPD) for the calculation of fire danger. The VPD includes, in comparison to the Nesterov index, a representation of the air humidity and the vegetation density. This work shows the successful application of a systematic model-data integration setup, as well as the integration of a new fire danger formulation, in order to optimize a process-based fire-enabled DGVM. It further highlights the potential of this approach to achieve a new level of accuracy in comprehensive global fire modelling and prediction.

1 Introduction

Fire in the Earth system is an important disturbance leading to many changes in the vegetation and has substantial impact on biodiversity, human health and ecosystems (Langmann et al., 2009). Fire is responsible for ca. 2 Pg of carbon emissions, which constitutes 20 % of global carbon emissions (Giglio et al., 2013; Werf et al., 2010). Fire-induced aerosol emissions and land surface changes modify evapotranspiration and surface albedo and have therefore a crucial impact on global climate (van der Werf et al., 2008; Yue and Unger, 2018). Despite ~~of a recent decline in global burnt area (Andela et al., 2017)~~ a tendency for globally declining burnt area (Andela et al., 2017; Forkel et al., 2019b), more frequent and intense drought-periods lead ~~worldwide~~ to increasing fire-prone weather and surface conditions worldwide and therefore fire danger (Jolly et al., 2015). Growing fire danger along with land-use change are increasing ecosystem's vulnerability, which ~~can~~ could in turn shift entire regions into a less vegetated state (Silvério et al., 2013). To account for these effects, it is extremely important to include well performing fire modules in Dynamic Global Vegetation Models (DGVMs).

Especially in South America, tropical forests, woodlands and other ecosystems are vulnerable to increasing fire danger and land use change ~~-This paper (Cochrane and Laurance, 2008). This study~~ focuses on the fire ~~behaviour~~ behavior in central-northern South America and especially on the Brazilian ~~fire-prone~~ biomes Caatinga and Cerrado, which are the most fire-prone regions in South America (Fig. 1). Together with the Amazon rainforest they form an area of very high biodiversity and have a large impact on the global carbon-cycle and the regional water cycle (Lahsen et al., 2016). Compared to the Amazon, the Cerrado

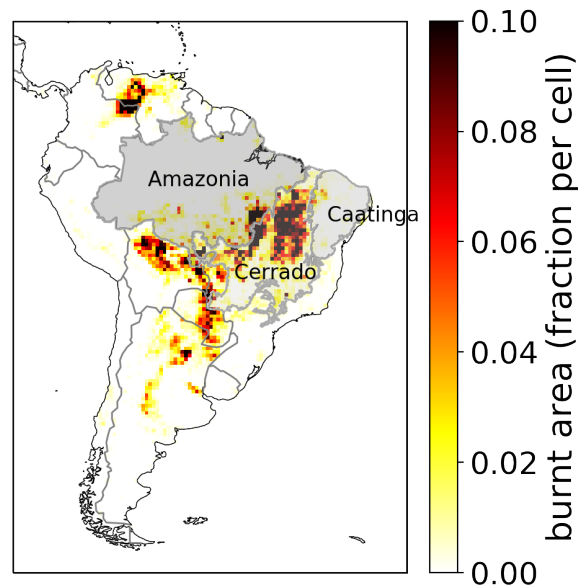


Figure 1. Overview of the mean annual burnt area in Brazil from 2005-2015 (Werf et al., 2017; Giglio et al., 2013) and the biomes Amazonia, Cerrado and Caatinga (IBGE, 2019; Harvard, 2019)

and Caatinga are both less densely vegetated and drier biomes, but with very different vegetation and precipitation dynamics.

The Cerrado is a savanna-like biome with a mixture of shrubs, high grasses and dry forest parts. With a precipitation of ca. 1500 mm per year the Cerrado does experience a rainy season. The Caatinga, on the other hand, has a semi-arid climate with irregular rainfall between 500 and 750 mm per year, mostly within only a few months of the year. The vegetation is heterogeneous and characterized by deciduous dry forest and shrubs (Alvares et al., 2013; Prado, 2003). The different vegetation types [of the Caatinga and the Cerrado](#) lead to different fire spread, fire intensity, fire resistance and fire mortality properties. While within the Cerrado fire is a frequent event and the plants are mostly adapted to it (70 % of burnt area in Brazil is within the Cerrado, Moreira de Araújo et al., 2012), the Caatinga has a lower fire intensity and fire spread due to a lower biomass, ~~which is~~ available for fuel. Such variability in the vegetation and dead fuel composition, within and between biomes, poses a challenge to global fire models to correctly simulate observed fire patterns for a variety of biomes. Both, the Caatinga and the Cerrado depend on a strict equilibrium of fire-vegetation-climate feedbacks (Lasslop et al., 2016), which is threatened to be disturbed by human impact through climate change and land use change (Beuchle et al., 2015). While the Amazon is the focus of various national and international conservation efforts and at least by law well protected, the Cerrado is currently over-exploited by the agribusiness and its importance for regional climate, biodiversity and the water cycle is often neglected (Lahsen et al., 2016). In particular the disturbance of increasing fire regimes by climate change and land-use change might accelerate biome degradation. These effects on the Cerrado might also impact the Amazon rainforest by shifting the position of the savanna-forest biome boundary towards forest, putting the functioning of the Amazon rain forest at risk (Chambers and Artaxo, 2017). Parts of the Cerrado are also itself vulnerable to increasing fire regimes, and might shift to a less vegetated state, similar to the Caatinga (Hoffmann et al., 2000). To model these ~~feedback-processes~~ [feedback processes](#) and to study the range of biome-stability under certain drought-induced perturbations, a realistic fire representation in climate and vegetation models is essential. However ~~the modelling of the fire behaviour~~, [modelling fire behavior](#) of the Brazilian Cerrado and Caatinga presents a huge challenge.

The fire occurrence depends on many interconnected parameters as humidity, precipitation, temperature, ignition sources (lightning and human) and windspeed, but also on fuel load, fuel moisture and the adaption of plant traits to fire (Keeley et al., 2011), which makes the development of fire models a complex task (Forkel et al., 2019a; Hantson et al., 2016; Lasslop et al., 2015; Krawchuk and Moritz, 2011; Jolly et al., 2015). Global fire modelling is done either by empirical models (e.g. Thonicke et al., 2001; Knorr et al., 2016; Forkel et al., 2017) or by process-based models (e.g. Venevsky et al., 2002; Thonicke et al., 2010). Empirical fire models are simplified statistical representations of fire processes and are based on empirical relationships between variables (e.g. soil moisture and fire occurrence). Process-based fire models attempt to simulate fire via explicit process-based relations: Fire ignitions are calculated by taking into account lightning flashes as natural sources and human ignitions. The chance of an ignition to become a spreading fire is then determined by the fire danger index. Sophisticated fire models ~~have further a function for calculating~~ [calculate](#) the rate of spread by taking into account wind speed and then translate these results into an area burnt, fuel consumption and fire carbon emissions (~~Thonicke et al., 2010; Hantson et al., 2016~~) [\(e.g. Thonicke et al. \(2010\); see Hantson et al. \(2016\) for an overview of global fire models\)](#).

Weather conditions control the moisture content of fuels and the danger of fire ~~ignitions to ignite~~ and spread. Hence the simulation of fire danger plays an important role to simulate the occurrence of fire within global process-based fire models

(Pechony and Shindell, 2009). Temperature, precipitation, humidity and vegetation-related variables are often used to compute fire weather indices and hence to estimate the risk of ~~fire ignitions or spread ignitions to become a spreading fire~~ (Chuvieco et al., 2010). Various fire weather indices ~~have been previously developed and~~ are used within operational fire danger assessment systems (e.g. Canadian Fire Weather index (~~FWI~~), FWI (Wagner et al., 1987), the Keetch Byram Drought Index (Keetch and Byram, 1968), the Angström Fire Danger Index (Arpaci et al., 2013), and the Nesterov index (Venevsky et al., 2002)). However, regional studies show that fire weather indices tend to have different predictive performances for fire occurrence (Arpaci et al., 2013). Hence, the performance of different fire weather indices should be ideally ~~assessed~~ tested in order to accurately represent fire danger in DGVMs. However, not all fire weather indices can be easily adapted for global fire models because they require input variables that are not available within a DGVM framework. Hence a fire danger index for a DGVM should be as complex as necessary but still relatively easy to implement. As a result, the relatively simple Nesterov index has been widely used within global fire models (Venevsky et al., 2002; Thonicke et al., 2010).

Here, we aim to improve the simulated occurrence of fire (i.e. burnt area) in the LPJmL4-SPITFIRE model for South America and in particular for the fire-prone biomes Cerrado and Caatinga. We aim to ~~test~~ evaluate the performance of two alternative fire danger indices within SPITFIRE (~~based on the already implemented Nesterov index (Venevsky et al., 2002)~~ and the newly implemented water ~~vapour pressure deficit (Pechony and Shindell, 2009; Ray et al., 2005)~~ vapor pressure deficit (VPD thereafter, Pechony and Shindell, 2009; Ray et al., 2005)). Furthermore, we apply a formal model-data integration framework (LPJmLmdi, Forkel et al., 2014) to estimate model parameters that control fire danger, fire ~~behaviour~~ behavior, fire resistance and mortality against satellite-based data sets of burnt area and above-ground biomass (Fig. 2). Our approach is likely to improve the representation of spatial-temporal variations in fire ~~behaviour in different climates and~~ behavior in different biomes to enable a much better modelling of the impact of climate change on fire-vegetation interactions in the current century.

2 Materials and Methods

2.1 The coupled vegetation-fire model LPJmL4-SPITFIRE

2.1.1 LPJmL 4.0

The LPJmL 4.0 model (Lund-Potsdam-Jena managed Land, Schaphoff et al., 2018a, b), is a well established and validated process-based DGVM, which globally simulates the surface energy balance, water fluxes, carbon fluxes and stocks, and natural and managed vegetation from climate and soil input data. LPJmL simulates global vegetation distribution as the fractional coverage of plant functional types (PFT), which is called foliage projective cover (FPC), and managed land as fractional coverage of crop functional types (CFT). The establishment and survival of different PFTs is regulated through bioclimatic limits and effects of heat, productivity and fire on plant mortality. Therefore, it enables LPJmL to investigate feedbacks, for example between vegetation and fire. In standard settings, which are also used here, the model operates on the grid of $0.5^\circ \times 0.5^\circ$ ~~lat-lon~~ latitude-longitude with a spinup time of 5000 years, repeating the first 30 years of the given climate data set.

Since its original implementation by Sitch et al. (2003), LPJmL has been improved by a representation of the water balance

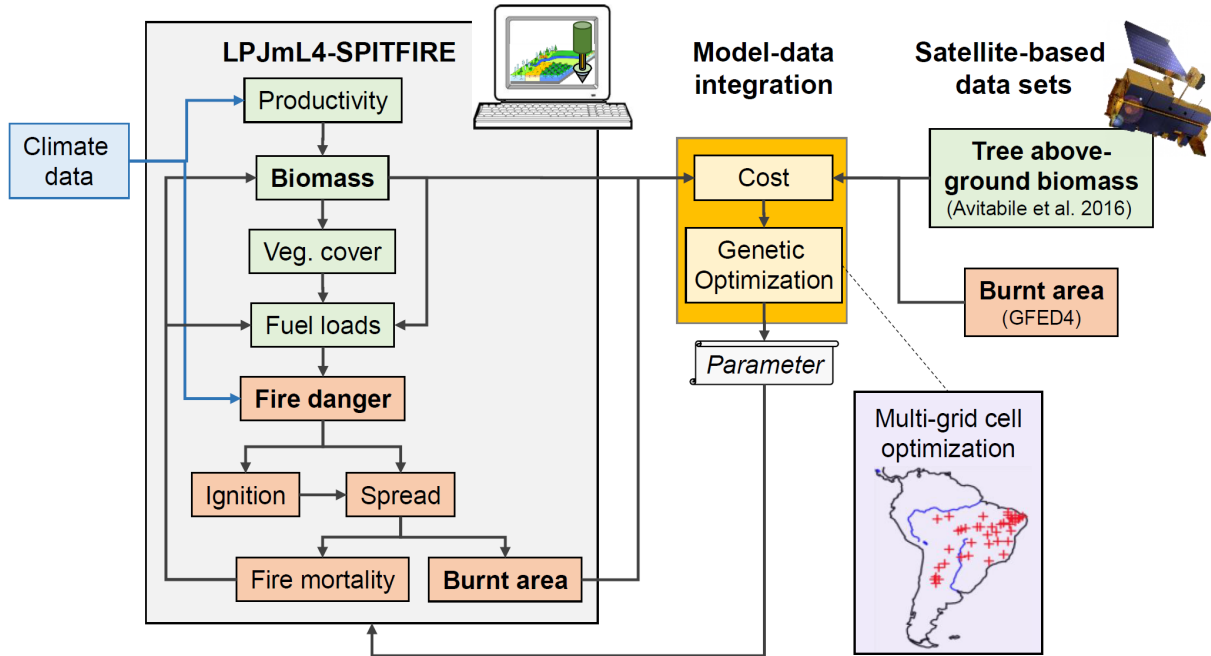


Figure 2. Schematic overview ~~about~~ of the model-data integration approach to estimate parameters of LPJmL4-SPITFIRE against satellite-based data sets of burnt area and ~~above-gorund~~ above-ground biomass

(Gerten et al., 2004), a representation of the agriculture (Bondeau et al., 2007), and new modules for fire (Thonicke et al., 2010), permafrost (Schaphoff et al., 2013) and phenology (Forkel et al., 2014).

2.1.2 SPITFIRE

- 5 SPITFIRE (SPread and InTensity of FIRE, Thonicke et al., 2010) is a process-based fire module, used in various vegetation models (e.g. Lasslop et al., 2014; Yue and Unger, 2018), including LPJmL4. We describe here its main features, which are published in Thonicke et al. (2010). SPITFIRE calculates fire disturbance by simulating the ignition, the danger, the spread and the effects of fire separately. As ignition sources SPITFIRE considers human ignition and lightning flashes. Human ignitions ($n_{h,ig}$) are calculated as a function of population density:

$$10 \quad n_{h,ig} = P_D \cdot k(P_D) \cdot a(N_D)/100, \quad (1)$$

where

$$k(P_D) = 30.0 \cdot \exp(p_h \cdot \sqrt{P_D}). \quad (2)$$

P_D is the human population density (individuals km^{-2}) and $a(N_D)$ (ignitions $\text{individual}^{-1} \text{day}^{-1}$) describes the inclination of humans to cause fire ignitions (Eq. 3 and 4 in Thonicke et al., 2010). p_h is a parameter, which is set to -0.5 in Thonicke

et al. (2010). This relationship assumes that human ignitions are lowest on very low populated regions and on high populated regions through a higher level of urbanization and landscape fragmentation. The ignition is highest for a medium-small population density. Lightning-caused ignitions are prescribed by lightning data from the OTD/LIS Gridded Climatological data set (Christian et al., 2003), assuming that 20 % of the flashes reach the ground and ~~their effectiveness to ignite a fire is 0.044 %~~

5 ~~of cloud-to-ground strikes can start a fire~~. In the study area of South America human ignitions are by far the most dominant ignition source, due to missing lightning in the dry season.

Fire danger is by default computed by using the Nesterov index which accounts for the maximum and dew point temperatures as well as scaling factors for different PFTs on a daily time step. In the following section, we describe the calculation of the fire danger indices in detail (Sect. 2.2). Fire duration t_{fire} (min) is calculated as a function of the fire danger index, assuming
10 that fires burn longer under a high fire danger:

$$t_{fire} = \frac{241}{1 + 240 \cdot \exp(p_{dt} \cdot FDI)}, \quad (3)$$

where ~~p_{dt}~~ is set to -11.06 in Thonicke et al. (2010). The maximum fire duration per day is 240 minutes.

The calculation of the forward rate of spread $ROS_{f,surface}$ (m min^{-1}) is based on the Rothermel equations (Rothermel, 1972; Pyne et al., 1996; Wilson, 1982):

$$15 \quad ROS_{f,surface} = \frac{I_R \cdot \xi \cdot (1 + \Phi_w)}{\rho_b \cdot \epsilon \cdot Q_{ig}}, \quad (4)$$

where I_R is the reaction intensity, ξ the propagation flux ratio, Φ_w a multiplier that accounts for the effect of wind, ϵ the effective heating number, Q_{ig} the heat of pre-ignition and ρ_b the fuel bulk density (Eq. 9 in Thonicke et al., 2010). ~~ρ_b is a PFT dependent (kg m^{-3}) is a PFT-dependent~~ parameter and describes the density of the fuel, which is available for burning. It is weighted over the different fuel classes. Hence, a changing PFT distribution has an impact on ~~the ROS~~ $ROS_{f,surface}$.

20 The simulated fire ignitions, fire danger and fire spread are then used to calculate the burnt area, fire carbon emissions, and plant mortality. Plant mortality depends on the scorch height (SH) and the probability of mortality due to crown damage $P_m(CK)$. ~~The scorch height SH~~ SH describes the height of the flame at which canopy scorching occurs. It increases with the 2/3 power of the surface intensity $I_{surface}$:

$$SH = F \cdot I_{surface}^{0.667}, \quad (5)$$

25 where ~~F~~ is a PFT-dependent parameter. Assuming a cylindrical crown, the proportion ~~CK~~ CK affected by fire is ~~calculated~~ as:

$$CK = \frac{SH - H + CL}{CL}, \quad (6)$$

where ~~H~~ H is the height of the average woody PFT and ~~CL~~ CL the crown length. The probability of mortality $P_m(CK)$ due to crown damage is then calculated by:

$$30 \quad P_m(CK) = rCK \cdot CK^p, \quad (7)$$

where r_{CK} is a PFT depended resistance factor between 0 and 1, and p in the range of 3 to 4. Disturbance by fire mortality has a large impact on the vegetation dynamics, which are calculated within LPJmL. SPITFIRE further includes a surface intensity threshold (10^6 , fraction burnt area per grid cell), which describes the threshold of the possible area burnt below which the surface intensity is set to zero and hence burnt area, emissions and fuel consumption is set to zero.

- 5 SPITFIRE considers anthropogenic effects on fire by taking into account human ignitions but does not account for fire suppression. Only wildfires occurring in natural vegetation are simulated. Fire on managed land like agriculture or pasture areas is not implemented, which has to be taken into account if simulated burnt area is compared with satellite observation. Furthermore, we introduced a small technical change in the LPJmL4 interaction with SPITFIRE compared to the original SPITFIRE implementation: In the version 4.0 of LPJmL the fire litter routine calculates the leaf and litter carbon pools in a
- 10 daily time step. Since the LPJmL tree allocation works at a yearly time step, this implementation leads to an incorrect LPJmL4-SPITFIRE interaction. We now split the fire-litter routine into two parts; the first one allocates burnt matter into the litter at a daily time step without recalculating the pools and the second one calculates the leaf and root carbon pools at a yearly time step.

2.2 Fire danger indices

- 15 The fire danger index (FDI) is a key parameter within process-based fire models such as SPITFIRE. The FDI determines the probability and the intensity of a spreading fire, which impacts fire ~~behaviour~~ behavior.

2.2.1 Nesterov index-based fire danger index (FDI_{NI})

- The fire danger index within SPITFIRE is based on the daily (d) calculated Nesterov Index ~~NI(d)~~ $NI(d)$ (Venevsky et al., 2002), which is widely used in numeric fire simulations. The NI is a cumulative function of daily maximum temperature
- 20 $T_{max}(d)$ ($^{\circ}C$) and dew-point temperature $T_{dew}(d)$ ($^{\circ}C$) and set to zero at a precipitation ≥ 3 mm or a temperature ≤ 4 $^{\circ}C$:

$$NI(d) = \sum T_{max}(d) \cdot (T_{max}(d) - T_{dew}(d)), \quad (8)$$

$$T_{dew} = T_{min}(d) - 4. \quad (9)$$

- ~~The final FDI is then calculated~~ The resulting fire danger index has been calculated as in Schaphoff et al. (2018a) (slightly
- 25 ~~different compared to Thonicke et al. (2010))~~ by taking into account the ~~Nesterov index NI~~ as measure for ~~the fire weather conditions~~, the total dead fuel load for the different fuel classes i (ω_{0i}) and ~~weather conditions and a PFT-dependent scaling factor~~ α_{NIi} :

$$FDI_{NI} = \max \left(0, 1 - \frac{1}{\sim} m_e \exp \left(- \sum \alpha_{NIi} \cdot \frac{NI}{\sim} \right) \right), \quad (10)$$

where n is the number of PFTs and m_e the moisture of extinction m_e , the latter is a PFT-dependent, which is a PFT-dependent parameter and is weighted over the litter amount. The fire danger index is scaled by a PFT-dependent constant, α_i , over the number of PFTs n (Thonicke et al., 2010):

$$FDI_{NI} = \max \left(0, \left(1 - \frac{1}{m_e} \exp \left(- \sum_{i=1}^n \alpha_i \cdot \frac{\omega_{0i}}{\omega_0} \right) \cdot NI \right) \right).$$

5 We will use the scaling factors $\alpha_i \cdot \alpha_{NI}$ in the parameter optimization (Sect. 2.4).

2.2.2 Vapor pressure deficit-based fire danger index (FDI_{VPD})

We implemented a new fire danger index, based on the water vapor pressure deficit (VPD). The VPD describes the difference of the saturation water vapor pressure e_s and the actual water vapor pressure in the air. For the parameterization of the VPD
10 we used an approach based on Pechony and Shindell (2009):

$$VPD \propto 10^{Z(T)} (1 - RH/100), \quad (11)$$

where T is the air temperature, RH the relative humidity and Z the Goff-Gratch equation (Goff and Gratch, 1946) to calculate the saturation vapor pressure. The flammability F at time step t for each grid cell can then be expressed as:

$$F(t) = 10^{Z(T(t))} \left(1 - \frac{RH(t)}{100} \right) VD(t) e^{-c_R R(t)}, \quad (12)$$

15 where VD is the vegetation density, R the total precipitation in mm/day and c_R is a constant factor ($c_R = 2$ day/mm). Here we used the simulated FPC from LPJmL4 as a proxy for the VD and a monthly mean for R to VD . The soil is a natural buffer for drought periods and heavy rainfall events. In the Nesterov index this was taken into account by the cumulative nature of this index. Since the VPD-based fire danger index is not cumulative, this buffering effect is taken into account by taking the monthly mean of the precipitation. In doing so we avoid unrealistic high flammability fluctuations in time steps with
20 isolated events of very low R or very high precipitation (R).

Based on this implementation in SPITFIRE, the resulting FDI was much smaller than the original FDI_{NI} . Hence, we scaled the VPD up with a PFT-dependent scaling factor $\alpha_i \cdot \alpha_{VPD}$, weighted over the corresponding FPC :

$$FDI_{VPD} = \frac{\sum \alpha_i \alpha_{VPD_i} \cdot FPC_i}{\sum FPC_i} \cdot F(t). \quad (13)$$

$\alpha_i \cdot \alpha_{VPD}$ for the FDI_{VPD} was not included in (Pechony and Shindell, 2009), but is important in order to allow different fire responses for different tree and grass types. We will use the scaling factors $\alpha_i \cdot \alpha_{VPD}$ in the parameter optimization (Sect. 2.4). In comparison to the NI, the VPD FDI_{VPD} requires more climate variables as input as it uses relative humidity and vegetation cover as additional fire-relevant variables. Vegetation cover has a direct link to fire risk by providing the number of available

fuel for burning. According to many studies (e.g. Ray et al., 2005; Sedano and Randerson, 2014; Seager et al., 2015) the FDI_{VPD} is a very accurate fire danger index with a high correlation with fire occurrence, while still being relatively easy to implement in a global fire model.

The general ~~behaviour~~ behavior of the two indices as modelled by LPJmL in dependence of relative humidity and temperature is shown in Fig. 3. The Nesterov index shows a strong but very localized maximum for high temperatures and a small humidity. Hence a spreading fire is only possible in a very small climate range (here ca. from 25° Celsius and a relative humidity smaller than 0.5). The VPD on the other hand shows a less pronounced maximum but a medium fire danger also for wetter and colder regions. The slope of towards lower VPD values is also smaller compared to the Nesterov index. Especially in regions with temperatures colder than 20°C and a relative humidity smaller than ca. 0.6 a fire is still possible. This might increase the area in which fires can occur compared to the Nesterov index, which could be an important improvement, enabling SPITFIRE to simulate ~~also~~ more fire in wetter ~~regions~~ and colder regions. The calculated VPD and NI values shown in Fig. 3 are based on a LPJmL-SPITFIRE run, and thus the influence of vegetation distribution on both fire danger indices.

2.3 Model input data

LPJmL4-SPITFIRE requires input data on daily air temperature, precipitation, long-wave and shortwave downward radiation, wind and specific humidity, which ~~we took~~ are taken from the NOAA Global Land Assimilation System (GLDAS, Rodell et al., 2004). The data has a spatial resolution of $0.25^\circ \times 0.25^\circ$ and the time step is 3 hours. We regridded and aggregated the data set to the LPJmL resolution of $0.5^\circ \times 0.5^\circ$ and to a daily time step. We used the GLDAS 2.0 for the years 1948-1999 and the version GLDAS 2.1 for the years 2000-2017. GLDAS 2.1 uses multiple satellite- and ~~ground-based~~ ground-based observational data as well as advanced land surface modelling and data assimilation techniques. GLDAS 2.0 is forced entirely with the Princeton meteorological forcing data (Civil and Environmental Engineering/Princeton University, 2006). Because LPJmL4 requires at least 30 years of climate data for its spin-up (Sect. 2.1.1), the time span covered by GLDAS 2.1 is too short. To run the model, we used both climate data sets, but used the years 2003-2013 from GLDAS 2.1 for the optimization and 2005-2015 for the evaluation period.

Furthermore, LPJmL4-SPITFIRE is forced with gridded constant soil texture (Nachtergaele et al., 2009) and annual information on land use from Fader et al. (2010). Atmospheric CO_2 concentrations are used from Mauna Loa station (Quéré et al., 2015) and applied globally. The population density is taken ~~form~~ from Goldewijk et al. (2011) and the lightning flashes are taken from the OTD/LIS satellite product (Christian et al., 2003).

2.4 Model optimization

To estimate parameters of LPJmL4-SPITFIRE, we aimed to calibrate model results against satellite observations of burnt area (GFED4: Giglio et al., 2013; Werf et al., 2017). However, as fire occurrence and spread impact and depend on vegetation productivity, hence fuel load, we wanted to ensure to not over-fit LPJmL4 against burnt area but to additionally achieve a realistic vegetation distribution. Therefore, we additionally included a satellite-derived data set on above-ground biomass (~~AGB, Avitabile et al., 2016~~) of trees of trees (AGB, Avitabile et al., 2016) in the optimization. We combined burnt area and

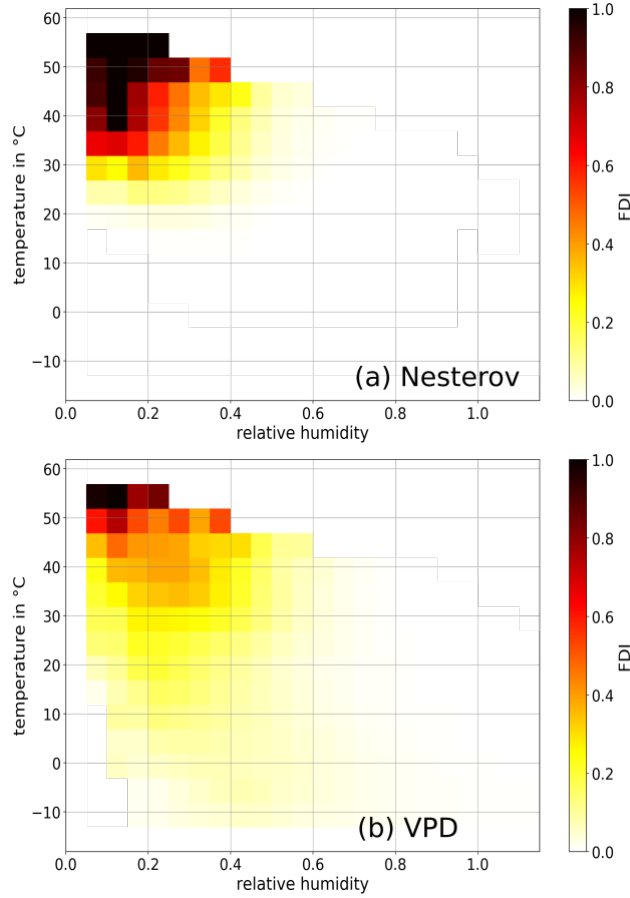


Figure 3. Dependence of the simulated fire danger index on monthly mean relative humidity and temperature for (a) the Nesterov-based index and (b) the VPD-based index. Both indices were calculated with monthly data for the years 2000-2010.

AGB with the corresponding model outputs within a joint cost function and applied a genetic optimization algorithm to estimate model parameters (Fig. 2). The implementation of the genetic optimization algorithm (Mebane and Sekhon, 2011) (GENetic Optimization Using Derivatives (GENOUD), Mebane and Sekhon, 2011) for LPJmL is described in (Forkel et al., 2014) Forkel et al. (2014). The used cost function is based on the Kling-Gupta efficiency (KGE), which is the Euclidean distance in a three-dimensional space of model performance measures that account for the bias, ratio of variance and correlation between simulations and the observations. Gupta et al. (2009) showed that the KGE performs in an optimization setup better than is better than, e.g., the Nash-Sutcliffe efficiency (and hence MSE). We extended the KGE by defining it for multiple data sets d (i.e. burnt area and AGB):

$$Cost = \sqrt{\sum_{d=1}^N \left(\frac{s'_d}{o_d} - 1 \right)^2 + \left(\frac{\sigma_{s,d}}{\sigma_{o,d}} - 1 \right)^2 + (r(s_d, o_d) - 1)^2} \quad (14)$$

where \bar{s} and \bar{o} are mean values (bias component) over space (i.e. different grid-cells) and time (e.g. months) of simulations s and the observations o , respectively. σ_s and σ_o are variances (variance component) and r is the Pearson correlation coefficient over space and time. The optimization was performed for 40 grid-cells in South America to represent a variety of fire regimes (Fig. 2). ~~Most of them were selected in~~ We selected the grid cells manually to cover active fire regions ~~, especially (either in~~
5 the model or in the evaluation data), specifically in the Cerrado and Caatinga. ~~In addition a few pixels with no or almost no fire~~
~~occurrence (e.g. We selected a high density of grid cells in the Caatinga region to improve the very poor model performance~~
~~in this region. To make sure that the model performance in the Caatinga and Cerrado was not achieved at the cost of a poor~~
~~performance in other areas, we also additionally selected some cells in areas where initial fire modeling gave good results,~~
~~as well as in areas where minimal or no fire occurs (central Brazilian Amazon) were chosen.~~ After inspection of the results,
10 minor adjustments were made and the selection of the grid cells was modified to account for neglected regions (which showed
worsening of the model performance). These initial analysis actually demonstrate that the choice of grid cells is important
for the model optimization and requires the development of a more thorough selection method in future model optimization
applications.

Several parameters of LPJmL4-SPITFIRE were included in the optimization that cover different fire processes (see Tab. 2):
15 ignition (human ignition parameter p_h , Eq. 2), fire danger (scaling factors FDI (α_{NL} and α_{VPD}), Eq. 10 and 13), fire spread
(fire duration p_{ad} , Eq. 3), fuel bulk density (ρ_b , Eq. 4), surface intensity threshold and fire effects (scorch height parameter
FF, Eq. 5; crown mortality parameter r_{CK} , Eq. 7). While p_{ad} , p_h and the surface intensity threshold are global parameters
(for all PFTs), the others were optimized for each PFT separately. Since we focus here on tropical South America, we used
20 tropical broadleaved evergreen (TrBE), tropical broadleaved raingreen (TrBR) and tropical herbaceous (TrH) PFTs for the op-
timization.

In genetic optimization algorithms, each model parameter is called an individual with a corresponding fitness, which represents
the cost of the model against the observations. At the beginning of the optimization process, the GENOUD algorithm creates
a generation of individuals based on random sampling of parameter sets within the prescribed parameter ranges. After the
calculation of the cost of all individuals of the first generation, a next generation is generated by cloning the best individuals,
25 by mutating the genes or by crossing different individuals (Mebane and Sekhon, 2011). This results, after some generations, in
a set of individuals with highest fitness, i.e. parameter sets with minimized cost. To find an optimum parameter set also used
the BFGS gradient search algorithm (named after the authors ????) within the GENOUD algorithm. An optimized parameter
set of the BFGS algorithm is used as individual in the next generation. We were applying the GENOUD algorithm with 20
generations and a population size of 800 individuals per generation, which corresponds to 16000 single model runs. We decided
30 on this amount of iterations, because the cost kept almost constant in the last iterations and the parameter values did not change
to the 6th digit, beyond which changes are not really relevant for model applications. During the optimization we ran the model
parallel for each grid cell (40 grid cells and CPU's, 3.2GHz) and had a total optimization time of ca. 24 hours.

The comparison of the two presented fire danger indices is the main objective of this study. Hence the optimization of the
~~PFT-dependent~~ PFT-dependent FDI scaling factors ~~α_{NL} is very important and~~ α_{NL} and α_{VPD} is crucial and obligatory for
35 the VPD ~~obligatory~~ because of no prior values. Accordingly, we conducted two different optimization experiments using

LPJmLmdi: First, using the VPD as FDI a FDI based on the VPD (VPD_{optim} hereafter) and secondly using the Nesterov index as FDI a FDI based on the NI (NI_{optim} hereafter). Both resulting parameter sets were then used for a corresponding LPJmL4 run-runs and were compared to the unoptimized original model version using the NI (NI_{orig} hereafter) and various evaluation data sets.

5

2.5 Evaluation data

We used burnt area from the global fire emission database (GFED4; Giglio et al., 2013; Werf et al., 2017), in the model optimization and to evaluate model results. The global data set is available at a resolution of $0.25^\circ \times 0.25^\circ$ in a monthly time step from 1997 until 2016. The GFED burnt area product is based on the 500 m Collection 5.1 MODIS direct broadcast (DB) burnt area product (MCD64A1, after 2001). We used data for the years 2003-2013 in the optimization in order to not include potential inconsistencies between the GLDAS 2.0 and 2.1 climate data sets or between burnt area observations within GFED that originate from different satellite sensors. The GFED product comes with a stratification of burnt area by land cover from the MODIS land cover map in the resolution of 500 m (Giglio et al., 2013). As LPJmL does not simulate fire on managed lands, we excluded burnt area on cropland classes from model-data comparisons. Due to lack of data we however did not account for the proportion of pastures. To constrain the simulated vegetation distribution, we used the AGB data set from Avitabile et al. (2016). This data set is approximately representative for the late 2000s and therefore we compared it against simulated AGB for the years 2009-2011. We regridded all data set to a $0.5^\circ \times 0.5^\circ$ resolution. In addition, we used maps of PFTs as derived from the ESA CCI land cover map V2.0.7 (Li et al., 2018; Forkel et al., 2014).

15

2.6 Evaluation metrics

To quantify the performance of the model output, we applied the Pearson Correlation between two time series, the normalized mean square error (NMSE; Kelley et al., 2013) and the Willmott coefficient of agreement (W; Willmott, 1982) to describe differences between the model simulation and the reference data sets. The NMSE is calculated by:

20

$$NMSE = \frac{\sum_{i=1}^N (y_i - x_i)^2}{\sum_{i=1}^N (x_i - \bar{x})^2} \quad (15)$$

where y_i is the simulated and x_i the observed value in the grid cell i . \bar{x} is the mean observed value. The NMSE is zero for perfect agreement between simulated and modelled results, 1.0 if the model is as good as using the observed mean as a predictor and larger than 1.0 if the model performs worse than that. We chose the NMSE to represent and compare the model errors, as it has a squared error term, which puts a stronger emphasis on large deviations between simulations and observations as compared to a linear term, and due to its normalization it is comparable across different parameters. Especially for fire simulations we have a relatively large deviation between simulations and observations.

25

The Willmott coefficient of agreement is given by:

30

$$W = 1 - \frac{\sum_{i=1}^N (y_i - x_i)^2 \cdot A_i}{\sum_{i=1}^N (|y_i - \bar{x}| + |x_i - \bar{x}|)^2 \cdot A_i} \quad (16)$$

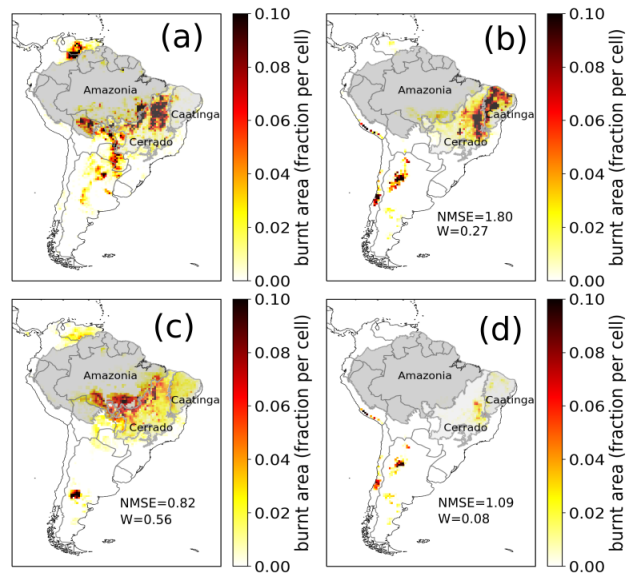


Figure 4. Yearly burnt area over a mean from 2005-2015 as fraction per cell. (a) GFED4 evaluation data of burnt area excluding crops and simulated burnt area by SPTIFIRE using the (b) NI_{orig} version, (c) VPD_{optim} version, (d) NI_{optim} version

which additionally accounts for the area weight A_i of the grid cell i . The Willmott coefficient is a squared index, where a value of 1 stands for perfect agreement between simulated and modelled runs and gets smaller for worse agreements with a minimum of 0. Unlike the coefficient of determination, the Willmott coefficient is additionally sensitive to biases between simulations and observations.

5 3 Results

3.1 Performance of optimized fire danger index formulations

Overall, the yearly burnt area simulated by the standard SPITFIRE model (using the original Nesterov index, NI_{orig}) showed poor ~~simulations~~ simulation results over South America as compared to the GFED4 evaluation data set (Fig. 4 a and b: NMSE=1.80, W=0.27). The average yearly burnt area (without croplands) for South America was with ca. 14 million ~~ha~~ ha about 25 % smaller than the observed burnt area with 19 million ha in the shown period from 2005-2015. The spatial pattern of the modelled burnt area agreed well with the GFED4 data ~~around the Cerrado~~ in the region of the Cerrado that is close to the Caatinga border, while the fires in other semi-arid regions of the continent were underestimated. For example, simulated fire is underestimated in the savanna-areas in the northern part of South America (on the Columbian-Venezuelian border) even though there is a strong signal visible in the satellite observations. The biomes ~~Caating~~ Caatinga and Cerrado, which are of
 15 special interest in this study, showed very different results: while fire in Caatinga was ~~underestimated~~ overestimated, it was ~~overestimated~~ underestimated in the Cerrado.

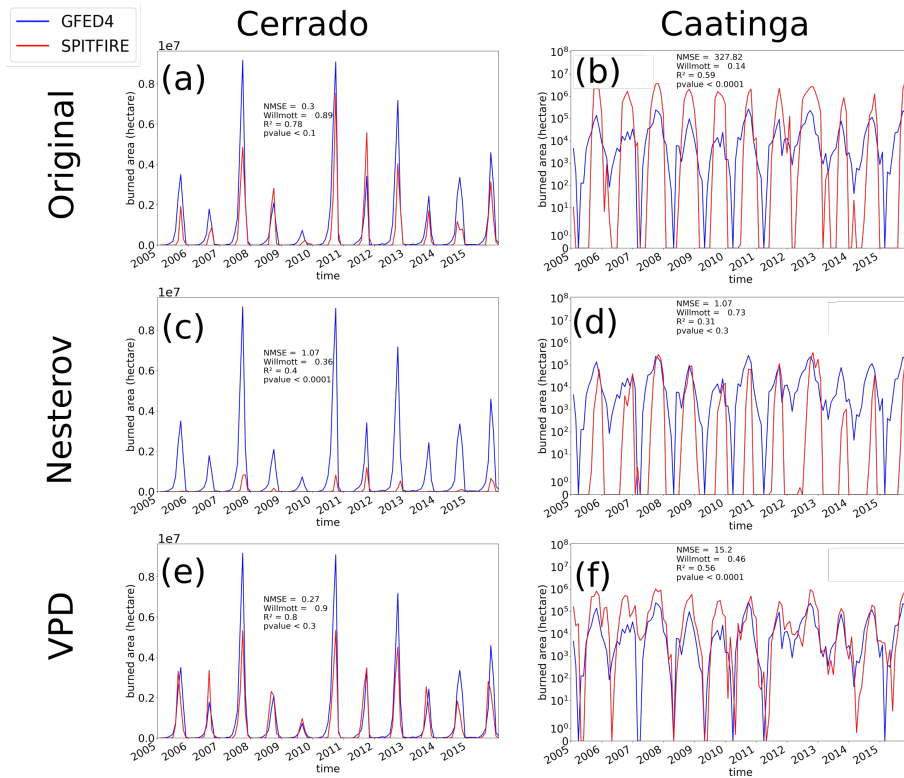


Figure 5. Timeseries-Time-series of monthly burnt area from 2005 - 2015 simulated by SPITFIRE (red lines) compared to GFED4 evaluation data (blue lines) for: (a) The Cerrado region, using NI_{orig} . (b) The Caatinga region, using the NI_{orig} . (c) The Cerrado region, using $VPDNI_{optim}$. (d) The Caatinga region, using $VPDNI_{optim}$. (e) The Cerrado region, using $NI_{VPD_{optim}}$. (f) The Caatinga region, using $NI_{VPD_{optim}}$. Note different the logarithmic scale for each Figure the Caatinga, which was applied in order to account for the large differences between the different model versions (for a non-logarithmic version see Fig. A6).

The optimized version using NI_{optim} (Fig. 4 d), led to an overall decrease of fire, with a slight improvement of NMSE (1.09) as compared to NI_{orig} and a worse Willmott coefficient of 0.08. While the overestimation of fire in Caatinga was reduced, all the fires across South America also have decreased significantly, which led to a general underestimation of fire by 90 % (2 million ha). The optimized version, using VPD_{optim} (Fig. 4 c), clearly improved the model performance, mainly by shifting much

5 of the simulated burnt area from the sparsely vegetated Caatinga towards the Cerrado region (NMSE=0.82 and W=0.56). In addition, by using VPD_{optim} , the model results also showed fire occurrence in northern South America, where fire was not at all or only minimally simulated when using NI_{optim} or NI_{orig} . The total burnt area was in this model version ca 20 % smaller than the evaluation data set (16 million ha).

The burnt area time series from 2005 to 2015 provides a more detailed view on the model performance for the fire-prone

10 Cerrado and Caatinga region (Fig. 5). While model performance was relatively good for the Cerrado region with NI_{orig} (NMSE=0.3, W=0.89, $R^2=0.78$), the simulated burnt area was strongly overestimated in the Caatinga region throughout the

whole period (NMSE=327.82, W=0.14, R²=0.59). After the optimization of the NI, the model performance indeed improved for the Caatinga (NMSE=1.07, W=0.73, R²=0.31), but at the same time the performance for the Cerrado ~~got much worse~~ declined (NMSE=1.07, W=0.36, R²=0.4). On the other hand VPD_{optim} showed an improved fire representation compared to the standard settings in the Cerrado (NMSE=0.27, W=0.9, R²=0.8) as well as in the Caatinga (NMSE=15.2, W=0.46, R²=0.56).

5 Even though fire in the Caatinga was still overestimated, the NMSE decreased by a factor of six.

Overall, the total amount of burnt area in the Cerrado was for all three model versions smaller than in the evaluation data set. Fire occurrence in the Caatinga was, on the other hand, largely overestimated by the NI_{orig} and the VPD_{optim} version. Just in the NI_{optim} version the burnt area of the Caatinga is in the same order of magnitude as the evaluation data set, which also led, however, to a large underestimation in the Cerrado (Tab. 1 ~~)~~ and Fig. 4). Also the Amazonia region mostly improved

10 by using the VPD_{optim} version (Tab. 1, Fig. A3). The R² and the Willmott coefficient improved, while the NMSE increased slightly. With the Nesterov index fire was strongly underestimated in the Amazon region, while the optimized VPD fixes this underestimation. The fire is only modelled (and also observed, see Fig. 4) at the edges to the Amazon, where wood density is lower and deforestation takes place. In the closed continuous forest area towards the center of the Amazon almost no fire is observed and neither modelled. The total burnt area increased from 0.7 million ha to 4.8 million ha (for VPD_{optim}), which is

15 now a bit overestimated to the observed burnt area of 3.4 million ha. Using the NI_{optim} all error metrics as well as the total burnt area decreased.

3.2 Optimized model parameters

Seven fire-related parameters were optimized, in order to improve the fire representation in the LPJm4-SPITIFRE model. Here we compare the optimized parameters for the different model versions in order to evaluate and discuss parameter variability and

20 changes. Table 2 shows all parameters, used for the optimization, their lower and upper boundary and the resulting optimized value. Since the FDI directly controls the amount of modelled fire, the FDI scaling factors for the different PFTs are central for this analysis. For both optimization experiments the boundaries were, hence, set rather generously within one magnitude of the original value. In the NI_{optim} experiment, all scaling factors generally decreased compared to the standard values used for NI_{orig}. Here, TrH displayed the smallest scaling factor ($9.39 \cdot 10^{-6}$), followed by TrBE ($2.48 \cdot 10^{-5}$) and TrBR ($4.76 \cdot 10^{-5}$).

25 Since the VPD is a newly implemented fire danger index, we have no standard values to compare the optimized scaling factors with. Here, the TrBE showed the largest value (22.41), ca. 20 times as large as the TrBR (1.21) and TrH (1.13) (Tab. 2).

In case of the other optimized parameters the boundaries were set smaller in order to decrease the possibility that a large error in the estimation of several parameters would lead to a better overall cost in the optimization procedure. The human ignition parameter became smaller for both optimizations, which led to a smaller amount of human ignitions (from -0.5 to -0.54 in

30 NI_{optim} and -0.53 in VPD_{optim}). The fuel bulk density increased for all three tropical PFTs in the NI_{optim} version, while for VPD_{optim} the fuel bulk density of the TrBE and TrH PFTs decreased and for the TrBR increased. For the NI_{optim} version, the fire duration parameter (τ_{FDt}) increased, leading to a shorter fire duration (from -11.06 to -9), while the value for the VPD_{optim} version stayed relatively similar (-11.37) to the prior value. The surface intensity ~~threshold~~ threshold became slightly larger for the NI_{optim} version than the original value (from 10^{-6} to $1.03 \cdot 10^{-6}$). For VPD_{optim} the parameter increased by a factor

Table 1. Comparison of the burnt area results in terms of NMSE, the Willmott coefficient of agreement and the sum (in ha per year) between NI_{orig} , VPD_{optim} , NI_{optim} and the GFED evaluation data

Region	NMSE	Willmott	Sum
Spatial - South America			
GFED			$1.9 \cdot 10^7$
NI_{orig}	1.80	0.27	$1.4 \cdot 10^7$
VPD_{optim}	0.82	0.56	$1.6 \cdot 10^7$
NI_{optim}	1.09	0.08	$0.2 \cdot 10^7$
Temporal - Cerrado			
GFED			$9.2 \cdot 10^6$
NI_{orig}	0.30	0.89	$5.2 \cdot 10^6$
VPD_{optim}	0.27	0.90	$6.4 \cdot 10^6$
NI_{optim}	1.07	0.36	$0.6 \cdot 10^6$
Temporal - Caatinga			
GFED †			$0.4 \cdot 10^6$
NI_{orig}	327.82	0.14	$6.0 \cdot 10^6$
VPD_{optim}	15.2	0.46	$1.6 \cdot 10^6$
NI_{optim}	1.07	0.73	$0.3 \cdot 10^6$
<u>Temporal - Amazonia</u>			
<u>GFED</u>			<u>$3.4 \cdot 10^6$</u>
<u>NI_{orig}</u>	<u>0.83</u>	<u>0.56</u>	<u>$0.7 \cdot 10^6$</u>
<u>VPD_{optim}</u>	<u>0.93</u>	<u>0.83</u>	<u>$4.8 \cdot 10^6$</u>
<u>NI_{optim}</u>	<u>1.22</u>	<u>0.32</u>	<u>$0.02 \cdot 10^6$</u>

Table 2. All optimized parameters with their standard values, the upper and lower boundary of the parameter ranges and the resulting optimized value including parameter for specific PFTs and global parameter, which have the same value for all PFTs. [All parameters except \$\rho_b\$ have no unit.](#)

Parameter	PFT	Standard value (as in Thonicke et al., 2010)	Lower bound.	Upper bound.	After optimization
NI_{optim}					
scaling factor FDI $\alpha_{\tau} \alpha_{NL_i}$	TrBE	$3.34 \cdot 10^{-5}$	$7 \cdot 10^{-6}$	$1.33 \cdot 10^{-4}$	$2.4885 \cdot 10^{-5}$
scaling factor FDI $\alpha_{\tau} \alpha_{NL_i}$	TrBR	$3.34 \cdot 10^{-5}$	$7 \cdot 10^{-6}$	$1.33 \cdot 10^{-4}$	$4.7649 \cdot 10^{-5}$
scaling factor FDI $\alpha_{\tau} \alpha_{NL_i}$	TrH	$6.67 \cdot 10^{-5}$	$7 \cdot 10^{-6}$	$1.33 \cdot 10^{-4}$	$9.3949 \cdot 10^{-6}$
fire duration parameter $\rho_a D_t$	all PFTs	-11.06	-13	-9	-9.0011
scorch height parameter F	TrBE	0.1487	0.01	0.6	0.1282
scorch height parameter F	TrBR	0.061	0.01	0.6	0.0752
crown mortality parameter rCK	TrBE	1.0	0.5	1	0.5030
crown mortality parameter rCK	TrBR	0.05	0	0.5	0.4038
fuel bulk density ρ_b (kg m^{-3})	TrBE	25.0	22.5	27.5	26.6473
fuel bulk density ρ_b (kg m^{-3})	TrBR	13.0	11.7	14.3	13.1896
fuel bulk density ρ_b (kg m^{-3})	TrH	2.0	1.8	2.2	2.0019
human ignition parameter p_h	all PFTs	-0.5	-0.6	-0.4	-0.5426
surface intensity threshold	all PFTs	10^{-6}	10^{-7}	10^{-5}	$1.0317 \cdot 10^{-6}$
VPD_{optim}					
scaling factor FDI $\alpha_{\tau} \alpha_{VPD_i}$	TrBE	-	1	50	22.4181
scaling factor FDI $\alpha_{\tau} \alpha_{VPD_i}$	TrBR	-	1	50	1.2135
scaling factor FDI $\alpha_{\tau} \alpha_{VPD_i}$	TrH	-	1	50	1.1299
fire duration parameter $\rho_a D_t$	all PFTs	-11.06	-13	-9	-11.3753
scorch height parameter F	TrBE	0.1487	0.01	0.6	0.1930
scorch height parameter F	TrBR	0.061	0.01	0.6	0.0799
crown mortality parameter rCK	TrBE	1.0	0.5	1	0.9983
crown mortality parameter rCK	TrBR	0.05	0	0.5	0.4801
fuel bulk density ρ_b (kg m^{-3})	TrBE	25.0	22.5	27.5	22.5923
fuel bulk density ρ_b (kg m^{-3})	TrBR	13.0	11.7	14.3	13.3750
fuel bulk density ρ_b (kg m^{-3})	TrH	2.0	1.8	2.2	1.8944
human ignition parameter p_h	all PFTs	-0.5	-0.6	-0.4	-0.5332
surface intensity threshold	all PFTs	10^{-6}	10^{-7}	10^{-5}	$3.6317 \cdot 10^{-6}$

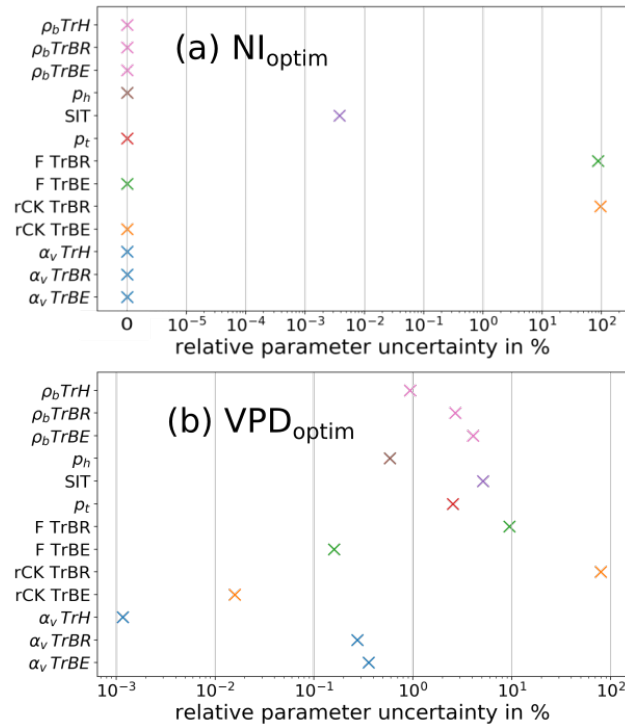


Figure 6. Relative uncertainty of model parameters after optimization for (a) $VPD_{NI_{optim}}$ and (b) $NI_{VPD_{optim}}$. The relative uncertainty is the ratio of the uncertainty after the optimization (range of all ~~parametersets-parameter sets~~ with low cost, below the 0.05 quantile) divided by the uncertainty before the optimization (range of the parameters for the optimization). Low and high values of relative uncertainty indicate strongly and weakly constrained parameters, respectively. ~~All parameters are defined by PFT (red dots) or global (grey dots). The black lines are added to support visual interpretation and show the relative uncertainty of each parameter.~~ SIT denotes the surface intensity threshold. ~~PFT-dependent parameters are grouped with the same color.~~

of three ($3.63 \cdot 10^{-6}$). The mortality related parameters ~~F and rCK~~ ~~F and rCK~~ led in the NI_{optim} version both to a decrease in the fire-related mortality for TrBE and an increase for TrBR PFTs. The optimized parameters for VPD_{optim} led to a decrease in the ~~fire-related~~ ~~fire-related~~ mortality for both PFTs ~~except for the TrBR rCK, which led to an increased mortality.~~

The relative uncertainties were for most optimized parameters very small (between 0 and ~~0.05~~ ~~10~~%), hence these parameters were strongly constrained (Fig. 6). Just the fire-mortality related parameters (~~F and rCK~~ ~~F and rCK~~) had large uncertainties for the TrBR, hence ~~were~~ weakly constrained. For VPD_{optim} the uncertainty of ~~rCK~~ ~~rCK~~ (TrBR) was 0.8 and for NI_{optim} the uncertainty of ~~F~~ ~~F~~ was 0.9 and for ~~rCK~~ ~~rCK~~ 1 (TrBR).

~~The decrease in the model error (cost) due to the optimization process has been mainly due to improvement in the burnt area. While for the NI_{optim} the cost of the burnt area dataset improved by 81%, the cost of the biomass dataset improved just by 6%. In case of the VPD_{optim} the cost of the burnt area dataset improved by 49%, whereas the biomass dataset improved by 19% (Fig. A5).~~

Table 3. Comparison of the results for AGB and the TrBE PFT cover in terms of NMSE and the Willmott coefficient of agreement between NI_{orig} , VPD_{optim} and NI_{optim} in South America (SA) ~~and~~, in the Cerrado ~~and in the Caatinga~~.

Region	NMSE	Willmott
AGB		
SA (NI_{orig})	0.97	0.83
SA (VPD_{optim})	0.91	0.84
SA (NI_{optim})	0.99	0.83
Cerrado (NI_{orig})	15.06	0.25
Cerrado (VPD_{optim})	12.36	0.28
Cerrado (NI_{optim})	16.06	0.24
<u>Caatinga (NI_{orig})</u>	<u>11.93</u>	<u>0.32</u>
<u>Caatinga (VPD_{optim})</u>	<u>8.57</u>	<u>0.36</u>
<u>Caatinga (NI_{optim})</u>	<u>10.44</u>	<u>0.33</u>
FPC - Evergreen <u>Evergreen</u>		
(TrBE)		
SA (NI_{orig})	0.42	0.82
SA (VPD_{optim})	0.41	0.82
SA (NI_{optim})	0.43	0.81
Cerrado (NI_{orig})	1.04	0.60
Cerrado (VPD_{optim})	0.70	0.64
Cerrado (NI_{optim})	1.40	0.55
<u>Caatinga (NI_{orig})</u>	<u>1.73</u>	<u>0.40</u>
<u>Caatinga (VPD_{optim})</u>	<u>1.54</u>	<u>0.29</u>
<u>Caatinga (NI_{optim})</u>	<u>2.05</u>	<u>0.44</u>

3.3 Model evaluation for South America

The modelled above-ground biomass (AGB) of trees in South America was throughout all model versions larger than the evaluation data set indicates (Fig. 7). Especially the biomass in the Amazon region is with an average of ca. 20 kgC/m² about one third overestimated. The drier savanna regions on the continent yielded a biomass of ca. 5-10 kgC/m², which also constitutes an overestimation in wide parts of the Cerrado and the Caatinga biome (evaluation data shows between 1-5 kgC/m², also see Roitman et al., 2018).

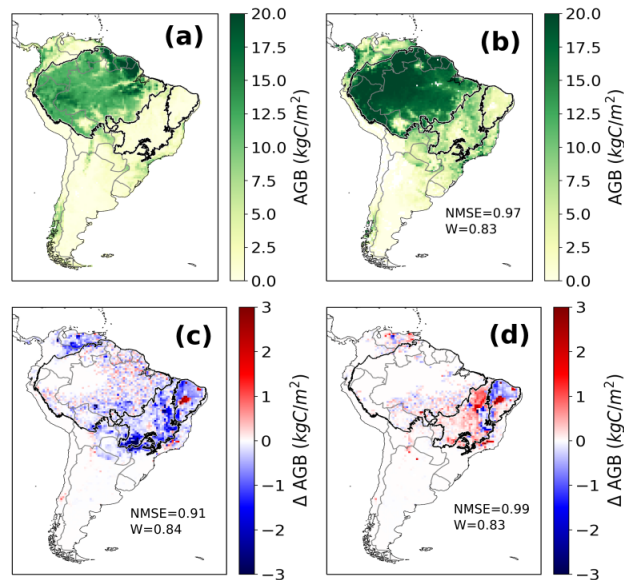


Figure 7. Annual above ground biomass (AGB) of trees over a mean from 2005-2015 in kgC/m². (a) Avitabile evaluation data. (b) Simulated AGB by LPJmL4-SPITFIRE in the NI_{orig} version. (c) Difference between VPD_{optim} and NI_{orig}. (d) Difference between NI_{optim} and NI_{orig}. Red (blue) color indicates a larger (smaller) biomass after the optimization.

The differences among the different model versions are marginal: The VPD_{optim} version had the best performance compared to the evaluation data set (NMSE= 0.91, W=0.84), the NI_{orig} version had the second best performance (NMSE=0.97, W=0.84) and the NI_{optim} the worst performance (NMSE=0.99, W=0.83). The model optimization scheme focuses on fire parameters, hence the model performance for AGB can only improve in fire-prone biomes, i.e. not in, e.g., wet tropical forest where fire is absent areas, where the fire occurrence has been modelled poorly and the vegetation-fire interactions have improved due to the optimization process. For example in the center of the Amazon rainforest almost no fire is found in the evaluation data nor is simulated. Hence no improvement of burnt area as well as AGB can be achieved. On the other hand, in regions where the modelling error of burnt area is now reduced, this can also improve simulated AGB, hence vegetation-fire interactions. In the fire-prone Caatinga and Cerrado the VPD_{optim} version mostly decreased the biomass by up to 3 kgC/m², showing a better performance compared to the evaluation data set (e.g. in the Cerrado the NMSE decreased from from 15.06 to 12.36 in the VPD_{optim} version compared to NI_{orig}, see Tab. 3).

The modelled foliage projective cover (FPC) showed for all three model versions a strong underestimation compared to the evaluation data set of the TrBE throughout the whole Amazonian region (ca. 50% compared to ca. 100% in the evaluation dataset). In the fire-prone biomes Cerrado and Caatinga, however, the TrBE PFT was sometimes overestimated (TrBE cover between 0 and 40 %, Fig. 8). In the regions with less TrBE the dominant PFT was mostly TrBR (Cerrado) or TrH (Caatinga) (see Fig. A1 and A2).

NI_{optim} led to an overall decrease in the model performance also in terms of FPC the TrBE distribution, as both, the NMSE

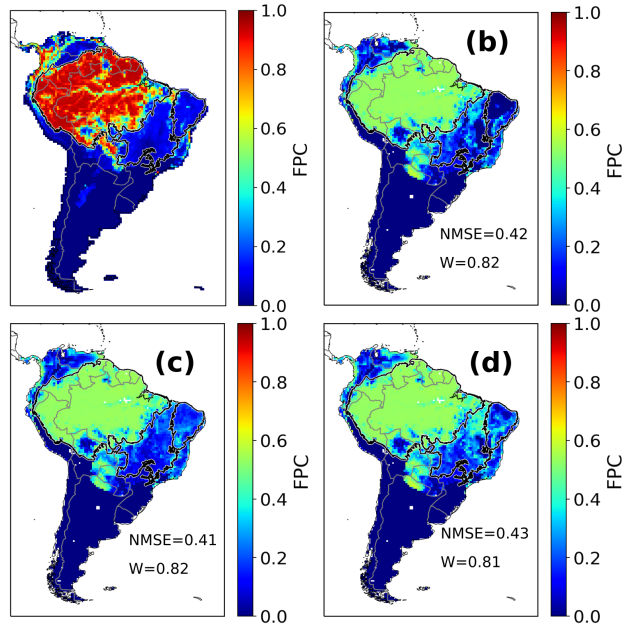


Figure 8. Annual FPC cover by tropical broadleaved evergreen PFT over a mean from 2005-2015 as fraction per cell. (a) ESA-CCI evaluation data (b) Simulated FPC by LPJmL4-SPITFIRE using the NI_{orig} version (c) Simulated FPC by LPJmL4-SPITFIRE using the VPD_{optim} version (d) Simulated FPC by LPJmL4-SPITFIRE using the NI_{optim} version

and the Willmott coefficient ~~worsened~~ declined compared to NI_{orig} (NI_{orig} : NMSE=0.42, W=0.82; NI_{optim} : NMSE=0.43, W=0.81).

The VPD_{optim} version, on the other hand, showed an slightly improved ~~FPC-TrBE~~ distribution (NMSE=0.41, W=0.82) but also ~~here we got in this case we obtained~~ an even larger improvement, when only the fire-prone regions Cerrado or Caatinga are considered (Tab. 3). Also for the TrBR and TrH PFT distributions the optimization lead to an improved performance using the VPD_{optim} in the Caatinga and Cerrado, whereas the PFT distribution in the Amazon remained similar to the prior PFT distribution. In the NI_{optim} version, parameter optimization only slightly reduced TrBR cover showing a worse performance compared to VPD_{optim} . However, herbaceous cover changed only slightly in all optimization experiments (Fig. A1 and A2).

4 Discussion

10 In summary, our results show that the implementation of a new fire danger index based on the water vapor pressure deficit FDI_{VPD} and its optimization against satellite data sets improved the simulations of fire in LPJmL4-SPITFIRE, both in terms of spatial patterns as well as temporal dynamics of burnt area. In the following, we discuss the model improvements, limitations and recommendations for future improvements of process-based global fire models within the DGVM framework.

4.1 Improvements in model performance

The VPD results showed a better model performance for fire in the spatial dimension, as well as in the temporal dimension (Tab. 1 and 3). Compared to the Nesterov index, FDI_{VPD} uses additional climate input as relative humidity and precipitation. In the calculation of the Nesterov index precipitation is just used as a threshold. This leads to a better accounting of the very different climatic conditions among various biomes. Furthermore, the FDI_{VPD} includes a direct representation of the vegetation density. The ~~significante~~ significance of this has been recently shown by findings of Forkel et al. (2019a) who have emphasized the importance of past plant productivity and fuel production for burnt area. This is ~~particular~~ particularly important for differentiating between fires in biomes with similar PFT distribution. For example, the vegetation density is much ~~higher~~ larger in the Cerrado, even though the Caatinga and Cerrado have a similar modelled PFT composition. ~~A larger VPD provides more fire,~~ which provides more fuel and therefore leads to a higher fire danger.

While the seasonal and interannual variability in the Caatinga has improved largely using the FDI_{VPD} (NMSE decreased by a factor of ca. 20), the improvement in the Cerrado was relatively small (NMSE decreased by ca. 10%). This is due to the fact, that the optimization tries to obtain a compromise between the different optimized cells. As the model performance was originally much better for the Cerrado, the largest improvement could be achieved for the Caatinga. We have also chosen a large amount of cells in the Caatinga, because the model performance was here ~~particular~~ particularly bad. This leads to a large improvement in the time series of the Caatinga region, while the improvement for the Cerrado was less significant. With the Nesterov index fire was strongly underestimated in the Amazonia region, while the optimized VPD increases the modelled burnt area. The fire is only present at the edges of the Amazon (both in model and observation, see Fig. 4), where tree density is lower and deforestation takes place. In the closed continuous forest area towards the center of the Amazon almost no fire is observed and also not simulated.

Another result of the optimizing procedure, using FDI_{VPD} , was the improvement of the PFT distribution and the above-ground biomass of trees especially in the fire-prone biomes Caatinga and Cerrado (Fig. 8). ~~The~~ For example, the central Amazon, where fire is a scarce event, shows almost no changes compared to the non-optimized model version. Here, it is the improvement of the vegetation model itself, and not the fire module, which can help to improve the model performance of LPJmL4-SPITFIRE. Hence, it emphasizes that ~~three-parameter-sets-determining-we-need-to-include-further-parameters-in-the-optimization,~~ which impact directly the PFT distribution, biomass and fire ~~are-required-in-the-optimization~~ to obtain a ~~realistic~~ significant improvement in the spatial and temporal distribution of both, vegetation and fire. ~~It-becomes-especially-visible~~ However, this study focused solely on the parameters within the SPITFIRE module. Due to the focus on fire related parameter, the cost of the burnt area dataset decreases much more than the cost of the biomass dataset (Fig. A5). Hence we only get a substantial improvement in model performance in semi-arid, ~~hence~~ fire-prone biomes, where vegetation dynamics and fire are strongly coupled.

During the optimization-process most of the optimized parameters were well constrained, except for the mortality-related parameters for the TrBR PFT (Fig. 6). The TrBR PFT is dominant in the fire-prone regions, where the mortality-related parameters have a large impact on vegetation dynamics. Hence, they impact multiple LPJmL routines, which are responsible for the

PFT distribution and carbon cycling. This leads in turn to a less certain parameter estimation. In order to better constrain these parameters also the optimization of vegetation model parameters would be ~~an approach necessary~~ to decrease the uncertainties. The fire danger index scaling factors (α_{NI} and α_{VPD}) convert the quantified fire risk (NI or VPD) into the actual fire danger (FDI). Both scaling factors thus set the magnitude of the fire danger for the different PFTs. Hence they impact directly the fire spread, burnt area and the number of fires as well as indirectly fire mortality. These very important parameters vary significantly for the different PFTs. TrH has the smallest scaling factor in case of both FDIs, which leads to a lower fire danger compared to the other PFTs. This indicates a prior overestimation of the fire danger of grass in tropical South America, as grasslands are generally parametrized to have a low fire resistance and moisture content and can hence burn very easily. This overestimation, compared to tree PFTs has been decreased by the optimization. In case of the VPD also the TrBR is scaled by a much smaller factor than the TrBE, which leads to a lower fire danger index. This is due to the fact, that the TrBR is dominant in dry and fire-prone regions, which experience frequent fires. Here the burnt area was often overestimated by SPITFIRE (e.g. Caatinga or eastern Cerrado) and is now decreased. On the other hand, a larger FDI for the TrBE allows more fire in wetter regions at the edge between the Cerrado and the Amazon rainforests, where TrBE is more dominant. The mortality risk of TrBE for VPD_{optim} remains close to the prior value of 1, confirming previous assumptions about its fire sensitivity. Whereas the rCK for TrBR increased to 0.48, close to the upper boundary, meaning that a mortality risk of 50% when the full crown is scorched and a 7% mortality risk when 50% of the crown is scorched, which makes the TrBR less resistant against crown damage than before. Due to this changes the overestimation of biomass in the original model for the Cerrado/Caatinga region decreased (see Fig. 7).

4.2 Limitations ~~during of~~ the optimization process

20 Generally, optimizing a model against burned area is challenging because 1) of the skewed statistical distribution of burned area and 2) because temporal or spatial mismatches in simulated burning can cause large model-data errors. These issues can be avoided with the choice of an appropriate cost function. For example, squared-error metrics tend to underestimate the variance of burned area in comparison to, e.g., the Kling-Gupta efficiency as it has been shown in the optimization of an empirical model for burned area (see Tab. A3 in Forkel et al. (2017)). Here, the optimum parameter set for the Nesterov index-based model resulted in almost no fires across South America. Thereby the optimization algorithm tries to decrease the model error by tending towards a conservative 'no fire strategy' for all biomes. This result nicely demonstrates the need to evaluate model optimization results against spatially and temporally independent data and independent variables (Keenan et al., 2011).
The Nesterov index is not able to capture fire variability within the Caatinga as well as the Cerrado at the same time. This shows that the difference in the PFT distribution between these two biomes is not adequately modelled by LPJmL or just using PFT dependent scaling factors did not sufficiently improve the model performance when using the Nesterov index. On the other hand, using the VPD fire danger index reduced the model error for burned area in both biomes, by improving the modelled performance for the Caatinga and maintaining the good performance of the Cerrado region. Since improved performance of the fire model mainly had minor effect on improving FPC of the tropical PFTs, the presented optimization scheme has to go along with process-based improvements in both, in the fire and in the vegetation modules of LPJmL.

Fire largely depends on the vegetation type and their associated flammability, fire tolerance and mortality. Hence an accurately modelled vegetation distribution is crucial for a good model performance in terms of burnt area and fire effects (Forkel et al., 2019a; Rogers et al., 2015). As shown in Fig. 8, [A1 and A2](#), the modelled PFT coverage showed an equal distribution of tropical raingreen and evergreen PFTs throughout wide parts of central-northern South America. Evaluation data shows, however, an ~~evergreen-TrBE~~ dominance in the wet rainforest regions and a ~~raingreen-TrBR~~ dominance in the Cerrado and Caatinga. ~~By choosing a large amount of optimization cells in the, by NI_{orig} , strongly overestimated Caatinga region, the burned area decreased there significantly after the optimization. Because the modelled proportion of tropical evergreen trees in the Cerrado is similar to the Caatinga biome, fire occurrence in the Cerrado has also over-proportional decreased for NI_{optim} (Fig. 4). As a result the modelled fire in whole South America was rather low, which improved the performance in the Caatinga but decreased it in other regions. Even though the VPD suffers from the same limitations, its optimized version has indeed improved simulated fire in wide parts of South America (including both Caatinga and Cerrado) and led to a better overall model performance for South America in terms of burnt area, biomass and PFT distribution. The reason for this behaviour is the above discussed impact of vegetation density and the more comprehensive representation of the climate impact. This emphasizes the potential of the VPD to be used as the standard fire danger index in SPITFIRE and to improve the fire modelling even further, based on a possible an improved PFT distribution. In the tropical rainforest the ~~modelled-TrBE proportion should be much larger. In the Cerrado and especially the Caatinga, however, trees suffer from water stress in the dry season and should shed their leaves to avoid mortality related to drought or growth efficiency. The resulting dominance of the TrBR PFT has a very different effect~~ TrBR proportion is overestimated, which leads to problems in the optimization procedure, since TrBR has very different effects on fire spread and is more fire-tolerant (different fuel characteristics and resulting fire intensity); ~~thus has~~. This leads to a lower fire-related mortality; which fits better to the drier and fire prone savanna-like regions (e.g. Cerrado). The poorly modelled PFT distribution also is responsible for the overestimation of the burnt area in the Amazon region. Because of the too large fraction of TrBE in the Cerrado/Caatinga region the scaling factor for this PFT is relatively high. This leads in turn to an overestimation in the Amazon region, where the fraction of the TrBE is larger. As stated before, the VPD fire danger index requires additional climate variables which are available from re-analysis data. We have chosen the GLDAS climate data set for its good performance to capture South-American climate. Because GLDAS 2.0 (covering 1948-2010) is just forced with Princeton meteorological data, while GLDAS 2.1 (2000-now) includes various observational and meteorological data, both versions result in a small offset. Since the offset is very small, the years 2000-2003 (first three years of GLDAS 2.1, before the optimization period) are enough for the model to recover from the offset and the carbon pools to return to equilibrium. To exclude the possibility that long-term trends within GLDAS 2.0 changed the modelled vegetation state significantly, we tested our optimization also just based on GLDAS 2.0 data (until 2010) and ~~just based~~ on GLDAS 2.1 data (2000-2017) only, using the same years for model spinup, optimization and evaluation. Both versions yielded similar results compared to the optimization presented in this study (results not shown).~~

Due to the fact that evaluation data are only available for the last 10-20 years, we are constrained to optimize the model in this relatively short time period. In South America these years were subject to an unusual high amount of severe droughts and other extreme events (Panisset et al., 2017). As a result, an optimization in this period could lead to a worse model per-

formance in a period with less pronounced droughts. This is due to the non-linear relationship between the drought signal in the input data set and the resulting modelled biosphere ~~behaviour~~behavior. Nonetheless, we were able to improve the inter-annual variability and hence, the model performance ~~during extreme years~~ to a great extent for the Caatinga and slightly for the Cerrado and ~~Caatinga regions (e. g. for 2007/2008, Fig. 5).~~ Amazon regions (Fig. 5 and A3). The Cerrado already had a ~~very good modelling performance before the optimization process, which now only slightly improved. The performance of the interannual and seasonal variability of burnt area for total South America improved substantially (Fig. A3).~~ The optimized SPITFIRE is now ~~able to model~~ better able to simulate accurately the climate dependent seasonal and interannual variability as well as the spatial extent of fire on natural land throughout the fire-prone woodlands of South America.

~~Systematic optimizations within a model-data integration setup of fire models which are embedded in a DGVM are still very rare. Previously, Rabin et al. (2018) optimized the fire model FINAL.1 within the land-surface model LM3. Our study differs from Rabin et al. (2018) in the conceptual design of the vegetation-fire models and the optimization process. While LM3 has been run on a 2° longitude by 2.5° latitude, which is much coarser than LPJmL with 0.5° by 0.5°. This difference allows us to account for a locally better climate input, vegetation and fire interaction. While FINAL.1 is a process-based model, many calculations (e.g. the fire spread routine) are done by multiplying the important factors and fitting the resulting values to observational data. SPITFIRE tries to model the important fire variables by simulating the underlying processes, and by taking the influence of climate and the different fire ignitions into account. An advantage of FINAL.1 is the inclusion of agricultural fires based on a statistical approach. Whereas Rabin et al. (2018) used a local search algorithm (Levenberg-Marquardt algorithm) to optimize their model, we used a global search algorithm (genetic optimization). Local search algorithms depend on the chosen initial parameter sets and might eventually end up in a local optimum. A genetic optimization algorithm allows to explore the full parameter space and hence gives a higher chance to find the global optimum. However, local search algorithms require less iterations than global search algorithms (300 in Rabin et al. (2018) vs. 16000 in our study). Forkel et al. (2014) tested the optimization of LPJmL with different optimization algorithms and found that it was not feasible to optimize LPJmL with a local search algorithm. Rabin et al. (2018) ran the model during the optimization process only for the period of 1991-2009, whereas we made complete model runs including 5000 years of spinup in order to get a model equilibrium for each tested parameter combination.~~

4.3 Outlook – the way ahead in improving ~~Limitations of fire modules modelling in DGVMs~~LPJmL4-SPITFIRE

In fire-prone regions the interactions between fire and vegetation dynamics ~~is-are~~ strong, hence ~~are~~ posing a challenge for global fire models embedded in DGVMs. By just focussing on fire-related ~~parameter~~parameters, an optimization approach can only to a certain extent ~~additionally~~ improve PFT distribution and simulated biomass. For a good fire representation e.g. in the Cerrado and Caatinga, a shrub PFT could further improve the model performance. Most fires in this region occur, where shrub PFTs are abundant. LPJmL tries to account for this by establishing rather small raingreen PFTs as a shrub replacement. A much better option would be a ~~seperate~~separate shrub PFT with parameters leading to a high flammability, but also a low fire mortality. An optimization of LPJmL4-SPITFIRE, including shrub PFTs could yield ~~much~~ better results than shown in this study.

~~Even though Fire models embedded in DGVMs should build on a FDI which is complex enough to account for various fire dynamics, while it's parameterization should be simple enough to be accurately applied on a global scale. While the VPD is more complex than the Nesterov index and takes into account more climatic input, it would be possible to use an even more comprehensive fire danger index (e.g. Canadian Fire Weather Index; Wagner et al., 1987) or different fire danger indices for different biomes: as the Nesterov index, it is still relatively easy to implement in a global fire model.~~

There are various other fire danger indices used for modelling purposes, as well as real fire danger assessment and fire forecast purposes. For example, fire-prone countries have developed their own fire danger indices (e.g. Canada, Australia), which are suited to the unique local fire regimes and vegetation dynamics. In a global modelling approach, however, we need to find one fire danger index, which suits best for all regions of the world and has a relatively easy implementation to decrease computational cost and the number of input data sets (which might be unavailable or uncertain).

Currently, SPITFIRE does not account for fire in managed land like cropland or managed grassland. We accounted for this by excluding cropland fires from the evaluation burnt area data set. We do, however, not account for the proportion of grassland, which is used for e.g. cattle ranching. Since in SPITFIRE fire is not enabled on ~~managed grassland~~ pasture, our results show a slightly smaller ~~fire amplitude burnt area~~ throughout South America than could be expected with managed land included and hence also compared to the GFED4 evaluation data set. ~~The~~ This effect is however small, because pasture lands cover a substantial fraction only in very few grid cells (e.g. southern Cerrado; Parente et al., 2017). Fire on managed land is generally difficult to predict in a DGVM because the reason and timing of using fire ~~on managed land~~ depends less on climatic factors but mostly on social and political decisions, ~~which vary among different countries, as well as different regions within e.g. Brazil. With fires on managed land included we would expect to further improve the simulated interannual variability, which can vary between countries, regions and localities. We expect further improvement of model performance especially in regions of large land-use areas with fires on pastures included (e.g. Rabin et al., 2018; Pfeiffer et al., 2013).~~

5 Conclusions

We ~~have demonstrated a major improvement of~~ significantly improved the fire representation within LPJmL4-SPITFIRE, applied for South America, by implementing a new fire danger index and applying a model-data integration setup to optimize fire-related parameters. We improved the seasonal and interannual variability, as well as the spatial pattern of burnt area in South America. In addition, modelling of related vegetation variables, e.g., the biomass and the PFT distribution in the fire-prone Cerrado and Caatinga biomes have also been improved. ~~To our knowledge this is the first study that attempts to systematically optimize a process-based fire model within a complex DGVM. E.g. Rabin et al. (2018) did an optimization of a fire model, which is not embedded in a DGVM.~~

~~Optimizing fire parameter~~ Optimizing fire parameters has its limits due to error propagation of the PFT distribution and hence their fire traits influencing simulated fire spread and ~~behaviour~~ behavior. Furthermore, it remains a challenge to find a fire danger index that is physically interpretable and can be applied globally. In this study, the parameter-optimization by using FDI_{NI} led to a large underestimation of fire and a generally worse model performance, when focusing on the Cerrado and Caatinga

biome. However, implementing the more complex FDI_{VPD} and optimizing it thereafter, led to an improved model performance compared to the original SPITFIRE implementation for South America. Our results demonstrate that the improvement of model processes, as well as a systematic model-data optimization are required in order to obtain a more accurate fire representation within complex DGVMs, where observations or experimental evidence to constraint fire parameter ~~is~~are scarce.

- 5 This work highlights the potential for future model-data integration approaches to obtain a better fire model performance in a global setting, based on improved vegetation dynamics within LPJmL4. ~~A realistic representation of fire is also crucial for fire-vegetation-climate feedbacks and is hence necessary for DGVMs coupled within an comprehensive Earth system model.~~

Code availability. Upon request. With publication of this article the model code will be published at <https://github.com/PIK-LPJmL>, analogously to Schaphoff et al. (2018a). We will publish the code of LPJmLmdi along with the model code on the github page of LPJmL:

- 10 <https://github.com/PIK-LPJmL/LPJmLmdi>

Appendix A

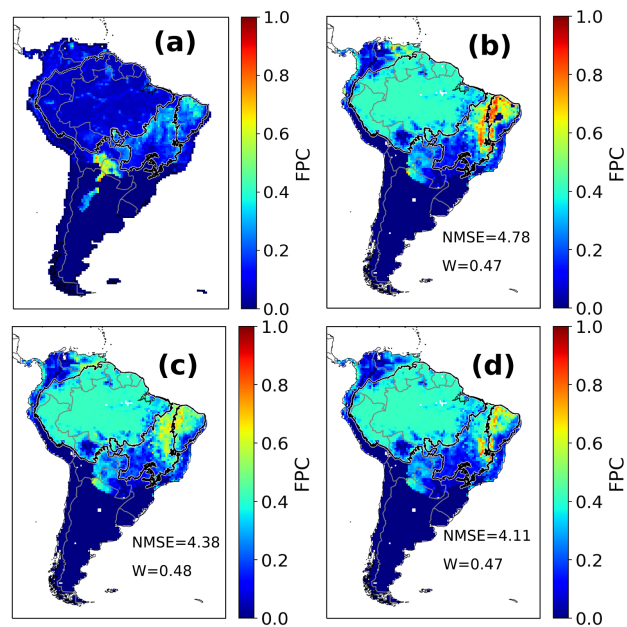


Figure A1. Annual FPC cover by tropical broadleaved raingreen PFT over a mean from 2005-2015 as fraction per cell. (a) ESA-CCI evaluation data (b) Simulated FPC by LPJmL4-SPITFIRE using the NI_{orig} version (c) Simulated FPC by LPJmL4-SPITFIRE using the VPD_{optim} version (d) Simulated FPC by LPJmL4-SPITFIRE using the NI_{optim} version

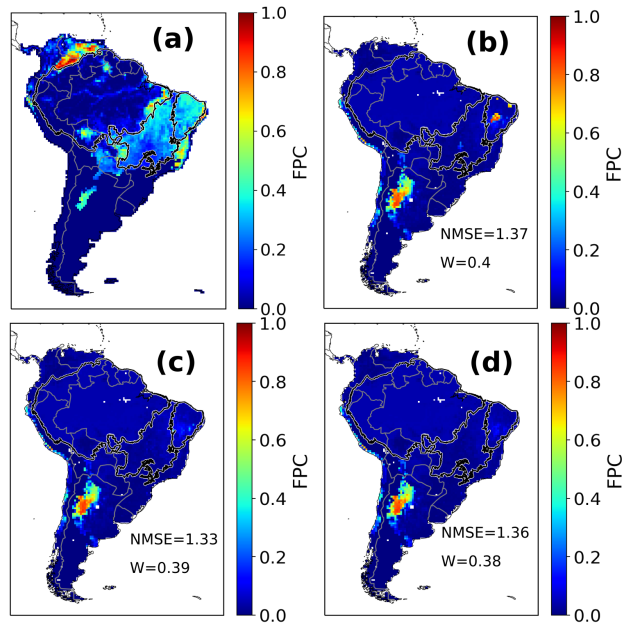


Figure A2. Annual FPC cover by tropical herbaceous PFT over a mean from 2005-2015 as fraction per cell. (a) ESA-CCI evaluation data (b) Simulated FPC by LPJmL4-SPITFIRE using the NI_{eria} version (c) Simulated FPC by LPJmL4-SPITFIRE using the VPD_{optim} version (d) Simulated FPC by LPJmL4-SPITFIRE using the NI_{optim} version

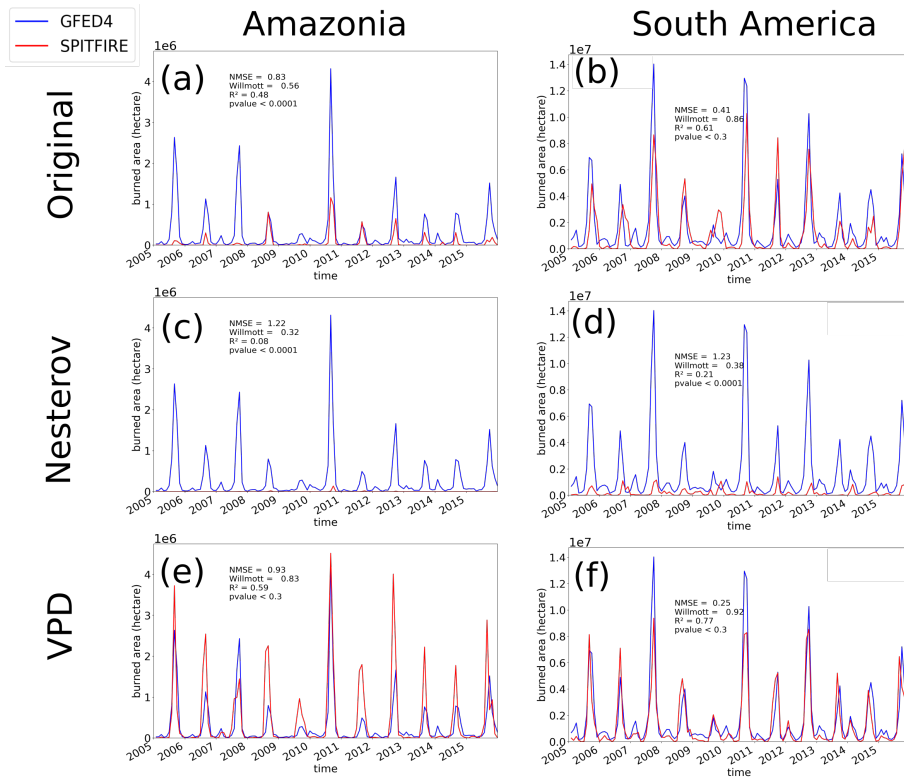


Figure A3. Time-series of monthly burnt area from 2005 - 2015 simulated by SPITFIRE (red lines) compared to GFED4 evaluation data (blue lines) for: (a) The Amazonia region, using NI_{orig} . (b) Total South America, using the NI_{orig} . (c) The Amazonia region, using NI_{optim} . (d) Total South America, using NI_{optim} . (e) The Amazonia region, using VPD_{optim} . (f) Total South America, using VPD_{optim} .

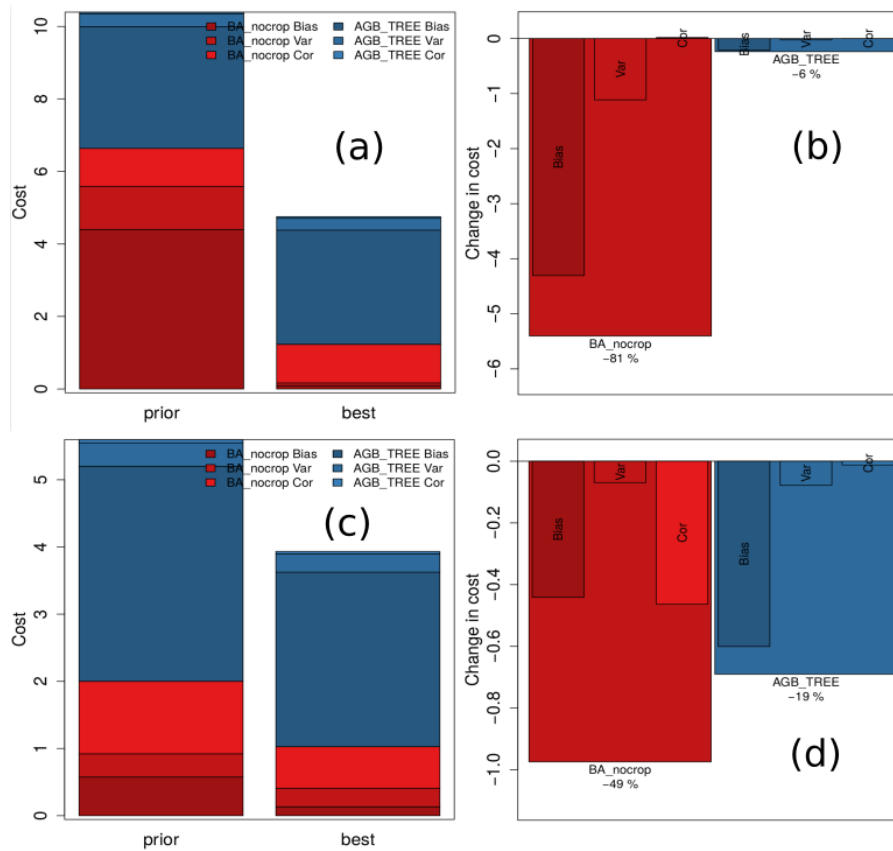


Figure A4. Cost reduction of the burnt area and the biomass during the optimization process, by showing the various components of the cost that are related to model-data bias, variance ratio and correlation. The cost for burnt area for NI_{optim} decreased by ca. 81%, whereas the cost of the biomass only decreases by ca. 6% (a and b). For VPD_{optim} the cost decreased by ca. 48% for burnt area and ca. 19% for the biomass (c and d). Hence the impact of the optimization process on burnt area is much larger due to the focus on fire parameters.

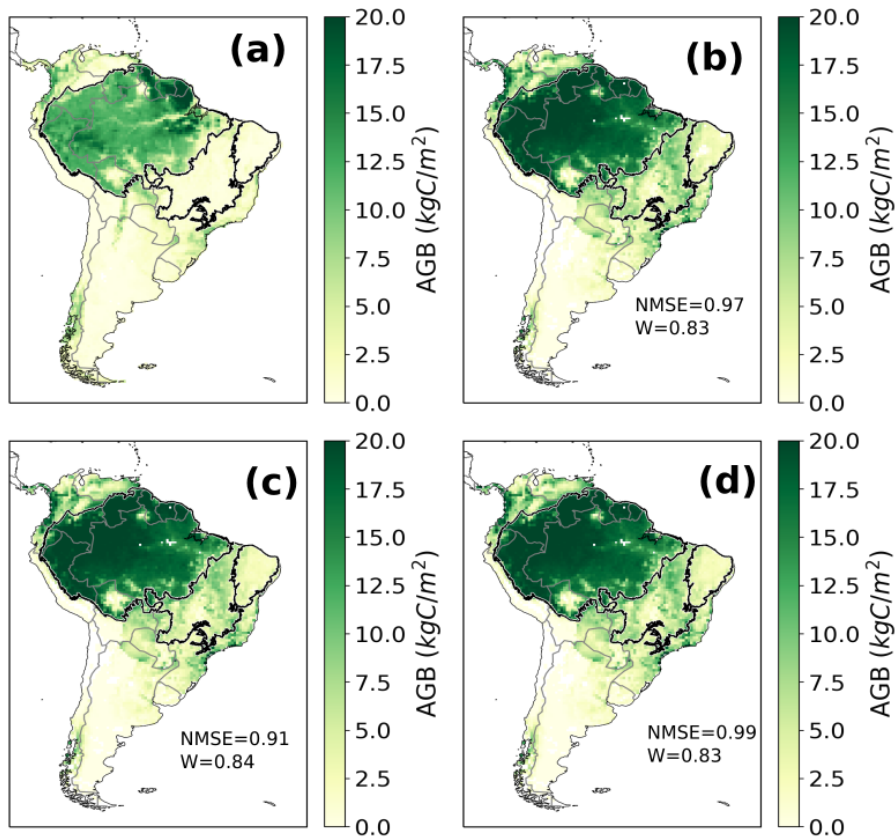


Figure A5. Annual above ground biomass (AGB) of trees over a mean from 2005-2015 in kgC/m^2 . (a) Avitabile evaluation data. (b) Simulated AGB by LPJmL4-SPITFIRE in the NI_{orig} version. (c) Simulated AGB by LPJmL4-SPITFIRE in the VPD_{optim} version. (d) Simulated AGB by LPJmL4-SPITFIRE in the NI_{optim} version.

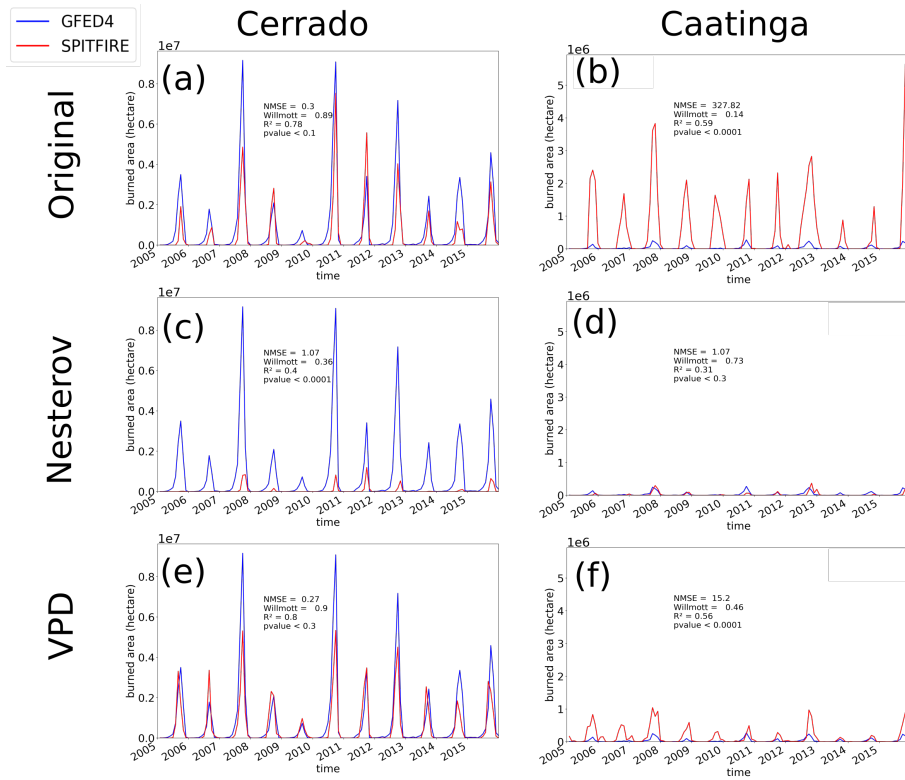


Figure A6. Time-series of monthly burnt area from 2005 - 2015 simulated by SPITFIRE (red lines) compared to GFED4 evaluation data (blue lines) for: (a) The Cerrado region, using NI_{orig} . (b) The Caatinga region, using the NI_{orig} . (c) The Cerrado region, using NI_{optim} . (d) The Caatinga region, using NI_{optim} . (e) The Cerrado region, using VPD_{optim} . (f) The Caatinga region, using VPD_{optim} .

Author contributions. MD, MF and KT designed the study in discussion with MC, MB, BS and JK. MD and MF implemented the model-data integration framework for LPJmL4. MD and WvB implemented the new fire danger index. MD performed the analysis with inputs from MF. MD wrote the paper with inputs from all Co-authors.

Competing interests. The authors declare that they have no conflict of interest.

- 5 *Acknowledgements.* This paper was developed within the scope of the IRTG 1740 / TRP 2015/50122-0, funded by the DFG / FAPESP (MD und KT). MC acknowledges the support from the projects FAPESP 2015/50122-0 (São Paulo Research Foundation), and CNPq 314016/2009-0 (Brazilian National Council for Scientific and Technological Development). MF acknowledges funding through the TU Wien Wissenschaftspreis 2015. MB acknowledges the support of the Brazilian Research Network on Global Climate Change (Rede Clima) and of the National Institute of Science and Technology for Climate Change Phase 2 under CNPq, Grant 465501/2014-1, FAPESP, Grant 2014/50848-9 and the

National Coordination for High Level Education and Training (CAPES) Grant, 16/2014. KT and BS acknowledge funding from the BMBF- and Belmont Forum-funded project “CLIMAX: Climate Services Through Knowledge Co-Production: A Euro-South American Initiative For Strengthening Societal Adaptation Response to Extreme Events”, Grant no. 01LP1610A.

References

- Alvares, C. A., Stape, J. L., Sentelhas, P. C., de Moraes Gonçalves, J. L., and Sparovek, G.: Köppen's climate classification map for Brazil, *Meteorol. Z.*, 22, 711–728, <https://doi.org/10.1127/0941-2948/2013/0507>, 2013.
- Andela, N., Morton, D. C., Giglio, L., Chen, Y., van der Werf, G. R., Kasibhatla, P. S., DeFries, R. S., Collatz, G. J., Hantson, S., Kloster, S.,
5 Bachelet, D., Forrest, M., Lasslop, G., Li, F., Mangeon, S., Melton, J. R., Yue, C., and Randerson, J. T.: A human-driven decline in global burned area, *Science*, 356, 1356–1362, <https://doi.org/10.1126/science.aal4108>, 2017.
- Arpaci, A., Eastaugh, C. S., and Vacik, H.: Selecting the best performing fire weather indices for Austrian ecoregions, *Theor. Appl. Climatol.*, 114, 393–406, <https://doi.org/10.1007/s00704-013-0839-7>, 2013.
- Avitabile, V., Herold, M., Heuvelink, G. B. M., Lewis, S. L., Phillips, O. L., Asner, G. P., Armston, J., Ashton, P. S., Banin, L., Bayol, N.,
10 Berry, N. J., Boeckx, P., de Jong, B. H. J., DeVries, B., Girardin, C. A. J., Kearsley, E., Lindsell, J. A., Lopez-Gonzalez, G., Lucas, R., Malhi, Y., Morel, A., Mitchard, E. T. A., Nagy, L., Qie, L., Quinones, M. J., Ryan, C. M., Ferry, S. J. W., Sunderland, T., Laurin, G. V., Gatti, R. C., Valentini, R., Verbeeck, H., Wijaya, A., and Willcock, S.: An integrated pan-tropical biomass map using multiple reference datasets, *Global Change Biol.*, 22, 1406–1420, <https://doi.org/10.1111/gcb.13139>, 2016.
- Beuchle, R., Grecchi, R. C., Shimabukuro, Y. E., Seliger, R., Eva, H. D., Sano, E., and Achard, F.: Land cover changes in the Brazilian
15 Cerrado and Caatinga biomes from 1990 to 2010 based on a systematic remote sensing sampling approach, *Appl. Geogr.*, 58, 116–127, <https://doi.org/10.1016/j.apgeog.2015.01.017>, 2015.
- Bondeau, A., Smith, P. C., Zaehle, S., Schaphoff, S., Lucht, W., Cramer, W., Gerten, D., Lotze-Campen, H., Müller, C., Reichstein, M., and Smith, B.: Modelling the role of agriculture for the 20th century global terrestrial carbon balance, *Global Change Biol.*, 13, 679–706, <https://doi.org/10.1111/j.1365-2486.2006.01305.x>, 2007.
- 20 Chambers, J. Q. and Artaxo, P.: Biosphere–atmosphere interactions: Deforestation size influences rainfall, *Nat. Clim. Change*, 7, 175, <https://doi.org/10.1038/nclimate3238>, 2017.
- Christian, H. J., Blakeslee, R. J., Boccippio, D. J., Boeck, W. L., Buechler, D. E., Driscoll, K. T., Goodman, S. J., Hall, J. M., Koshak, W. J., Mach, D. M., and Stewart, M. F.: Global frequency and distribution of lightning as observed from space by the Optical Transient Detector, *J. Geophys. Res. Atmos.*, 108, ACL4–1–ACL4–15, <https://doi.org/10.1029/2002JD002347>, 2003.
- 25 Chuvieco, E., Aguado, I., Yebra, M., Nieto, H., Salas, J., Martín, M. P., Vilar, L., Martínez, J., Martín, S., Ibarra, P., de la Riva, J., Baeza, J., Rodríguez, F., Molina, J. R., Herrera, M. A., and Zamora, R.: Development of a framework for fire risk assessment using remote sensing and geographic information system technologies, *Ecol. Modell.*, 221, 46–58, <https://doi.org/10.1016/j.ecolmodel.2008.11.017>, 2010.
- Civil, O. and Environmental Engineering/Princeton University, D.: Global Meteorological Forcing Dataset for Land Surface Modeling, UCAR/NCAR - Research Data Archive, <https://rda.ucar.edu/datasets/ds314.0>, 2006.
- 30 Cochrane, M. A. and Laurance, W. F.: Synergisms among Fire, Land Use, and Climate Change in the Amazon, *Ambio*, 37, 522–527, <https://doi.org/10.2307/25547943>, 2008.
- Fader, M., Rost, S., Müller, C., Bondeau, A., and Gerten, D.: Virtual water content of temperate cereals and maize: Present and potential future patterns, *J. Hydrol.*, 384, 218–231, <https://doi.org/10.1016/j.jhydrol.2009.12.011>, 2010.
- Forkel, M., Carvalhais, N., Schaphoff, S., V. Bloh, W., Migliavacca, M., Thurner, M., and Thonicke, K.: Identifying environmental controls
35 on vegetation greenness phenology through model–data integration, *Biogeosciences*, 11, 7025–7050, <https://doi.org/10.5194/bg-11-7025-2014>, 2014.

- Forkel, M., Dorigo, W., Lasslop, G., Teubner, I., Chuvieco, E., and Thonicke, K.: A data-driven approach to identify controls on global fire activity from satellite and climate observations (SOFIA V1), *Geosci. Model Dev.*, 10, 4443–4476, <https://doi.org/10.5194/gmd-10-4443-2017>, 2017.
- Forkel, M., Andela, N., Harrison, S. P., Lasslop, G., Marle, M. v., Chuvieco, E., Dorigo, W., Forrest, M., Hantson, S., Heil, A., Li, F., Melton, J., Sitch, S., Yue, C., and Arneith, A.: Emergent relationships with respect to burned area in global satellite observations and fire-enabled vegetation models, *Biogeosciences*, 16, 57–76, <https://doi.org/10.5194/bg-16-57-2019>, 2019a.
- Forkel, M., Dorigo, W., Lasslop, G., Chuvieco, E., Hantson, S., Heil, A., Teubner, I., Thonicke, K., and Harrison, S. P.: Recent global and regional trends in burned area and their compensating environmental controls, *Environ. Res. Commun.*, 1, 051005, <https://doi.org/10.1088/2515-7620/ab25d2>, 2019b.
- 10 Gerten, D., Schaphoff, S., Haberlandt, U., Lucht, W., and Sitch, S.: Terrestrial vegetation and water balance—hydrological evaluation of a dynamic global vegetation model, *J. Hydrol.*, 286, 249–270, <https://doi.org/10.1016/j.jhydrol.2003.09.029>, 2004.
- Giglio, L., Randerson, J. T., and van der Werf, G. R.: Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4), *J. Geophys. Res. Biogeosci.*, 118, 317–328, <https://doi.org/10.1002/jgrg.20042>, 2013.
- Goff, J. and Gratch, S.: List 1947, Smithsonian meteorological tables, *Transactions of the American Society of Ventilation Engineering*, 52, 15 95, 1946.
- Goldewijk, K. K., Beusen, A., van Drecht, G., and de Vos, M.: The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years, *Global Ecol. Biogeogr.*, 20, 73–86, <https://doi.org/10.1111/j.1466-8238.2010.00587.x>, 2011.
- Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, *J. Hydrol.*, 377, 80–91, <https://doi.org/10.1016/j.jhydrol.2009.08.003>, 2009.
- 20 Hantson, S., Arneith, A., Harrison, S. P., Kelley, D. I., Prentice, I. C., Rabin, S. S., Archibald, S., Mouillot, F., Arnold, S. R., Artaxo, P., Bachelet, D., Ciais, P., Forrest, M., Friedlingstein, P., Hickler, T., Kaplan, J. O., Kloster, S., Knorr, W., Lasslop, G., Li, F., Melton, J. R., Meyn, A., Sitch, S., Spessa, A., van der Werf, G. R., Voulgarakis, A., and Yue, C.: The status and challenge of global fire modelling, Copernicus Publications, <http://centaur.reading.ac.uk/53498>, 2016.
- Harvard: Harvard WorldMap, <http://worldmap.harvard.edu>, [Online; accessed 27. Mar. 2019], 2019.
- 25 Hoffmann, W. A., Jackson, R. B., Hoffmann, W. A., and Jackson, R. B.: Vegetation–Climate Feedbacks in the Conversion of Tropical Savanna to Grassland, [https://doi.org/10.1175/1520-0442\(2000\)013<1593:VCFITC>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<1593:VCFITC>2.0.CO;2), [Online; accessed 28. Mar. 2019], 2000.
- IBGE: Mapa de Biomas e de Vegetação, <https://ww2.ibge.gov.br/home/presidencia/noticias/21052004biomashtml.shtm>, [Online; accessed 14. Feb. 2019], 2019.
- Jolly, W. M., Cochrane, M. A., Freeborn, P. H., Holden, Z. A., Brown, T. J., Williamson, G. J., and Bowman, D. M. J. S.: Climate-induced variations in global wildfire danger from 1979 to 2013, *Nat. Commun.*, 6, 7537, <https://doi.org/10.1038/ncomms8537>, 2015.
- 30 Keeley, J. E., Pausas, J. G., Rundel, P. W., Bond, W. J., and Bradstock, R. A.: Fire as an evolutionary pressure shaping plant traits, *Trends Plant Sci.*, 16, 406–411, <https://doi.org/10.1016/j.tplants.2011.04.002>, 2011.
- Keenan, T. F., Carbone, M. S., Reichstein, M., and Richardson, A. D.: The model–data fusion pitfall: assuming certainty in an uncertain world, *Oecologia*, 167, 587, <https://doi.org/10.1007/s00442-011-2106-x>, 2011.
- 35 Keetch, J. J. and Byram, G. M.: A Drought Index for Forest Fire Control, Res. Pap. SE-38. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station. 35 p, 038, <https://www.fs.usda.gov/treesearch/pubs/40>, 1968.
- Kelley, D. I., Prentice, I. C., Harrison, S. P., Wang, H., Simard, M., Fisher, J. B., and Willis, K. O.: A comprehensive benchmarking system for evaluating global vegetation models, *Biogeosciences*, 10, 3313, <https://doi.org/10.5194/bg-10-3313-2013>, 2013.

- Knorr, W., Arneth, A., and Jiang, L.: Demographic controls of future global fire risk, *Nat. Clim. Change*, 6, 781, <https://doi.org/10.1038/nclimate2999>, 2016.
- Krawchuk, M. A. and Moritz, M. A.: Constraints on global fire activity vary across a resource gradient, *Ecology*, 92, 121–132, <https://doi.org/10.1890/09-1843.1>, 2011.
- 5 Lahsen, M., Bustamante, M. M. C., and Dalla-Nora, E. L.: Undervaluing and Overexploiting the Brazilian Cerrado at Our Peril, *Environment: Science and Policy for Sustainable Development*, 58, 4–15, <https://doi.org/10.1080/00139157.2016.1229537>, 2016.
- Langmann, B., Duncan, B., Textor, C., Trentmann, J., and van der Werf, G. R.: Vegetation fire emissions and their impact on air pollution and climate, *Atmos. Environ.*, 43, 107–116, <https://doi.org/10.1016/j.atmosenv.2008.09.047>, 2009.
- Lasslop, G., Thonicke, K., and Kloster, S.: SPITFIRE within the MPI Earth system model: Model development and evaluation, *J. Adv. Model. Earth Syst.*, 6, 740–755, <https://doi.org/10.1002/2013MS000284>, 2014.
- 10 Lasslop, G., Hantson, S., and Kloster, S.: Influence of wind speed on the global variability of burned fraction: a global fire model’s perspective, *Int. J. Wildland Fire*, 24, 989–1000, <https://doi.org/10.1071/WF15052>, 2015.
- Lasslop, G., Brovkin, V., Reick, C. H., Bathiany, S., and Kloster, S.: Multiple stable states of tree cover in a global land surface model due to a fire-vegetation feedback, *Geophys. Res. Lett.*, 43, 6324–6331, <https://doi.org/10.1002/2016GL069365>, 2016.
- 15 Li, W., MacBean, N., Ciais, P., Defourny, P., Lamarche, C., Bontemps, S., Houghton, R. A., and Peng, S.: Gross and net land cover changes in the main plant functional types derived from the annual ESA CCI land cover maps (1992–2015), *Earth Syst. Sci. Data*, 10, 219–234, <https://doi.org/10.5194/essd-10-219-2018>, 2018.
- Mebane, Jr., W. R. and Sekhon, J. S.: Genetic Optimization Using Derivatives: The rgenoud Package for R, *Journal of Statistical Software*, 42, 1–26, <https://doi.org/10.18637/jss.v042.i11>, 2011.
- 20 Moreira de Araújo, F., Ferreira, L. G., and Arantes, A. E.: Distribution Patterns of Burned Areas in the Brazilian Biomes: An Analysis Based on Satellite Data for the 2002–2010 Period, *Remote Sens.*, 4, 1929–1946, <https://doi.org/10.3390/rs4071929>, 2012.
- Nachtergaele, F. O., van Velthuisen, H. T., and Verelst, L.: Harmonized World Soil Database, <http://pure.iiasa.ac.at/id/eprint/8958>, [Online; accessed 4. Oct. 2018], 2009.
- Panisset, J. S., Libonati, R., Gouveia, C. M. P., Machado-Silva, F., França, D. A., França, J. R. A., and Peres, L. F.: Contrasting patterns of the extreme drought episodes of 2005, 2010 and 2015 in the Amazon Basin, *Int. J. Climatol.*, 38, 1096–1104, <https://doi.org/10.1002/joc.5224>, 2017.
- 25 Parente, L., Ferreira, L., Faria, A., Nogueira, S., Araújo, F., Teixeira, L., and Hagen, S.: Monitoring the Brazilian pasturelands: A new mapping approach based on the Landsat 8 spectral and temporal domains, *Int. J. Appl. Earth Obs. Geoinf.*, 62, 135–143, <https://doi.org/10.1016/j.jag.2017.06.003>, 2017.
- 30 Pechony, O. and Shindell, D. T.: Fire parameterization on a global scale, *J. Geophys. Res. Atmos.*, 114, <https://doi.org/10.1029/2009JD011927>, 2009.
- Pfeiffer, M., Spessa, A., and Kaplan, J. O.: A model for global biomass burning in preindustrial time: LPJ-LMfire (v1.0), *Geosci. Model Dev.*, 6, 643–685, <https://doi.org/10.5194/gmd-6-643-2013>, 2013.
- Prado, D. E.: As caatingas da América do Sul, *Ecologia e conservação da Caatinga*, 2, 3–74, 2003.
- 35 Pyne, S. J., Andrews, P. L., and Laven, R. D.: Introduction to wildland fire, 2nd edition revised, DigitalCommons@USU, <https://digitalcommons.usu.edu/barkbeetles/135>, 1996.
- Quéré, C. L., Moriarty, R., Andrew, R. M., Canadell, J. G., Sitch, S., Korsbakken, J. I., Friedlingstein, P., Peters, G. P., Andres, R. J., Boden, T. A., Houghton, R. A., House, J. I., Keeling, R. F., Tans, P., Arneth, A., Bakker, D. C. E., Barbero, L., Bopp, L., Chang, J., Chevallier, F.,

- Chini, L. P., Ciaia, P., Fader, M., Feely, R. A., Gkritzalis, T., Harris, I., Hauck, J., Ilyina, T., Jain, A. K., Kato, E., Kitidis, V., Goldewijk, K. K., Koven, C., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lima, I. D., Metzl, N., Millero, F., Munro, D. R., Murata, A., Nabel, J. E. M. S., Nakaoka, S., Nojiri, Y., O'Brien, K., Olsen, A., Ono, T., Pérez, F. F., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Rödenbeck, C., Saito, S., Schuster, U., Schwinger, J., Séférian, R., Steinhoff, T., Stocker, B. D., Sutton, A. J., Takahashi, T., Tilbrook, B., van der Laan-Luijkx, I. T., van der Werf, G. R., van Heuven, S., Vandemark, D., Viovy, N., Wiltshire, A., Zaehle, S., and Zeng, N.: Global Carbon Budget 2015, *Earth Syst. Sci. Data*, 7, 349–396, <https://doi.org/10.5194/essd-7-349-2015>, 2015.
- 5 Rabin, S. S., Ward, D. S., Malyshev, S. L., Magi, B. I., Shevliakova, E., and Pacala, S. W.: A fire model with distinct crop, pasture, and non-agricultural burning: use of new data and a model-fitting algorithm for FINAL.1, *Geosci. Model Dev.*, 11, 815–842, <https://doi.org/10.5194/gmd-11-815-2018>, 2018.
- 10 Ray, D., Nepstad, D., and Moutinho, P.: MICROMETEOROLOGICAL AND CANOPY CONTROLS OF FIRE SUSCEPTIBILITY IN A FORESTED AMAZON LANDSCAPE, *Ecol. Appl.*, 15, 1664–1678, <https://doi.org/10.1890/05-0404>, 2005.
- Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., Toll, D., Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and Toll, D.: The Global Land Data Assimilation System, *Bull. Am. Meteorol. Soc.*, <https://doi.org/10.1175/BAMS-85-3-381>, 2004.
- 15 Rogers, B. M., Soja, A. J., Goulden, M. L., and Randerson, J. T.: Influence of tree species on continental differences in boreal fires and climate feedbacks, *Nat. Geosci.*, 8, 228, <https://doi.org/10.1038/ngeo2352>, 2015.
- Roitman, I., Bustamante, M. M. C., Haidar, R. F., Shimbo, J. Z., Abdala, G. C., Eiten, G., Fagg, C. W., Felfili, M. C., Felfili, J. M., Jacobson, T. K. B., Lindoso, G. S., Keller, M., Lenza, E., Miranda, S. C., Pinto, J. R. R., Rodrigues, A. A., Delitti, W. B. C., Roitman, P., and Sampaio, J. M.: Optimizing biomass estimates of savanna woodland at different spatial scales in the Brazilian Cerrado: Re-evaluating allometric equations and environmental influences, *PLoS One*, 13, e0196742, <https://doi.org/10.1371/journal.pone.0196742>, 2018.
- 20 Rothermel, R. C.: A mathematical model for predicting fire spread in wildland fuels, *DigitalCommons@USU, Res. Pap. INT-115*, 1, <https://digitalcommons.usu.edu/barkbeetles/438>, 1972.
- Schaphoff, S., Heyder, U., Ostberg, S., Gerten, D., Heinke, J., and Lucht, W.: Contribution of permafrost soils to the global carbon budget, *Environ. Res. Lett.*, 8, 014026, <https://doi.org/10.1088/1748-9326/8/1/014026>, 2013.
- 25 Schaphoff, S., Bloh, W. v., Rammig, A., Thonicke, K., Biemans, H., Forkel, M., Gerten, D., Heinke, J., Jägermeyr, J., Knauer, J., Langerwisch, F., Lucht, W., Müller, C., Rolinski, S., and Waha, K.: LPJmL4 – a dynamic global vegetation model with managed land – Part 1: Model description, *Geosci. Model Dev.*, 11, 1343–1375, <https://doi.org/10.5194/gmd-11-1343-2018>, 2018a.
- Schaphoff, S., Forkel, M., Müller, C., Knauer, J., Bloh, W. v., Gerten, D., Jägermeyr, J., Lucht, W., Rammig, A., Thonicke, K., and Waha, K.: LPJmL4 – a dynamic global vegetation model with managed land – Part 2: Model evaluation, *Geosci. Model Dev.*, 11, 1377–1403, <https://doi.org/10.5194/gmd-11-1377-2018>, 2018b.
- 30 Seager, R., Hooks, A., Williams, A. P., Cook, B., Nakamura, J., Henderson, N., Seager, R., Hooks, A., Williams, A. P., Cook, B., Nakamura, J., and Henderson, N.: Climatology, Variability, and Trends in the U.S. Vapor Pressure Deficit, an Important Fire-Related Meteorological Quantity*, <https://journals.ametsoc.org/doi/abs/10.1175/JAMC-D-14-0321.1>, [Online; accessed 19. Oct. 2018], 2015.
- 35 Sedano, F. and Randerson, J.: Multi-scale influence of vapor pressure deficit on fire ignition and spread in boreal forest ecosystems, <https://doi.org/10.5194/bg-11-3739-2014>, [Online; accessed 11. Oct. 2018], 2014.

- Silvério, D. V., Brando, P. M., Balch, J. K., Putz, F. E., Nepstad, D. C., Oliveira-Santos, C., and Bustamante, M. M. C.: Testing the Amazon savannization hypothesis: fire effects on invasion of a neotropical forest by native cerrado and exotic pasture grasses, *Philosophical Transactions of the Royal Society B: Biological Sciences*, <https://royalsocietypublishing.org/doi/full/10.1098/rstb.2012.0427>, 2013.
- Sitch, S., Smith, B., Prentice, I. C., Arneeth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K., and Venevsky, S.: Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model, *Global Change Biol.*, 9, 161–185, <https://doi.org/10.1046/j.1365-2486.2003.00569.x>, 2003.
- Thonicke, K., Venevsky, S., Sitch, S., and Cramer, W.: The role of fire disturbance for global vegetation dynamics: coupling fire into a Dynamic Global Vegetation Model, *Global Ecol. Biogeogr.*, 10, 661–677, <https://doi.org/10.1046/j.1466-822X.2001.00175.x>, 2001.
- Thonicke, K., Spessa, A., Prentice, I. C., Harrison, S. P., Dong, L., and Carmona-Moreno, C.: The influence of vegetation, fire spread and fire behaviour on biomass burning and trace gas emissions: results from a process-based model, *Biogeosciences*, 7, 1991–2011, <https://doi.org/10.5194/bg-7-1991-2010>, 2010.
- van der Werf, G. R., Dempewolf, J., Trigg, S. N., Randerson, J. T., Kasibhatla, P. S., Giglio, L., Murdiyarso, D., Peters, W., Morton, D. C., Collatz, G. J., Dolman, A. J., and DeFries, R. S.: Climate regulation of fire emissions and deforestation in equatorial Asia, *Proc. Natl. Acad. Sci. U.S.A.*, 105, 20 350–20 355, <https://doi.org/10.1073/pnas.0803375105>, 2008.
- Venevsky, S., Thonicke, K., Sitch, S., and Cramer, W.: Simulating fire regimes in human-dominated ecosystems: Iberian Peninsula case study, *Global Change Biol.*, 8, 984–998, <https://doi.org/10.1046/j.1365-2486.2002.00528.x>, 2002.
- Wagner, C. E. V., Forest, P., Station, E., Ontario, C. R., Francais, R. U. E., and Davis, H. J.: Development and Structure of the Canadian Forest Fire Weather Index System, *Can. For. Serv., Forestry Tech. Rep.*, 1987.
- Werf, G. R. v. d., Randerson, J. T., Giglio, L., Collatz, G. J., Mu, M., Kasibhatla, P. S., Morton, D. C., DeFries, R. S., Jin, Y., and Leeuwen, T. T. v.: Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009), *Atmos. Chem. Phys.*, 10, 11 707–11 735, <https://doi.org/10.5194/acp-10-11707-2010>, 2010.
- Werf, G. R. v. d., Randerson, J. T., Giglio, L., Leeuwen, T. T. v., Chen, Y., Rogers, B. M., Mu, M., Marle, M. J. E. v., Morton, D. C., Collatz, G. J., Yokelson, R. J., and Kasibhatla, P. S.: Global fire emissions estimates during 1997–2016, *Earth Syst. Sci. Data*, 9, 697–720, <https://doi.org/10.5194/essd-9-697-2017>, 2017.
- Willmott, C. J.: Some Comments on the Evaluation of Model Performance, *Bull. Am. Meteorol. Soc.*, [https://doi.org/10.1175/1520-0477\(1982\)063<1309:SCOTEO>2.0.CO;2](https://doi.org/10.1175/1520-0477(1982)063<1309:SCOTEO>2.0.CO;2), 1982.
- Wilson, R. A.: A reexamination of fire spread in free-burning porous fuel beds /, <http://www.sidalc.net/cgi-bin/wxis.exe/?IsisScript=COLPOS.xis&method=post&formato=2&cantidad=1&expresion=mnf=001388>, [Online; accessed 4. Oct. 2018], 1982.
- Yue, X. and Unger, N.: Fire air pollution reduces global terrestrial productivity, *Nat. Commun.*, 9, 5413, <https://doi.org/10.1038/s41467-018-07921-4>, 2018.