Response to Reviews

We thank the two reviewers for their constructive comments to improve the manuscript. Their comments are reproduced below with our responses in blue.

Reviewer #1

This paper proposed a machine learning method to predict gridded monthly wildfire burned area during 2002-2015 over the South Central United States and identify the relative importance of the predicted factors on the burned area for both the winter-spring and summer fire seasons. The method is able to alleviate the problem of unevenly-distributed burned area data due to the grid-level resolution. The result is interesting and constructive to some extent. The authors said the machine learning method can achieve the R2 value of ~0.4. However, this result is hard to say it as a high accuracy. The authors can consider to compare more machine learning algorithms, such as AdaBoost, XGBoost. The low accuracy will also affect the reliability of the importance of predicted factors. Therefore, I would recommend a major revision.
We appreciate the feedback from the reviewer. It is easily understood that the prediction accuracy of wildfire burned areas will increase as the spatial and temporal resolution of the prediction model decreases. In comparison to previous studies that predict wildfire burned areas, our work is unique in two aspects: (1) our prediction performance (R^2 of ~0.4) was reported at the 0.5° x 0.5°-grid scale rather than over an aggregated spatial domain; (2) we did not exclude unburned grids or small burned grids from the prediction. Since very few published studies had both of these two features, we can only compare our results with the prior studies predicting burned area at similar spatial scales as our work (see the Table R1 below). These studies have R^2 values all below 0.3, despite using a coarser spatial resolution (~100 km x 100 km) and sometimes also a coarser temporal resolution (annually). Considering the complexity of wildfires and intrinsic nature of unevenly distributed burned area, we argue that R^2 value of around 0.4 from our work is a significant improvement over those previously published studies. Furthermore, the machine learning approaches developed by our work were motivated by the need to reduce the uneven distribution of burned area data so as to achieve a higher prediction accuracy.

We tested other boosting methods as suggested by the reviewer and they did not achieve significantly better results (see our response to comment #11). In addition, at this stage we do not foresee physical explanations to adopt these boosting methods. To address the reviewer's concern, we have rewritten the comparison of our work with others in the revised manuscript and elaborated in more detail how much better our model is compared with previously published works.

Table R1. Studies using statistical methods to estimate burned area at spatial scales similar as this study

Region	**period**	**Method**	**Spatial scale**	**Temporal scale**	**R^2**	**References**
Spain	1990-2008	MARS	25 km x 25 km ~100 km x 100 km	Monthly	Median correlation R=0.29 (R^2~0.08)	Bedia et al. (2014)
EU-Mediterranean	1985-2011	MLR	~108 km x 108 km	Annual	Median R^2=0.28	Urbieta et al. (2015)
Pacific western coast of USA	1985-2011	MLR	~108 km x 108 km	Annual	Median R^2=0.22	Urbieta et al. (2015)

More detailed discussion is presented below in the specific comment #11.

Regarding the reliability of the inferred importance of predicted factors, our results are robust based on the optimized model with the best results and the given set of predictor variables. To verify this, we also conducted 50 times 10-fold cross-validation by randomizing the order of all the data each time. The ranks of the variables by the median values of %IncMSE are identical to the ranks of the variables in our initial results. This indicates that feature importance identified by the random forest model is reliable and stable. We have included the above discussions in the revised manuscript.

1. P1L25: have been seen increasing?

We have changed the sentence to "many regions of the world have experienced an increase in frequency and intensity of wildfires …" (line 27).

2. P2L51: You said the machine learning methods were used to estimate total burned area aggregated over a large-scale domain in past studies. In this study, you focus on the grid-level resolution. Could you describe what the resolution of past studies is? Do these works have the issue of the unbalanced distributed burned area?

The resolution of past studies ranges from 25 km x 25 km to 2000 km x 2000 km, as listed in Table S1. The spatial resolution of the phytoclimatic zones in Bedia et al. (2014) ranges from 25 km x 25 km to 100 km x 100 km; the European countries in Amatulli et al. (2013) had the scales ranging from 300 km x 300 km to 1000 km x 1000 km. We have added the spatial and temporal resolution of these two studies mentioned in the main text (line 50-54).

For these studies, depending on the spatial scale, temporal scale, and wildfire characteristics (fire frequency, intensity etc.), the burned area distribution can be normal-distributed or right-skewed. Generally, studies predicting annual burned area on a country scale (spatial scale larger than 100 km x 100 km) do not have the issue of uneven-distributed burned area because the burned area are already aggregated. For example, Amatulli et al. (2013) shows the distributions of annual burned area for the countries in European Mediterranean (spatial scales ranging from 300 km x 300 km to 1000 km x 1000 km). For most of the countries, the burned area is normally distributed (Fig R1 below).

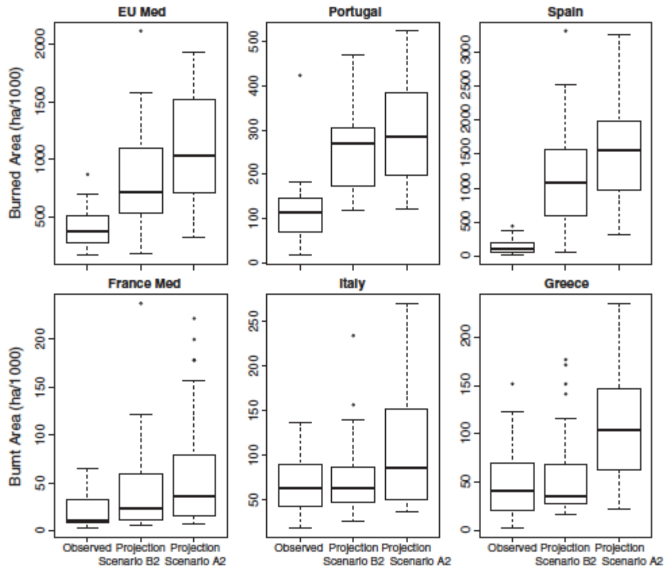

Fig. R1. Box plots of annual observed (left) and projected burned area in all study regions using the MARS models under B2 (middle) and A2 scenarios (right) for the European countries (adopted from Amatulli et al (2013); Fig. 8).

Another example of Carvalho et al. (2008) demonstrates that the distribution of burned area varies by districts in Portugal (spatial scales ranging from 25 km x 25 km to 100 km x 100 km). As Fig. R2 shows, the annual burned area distribution can be like a normal distribution for districts such as Braganca, or it can be very right-skewed for districts such as Evora and Portalegre. We also included the above example and associated discussions in the revised manuscript in line 322-325.

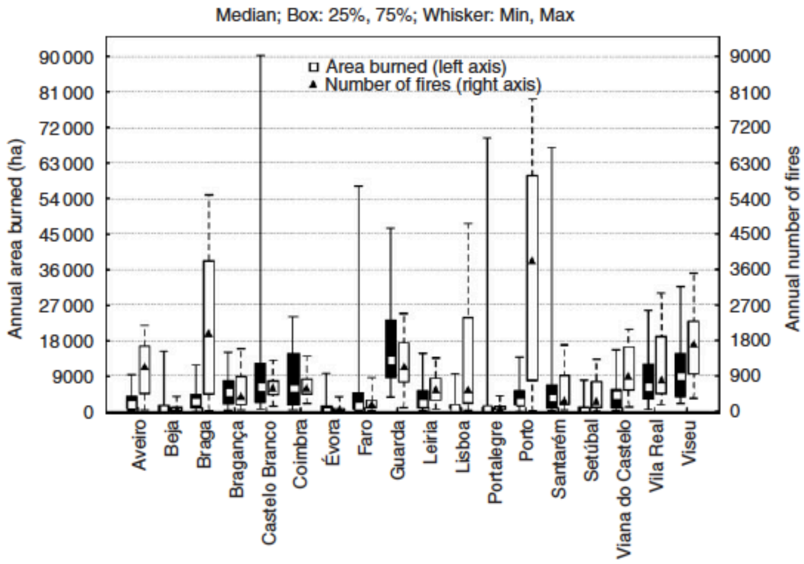

Fig. R2. Box plots of annual burned area and number of fires by Portuguese district for 1980-2004 period (adopted from Carvalho et al. (2008); Fig. 3b).

3. P3L83: The small fire is less than 10 ha, and the large fire is greater than 100 ha. However, the small fire is defined as less than 25 ha, and the large fire is greater than 150 ha in P2L58. Could you explain why they are different?
The definition of small fires at 25 ha and large fires at 150 ha was based on Steel et al. (2015) while the criterion of 10 ha was according to Yue et al. (2013). To avoid confusion and ensure the consistency, we changed the definition in line 183-186 from "10 ha" to "25 ha" and "100 ha" to "150 ha" and updated the statistics.

4. I don't think Figure 2 looks nice and it should be re-organized better. For example, the arrow between step 1 and step 2 confuses that there may be an input-output relationship. In fact, is it correct that they are independent processes?
We agree with the reviewer's concern and suggestions. We have re-organized and updated Figure 2, as shown below. Step 1 includes a quantile regression forest and step 2 includes a logistic regression with the same set of input variables. These two steps are independent processes but they are followed the order of the listed steps.

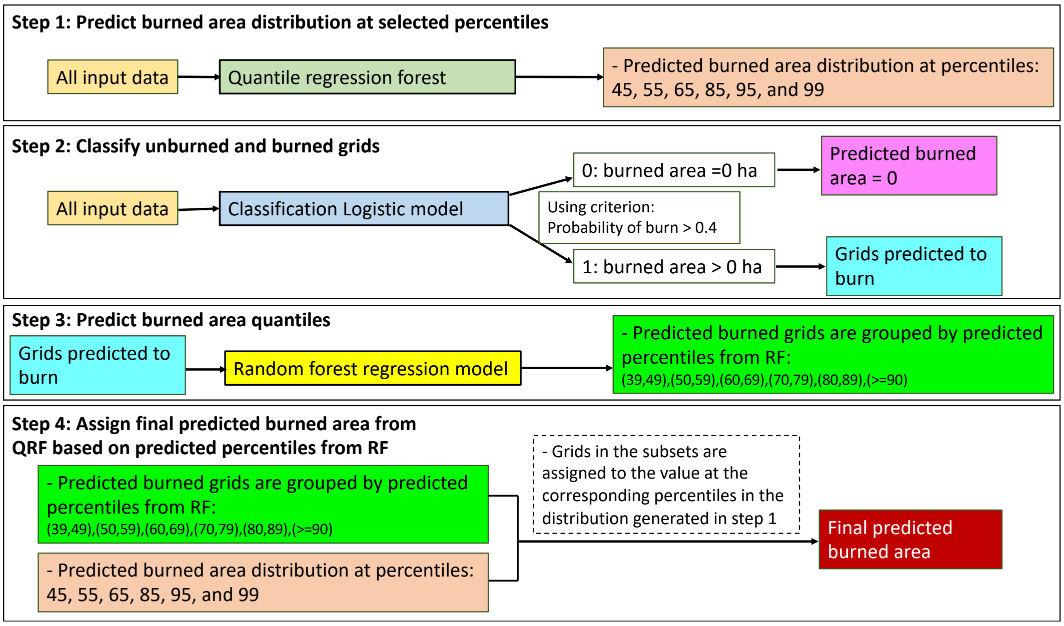

Fig. R3. Illustration of the steps in the developed model. The model includes four steps and three machine learning algorithms, including a logistic model (dark blue) classifying a grid with non-zero burned area or not, a random forest model (yellow) predicting quantiles of burned area, and a quantile regression forest (dark green) predicting conditional burned area distributions. (This figure is now Fig. 2. In the revised manuscript)

5. P3L90: The description of the four steps is not very clear. This paragraph should be rewritten. Is it correct the quantiles are the x-axis of frequency histogram? Why do you choose these quantiles? Will the pre-defined parameters induce uncertainties?

(1) Yes, the quantiles represent to the position of the predicted burned area in the cumulative distribution of all the burned area data. For example, Fig. R4 shows the empirical cumulative distribution functions of the burned area in the winter-spring fire season. The value of quantile at 0.7 can be determined by the value on the x axis, which is 3.74 (log of hectares). To better clarify the idea of quantiles, we have replaced 'quantile' with 'percentile' in the manuscript and rewrote the definition of percentile in line 192-194.

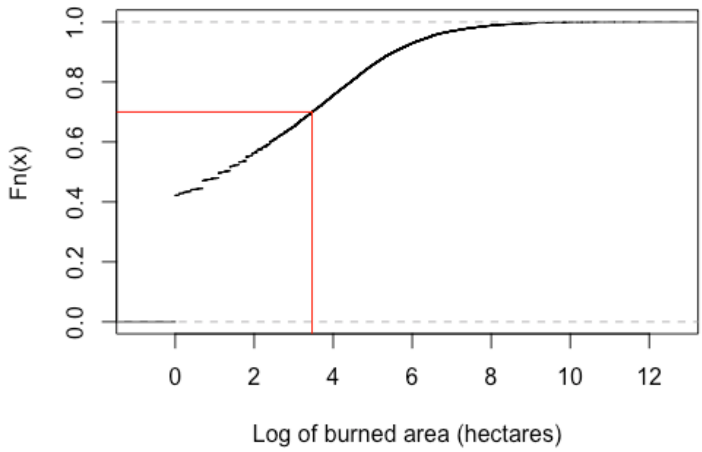

Fig. R4. The empirical cumulative distribution functions of the burned area in the winter-spring fire season. The x axis is the log of burned area and the y axis is the cumulative probability. The red lines here point the value of burned area at quantile 0.7.

(2) The quantiles were chosen to include the whole conditional distribution. The first three quantiles were selected to represent the median values between the lower and upper bounds for the first three subgroups in step 3. The last three quantiles (0.85, 0.95, and 0.99) were chosen based on the assumption that grids with larger predicted burned areas (predicted quantiles > 0.70) in the testing set will have burned area distributions that are more right-shifted than the distribution of the whole training set (Fig. R5). The quantiles were selected to reduce the model bias at the high end of burned areas. We have edited and added the above-mentioned explanations and Fig. R5 into the manuscript in line 199-205. In the revised manuscript, we have replaced 'quantile' with 'percentile'.

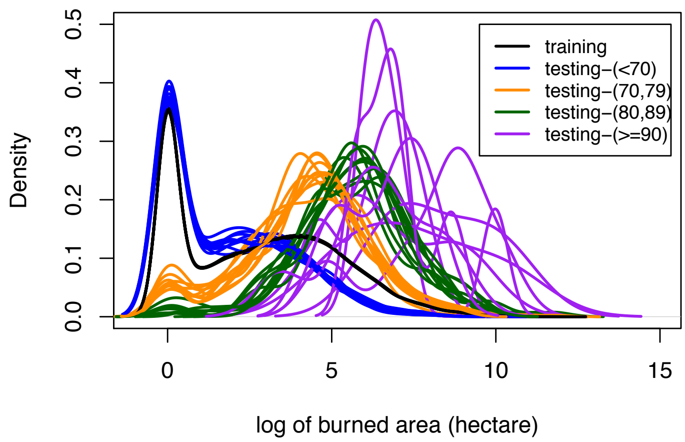

Fig. R5. Probability distribution of burned area for 10 folds of the training set (black line), testing set predicted to have percentiles less than 70 (blue), between 70 and 79 (yellow), between 80 and 89 (green), and equal to or larger than 90 (purple). (This figure is now Fig. S3. in the revised manuscript)

(3) To test the uncertainties that may be introduced by the pre-defined parameters (i.e. quantiles and subgroups), we switched the pre-defined quantiles but fixed the subgroups in the first sensitivity experiment. In this experiment, the last three quantiles were changed to the median values between a new set of lower and upper bounds, (0.75, 0.85, 0.95). As the Table R2 shows, the change of the quantiles has little effect on the overall MAE but affects the prediction of large burned areas with a smaller standard deviation in predicted values and larger MAE. Then we designed the second experiment by changing the number of subgroups, their ranges, and the corresponding quantiles. Changing both subgroups and quantiles has a marginal impact on MAE, although the standard deviation of the prediction is smaller than the results with the chosen quantiles and subgroups.

Generally, changing pre-defined parameters has little effect on overall MAE for the two fire seasons but the MAE of large burned area is larger and standard deviation of the predicted values becomes smaller. The pre-defined parameters may lead to uncertainties mostly affect the spread of the predictions and the magnitudes of large burned areas. Despite the sensitivities, the prediction model with the chosen quantiles is able to predict burned area at 0.5° x 0.5°-grid scale and achieve higher prediction accuracy compared to prior studies. Table R2 and the above discussions have been included into the manuscript (line 528-537).

Table R2. Comparison of MAE, MAE of large burned area, and standard deviation of predictions between the model with the chosen quantiles, quantile test 1, and quantile test set2 (This table is now Table S7 in the revised manuscript)

Model	**With the chosen quantiles***	**Quantile test set 1***	**Quantile test set 2***
MAE (log of area; winter-spring)	1.37	1.30	1.29
MAE of large burned area† (log of area; winter-spring)	2.13	2.64	2.81
Standard deviation of predictions (log of area; winter-spring)	2.42	2.09	2.09
MAE (log of area; summer)	1.17	1.12	1.11
MAE of large burned area† (log of area; summer)	2.25	2.42	2.52
Standard deviation of predictions (log of area; summer)	2.19	1.93	1.92

* Model developed in this study: Use the selected quantiles of 0.45, 0.55, 0.65, 0.85, 0.95, and 0.99 and six subgroups of (0.39, 0.49), (0.50, 0.59), (0.60, 0.69), (0.70, 0.79), (0.80, 0.89), (>=0.90).
* Set 1: Use the selected quantiles of 0.45, 0.55, 0.65, 0.75, 0.85, and 0.95 and six subgroups of (0.39, 0.49), (0.50, 0.59), (0.60, 0.69), (0.70, 0.79), (0.80, 0.89), (>=0.90).

* Set 2: Use the selected quantiles of 0.475, 0.63, 0.78, and 0.93 and four subgroups of (0.39, 0.55), (0.56, 0.70), (0.71, 0.85), (0.86, 1.00).

† Large burned area here is defined as the burned area larger than 90th percentile.

6. P3L90: Although the authors claim that the four steps method will alleviate the problem of uneven-distributed dataset, the multi-steps will introduce some risks. For example, if the second step wrongly classifies the burned area as the nonburned area, the bias will be amplified because it won't enter into step 3.

The reviewer is correct that biases from one step could be propagated to the subsequent steps, for example when the burned grids are predicted not to burn or when the unburned grids are predicted to burn. For the first case, when the burned grids are incorrectly predicted not to burn, the low bias is introduced because the burned grids would not proceed to step 3. For the second case, inclusion of unburned grids in step 3 may introduce a positive bias. We have included the discussions of the error propagation in section 6 (line 511-513).

To demonstrate our four-step model can achieve a higher accuracy and alleviate the issue of uneven-distributed dataset, we compare the prediction performance using random forest alone with that of the four-step model developed in this study, as shown in the Table R3 below:

Table R3. Comparison of MAE and skewness between the RF model and the developed model (The information of this table is now included into the Table S2 in the revised manuscript)

Model	**RF alone**	**Model developed in this study**
MAE (winter-spring)	1.34	1.13
Skewness (winter-spring)	37.40 (burned area)	0.70 (quantiles)
MAE (summer)	0.70	0.57
Skewness (summer)	33.83 (burned area)	0.96 (quantiles)

Skewness is a measure of the asymmetry of the probability distribution of a random variable about its mean. The skewness of a random variable X is the third standardized moment $\widetilde{\mu_3}$, defined as:

$$\widetilde{\mu_3} = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = \frac{\mu_3}{\sigma^3} = \frac{E[(X-\mu)^3]}{(E[(X-\mu)^2])^{3/2}} = \frac{\kappa_3}{\kappa_2^{3/2}}$$

where μ is the mean, σ is the standard deviation, E is the expectation operator, μ_3 is the third central moment, and κ_t are the *t*-th cumulants. If skewness is less than -1 or greater than +1, the distribution is highly skewed. If skewness is between -1 and -0.5 or between +0.5 and +1, the distribution is moderately skewed. If skewness is between -0.5 and 0.5, the distribution is approximately symmetric. The positive value indicates that the tail is on the right side of the distribution while negative value indicates that the tail is on the left.

Our model has a lower MAE, by 15% and 19% for the winter-spring and summer fire season, respectively, compared to the single RF model. The distribution of the quantiles in the developed model is more uniform than the distribution of the burned area, as shown in Figure R6 below and the skewness. We have added a discussion of this issue in the text (line 355-365) and included Table R3 in the supplementary information. The information of skewness calculation has been included in the supplementary information. Note that we have replaced ‘quantile’ with ‘percentile’ in the manuscript to better clarify the idea of quantiles.

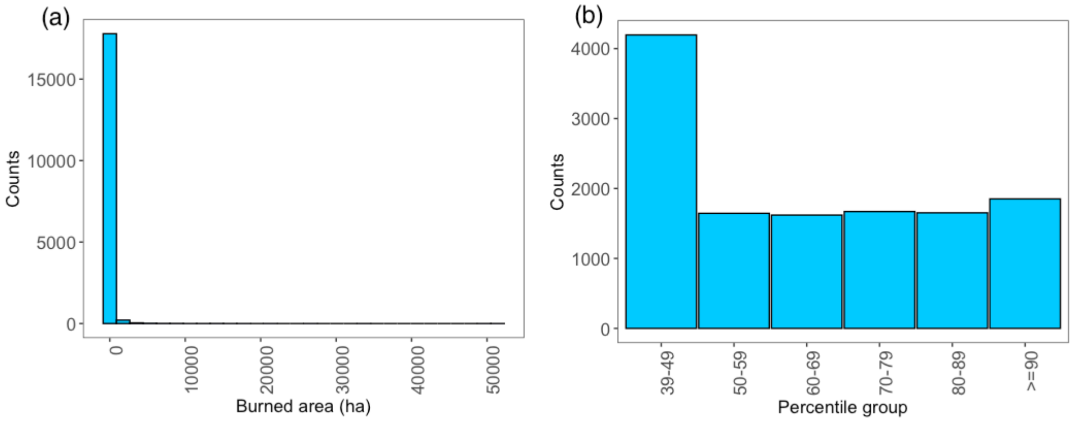

Fig. R6. (a) Histogram of burned area (b) Histogram of the percentile groups of burned area for the winter-spring fire season. (This figure is now Fig. S2. In the revised manuscript)

7. P4L126: Please explain the assumption or give the reference that grids with larger burned area will have more right-shifted burned area distribution than the distributions of the training set.

The above-mentioned Figure R5 shows the burned area distributions of training sets and testing sets categorized by predicted quantile groups. Grids that are predicted to have larger burned area (predicted quantiles larger than 0.70) have more right-shifted burned area distributions compared to the distribution of the training set. We have included the figure and explained the assumption in the manuscript (line 199-205 and Fig S3).

8. P6L171: Please add references.

We have added references in line 106-108.

9. P8L223: Please add the F1-score performance criteria because you mentioned it in L235.

We have added the description of F1-score in line 267-270.

10. P8L235: Could you please plot the AUC curve so that it could help to analyze the TP rate and FP rate. You can also analyze the F1-score performance by Precision

and Recall. This will help to understand whether the classifier is underprediction or overprediction.

The ROC curves are included in the supplementary (Fig S4), as demonstrated below.

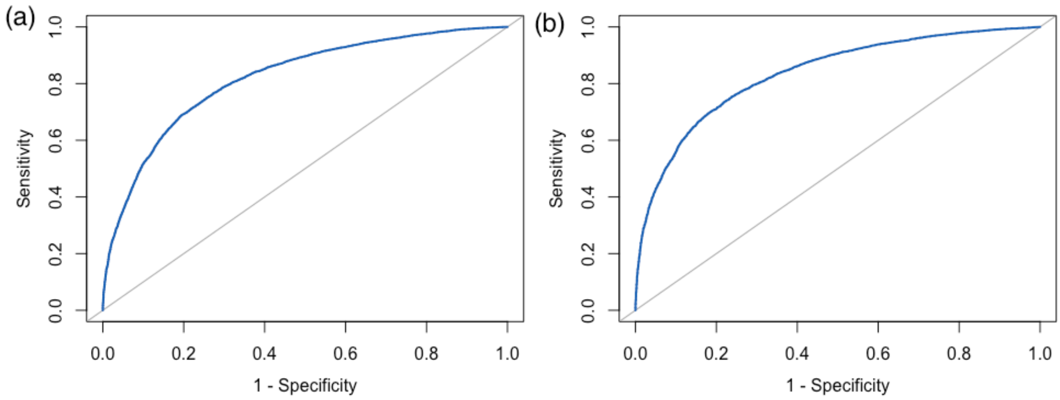

Fig. R7. The ROC curve of burning classification for the (a) winter-spring and (b) summer fire season. (This figure is now Fig. S4. In the revised manuscript)

The ROC curves show good performance of the models, given that the ROC curves are toward the upper left corner and the AUC for the two fire seasons are 0.82 and 0.83. The accuracy and F1-score are 0.74 and 0.79 for the winter-spring fire season. For the summer fire season, the accuracy and F1-score are 0.74 and 0.77. The above results demonstrate the model ability of predicting burned grids with the optimal balance of recall and precision. The values of AUC, recall, precision, and F1-score are also updated in Table 2. We have included the results and discussions of model performance in the main text (line 292-298).

11. P8L235: The performance accuracy of the classifier and the regressor in Table 2 is not very high. Typically, the F1-score of a good classifier can achieve over 0.8 and the RMSE of a good regressor is lower than 0.2. Could you compare your results with some other machine learning methods, such as Adaboost, XGBoost.

We compared our classification model results with some other machine learning methods (the parameters of each model have been optimized):

Table R4. Comparison of accuracy, recall, precision, and F1-score between the logistic regression model, RF model and XGBoost model

Model	Logistic regression (model for this study)	Random forest	XGBoost
Winter-spring fire season			
Accuracy	0.74	0.82	0.80
Recall	0.88	0.86	0.86
Precision	0.73	0.83	0.80
F1-score	0.79	0.84	0.83
Summer fire season			
Accuracy	0.74	0.81	0.81
Recall	0.84	0.82	0.82
Precision	0.71	0.82	0.81
F1-score	0.77	0.82	0.81

RF and XGboost show better performance in terms of accuracy, precision, and F1-score. However, they identify less burned grids (value of Recall) and require more parameter tuning and runtime, compared to logistic model. Considering the comparable performance, more identified burned grids, less model tuning, and less runtime, we chose logistic model as our classification model because it can be simply applied to different regions which is a competitive advantage for future applications of the prediction model.

As for the burned area prediction, we've compared our results with RF model in question 6. Here we included the model results from XGBoost, as it shows below:

Table R5. Comparison of MAE and skewness between the developed model, RF model, XGBoost model (This table is now Table S2 in the revised manuscript)

Metrics	Model developed in this study	RF alone	XGBoost alone
MAE (winter-spring)	1.13	1.34	1.26
Skewness (winter-spring)	0.70 (quantiles)	37.40 (burned area)	37.40 (burned area)
MAE (summer)	0.57	0.70	0.67
Skewness (summer)	0.96 (quantiles)	33.83 (burned area)	33.83 (burned area)

Our four-step model has a lower MAE, which deceases by 11% and 15% for the winter-spring and summer fire season, respectively, compared to the XGboost model. The developed model shows better performance in predicting burned area, compared to using RF or XGboost model. The model results from XGboost were also included in the Table S2 and corresponding discussion in the manuscript (line 355-365).

12. What does "630" mean in Table 2?
We have removed the misplaced line number 630 from Table 2.

13. P8L247: Please add some references for past studies.
The comparison of model performance to the previous studies has been rewritten and the associated references can be found in the manuscript (line 299-330).

14. P8L251: You compare your results with Chen et al. (2016) and Liu and Wimberly (2015). I wonder whether they are comparable if they are under different factors, different periods and different regions.
Since there are very few studies predicting gridded burned area directly and among them there is no study focusing on the South Central US, we chose to compare our results with these studies in terms of the approaches (i.e. excluding unburn grids or not), the temporal and spatial resolution, and the percent of variance explained by the model (i.e. R-square), regardless of their study regions, periods, and used predictors.

Although the study regions, study period, and the used predictors in this study are different from prior studies mentioned for comparison in the main text, the developed model in this study demonstrates some advantages compared to other models. First, both Chen et al. (2016) and Liu and Wimberly (2015) excluded unburned or small-burned grids when building their models, thus failing to capture the response of small fires size to predictor variables. Second, both studies focused on annual burned area in a spatial resolution of 1°x 1°, while the spatial and temporal resolution of our four-step model is finer both spatially and temporally (0.5° x 0.5° and monthly burned area). Our four-step model is able to resolve the fire-predictor relationship in a seasonal and a relatively-finer spatial scale.

We also included other studies with a similar spatial resolution for comparison. Urbieta et al. (2015) used multiple linear regression (MLR) to predict annual burned area of provinces and national forests in the southern countries of the European Union (EUMED) and Pacific Western US (PWUSA) (spatial resolution ~ 108 km x 108 km). For all the provinces/national forests, the median R^2 is 0.28 for EUMED and 0.22 for PWUSA. Carvalho et al. (2008) utilized MLR to predict monthly burned area of Portuguese districts (spatial resolution ranging from 25 km x 25 km to 100 km x 100 km) and their R^2 range from 0.43 to 0.80. Even though they achieved a better model performance for some districts, their models had poorer performance for the districts with very right-skewed burned area distribution (Figure R2 shown above), including Evora (R^2=0.43), Portalegre (R^2=0.45). Another example of Bedia et al. (2014) predicted monthly burned area of phytoclimatic zones in Spain (~ 25 km x 25 km to

100 km to 100 km) by using multivariate adaptive regression splines (MARS) and they obtained R^2 ranging from 0.01 to 0.37.

Although the model performance may vary depending on regions, fire characteristics, time scales, and predictors, the R^2 value of around 0.4 that we achieved to predict monthly burned area at a spatial resolution of 0.5°x0.5° is a significant improvement over previously published studies for burned area prediction at such spatiotemporal scale and the improvement was resulted from our efforts to alleviate the issue of unevenly-distributed burned area. We have rewritten the paragraphs to better explain the comparisons and the advantages of our models in line 299-330.

15. Please explain the meaning of the blue line in Figure 3.

The blue line is a best fit to the data by linear regression. We have added the descriptions of the blue line in the caption of Figure 3.

16. P9L281: Although the importance of Random Forest help to identify some key factors, they depend on the accuracy of the machine learning method. If the accuracy is not very high, it will reduce the reliability of the information. On the other hand, the importance can't provide how the change trend of factors affect the prediction.

(1) Based on the optimized model with the best results and the given set of predictor variables, the variable importance of predictors is reliable. Besides model accuracy, stability of the variable importance should also be considered (Han et al., 2012; He and Yu, 2010). To further ensure the variable ranking is stable, we conducted 50 times 10-fold cross-validation by randomizing the order of all the data each time. Figure R8 below shows the distributions of %IncMSE of each variable ranked by the median %IncMSE. Even though the feature importance varies a lot in different runs, the ranks by median values are identical to the variable ranks in our initial results, indicating the feature importance identified by the random forest model is stable. We have included the above discussions in the revised manuscript (line 382-385).

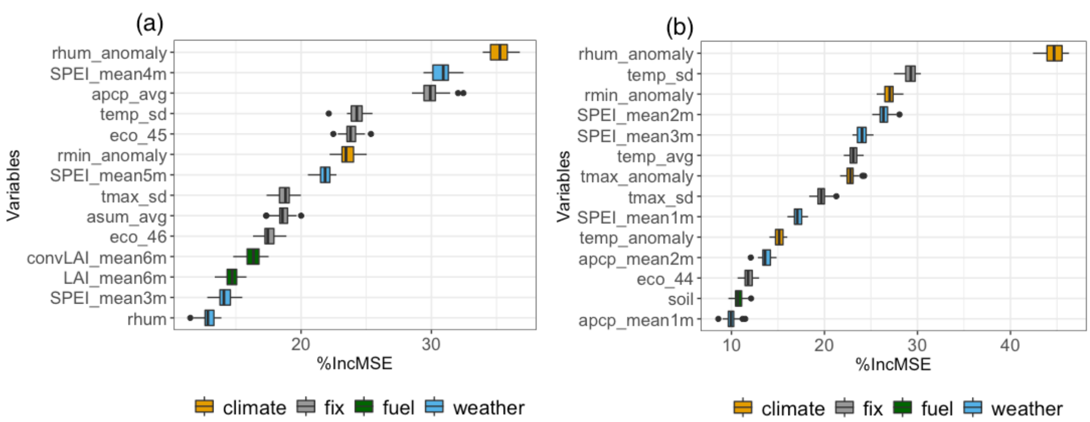

Fig. R8. Box plots of variable importance in %IncMSE from the 50 times 10-fold cross validation for (a) winter-spring and (b) summer fire season. (This figure is now Fig. S6. In the revised manuscript)

(2) Although the variable importance by RF cannot directly provide how the change trend of factors affect the prediction, like the coefficient in the linear regression, the partial dependence plots can be applied to the built model and show the marginal effect of a variable on the prediction performance (Friedman, 2001). The partial dependence plots consider a partial dependence function that is estimated by calculating averages in the training data and can be expressed as:

$$\widehat{f_{xS}}(xS) = \frac{1}{n}\sum_{i=1}^{n}\hat{f}(xS, x_C^{(i)}),$$

where xS is the feature we are interested in and $x_C^{(i)}$ are actual feature values for the features in which we are not interested. This partial function provides the average marginal effect on the prediction for given values of feature S.

Here we provide partial dependence plots for the burned area model and RH anomaly and mean SPEI of the preceding 4 months (the top two variables) for the winter-spring fire season (Fig. R9). For these two variables, there is a significant drop of fitted burned area when RH anomaly is larger than -1 and mean SPEI of the preceding 4 months larger than -0.6. The partial dependence plots demonstrate how change of a variable affects the predicted burned area. We have included the information and the above-mentioned examples in the revised manuscript (line 405-410).

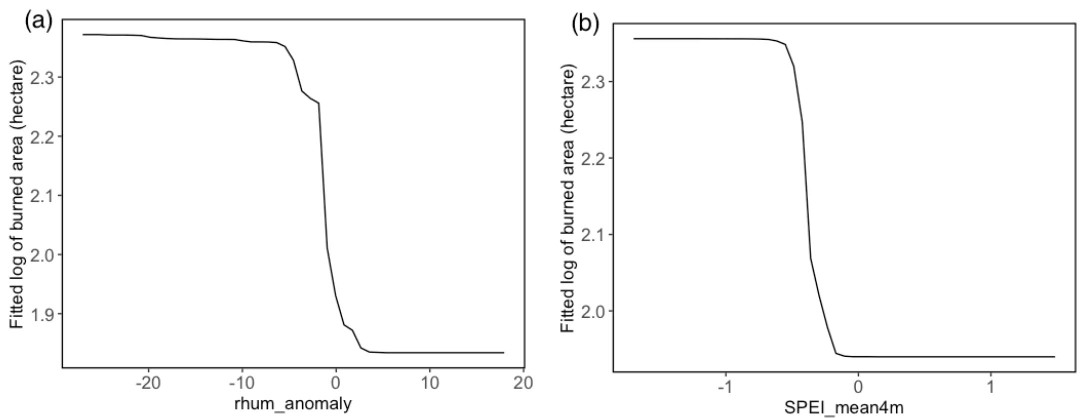

Fig. R9. Partial dependence plots for the burned area model and (a) RH anomaly and (b) mean SPEI of the preceding 4 months for the winter-spring fire season. (This figure is now Fig. S7. In the revised manuscript)

17. P9L299: I can’t find the reference “Westerling and Bryant (2008)” in the reference list.
We have included the reference “Westerling and Bryant (2008)” into the reference list.
18. There are several obvious typos in the manuscript, and the English language is poor. I think the authors should be asked to have the manuscript proofread by a native English speaker before the article can be considered for publication in a scientific journal.
We have revised the manuscript and the manuscript has been proofread by a native English speaker.

References:

Amatulli, G., Camia, A. and San-Miguel-Ayanz, J.: Estimating future burned areas under changing climate in the EU-Mediterranean countries, Science of The Total Environment, 450–451, 209–222, doi:10.1016/j.scitotenv.2013.02.014, 2013.

Bedia, J., Herrera, S. and Gutiérrez, J. M.: Assessing the predictability of fire occurrence and area burned across phytoclimatic regions in Spain, Natural Hazards and Earth System Sciences, 14(1), 53–66, doi:https://doi.org/10.5194/nhess-14-53-2014, 2014.

Carvalho, Flannigan, M., Logan, Miranda, A. and Borrego, C.: Fire activity in Portugal and its relationship to weather and the Canadian Fire Weather Index System, International Journal of Wildland Fire, 17, 328–338, doi:10.1071/WF07014, 2008.

Chen, Y., Morton, D. C., Andela, N., Giglio, L. and Randerson, J. T.: How much global burned area can be forecast on seasonal time scales using sea surface temperatures?, Environ. Res. Lett., 11(4), 045001, doi:10.1088/1748-9326/11/4/045001, 2016.

Friedman, J. H.: Greedy Function Approximation: A Gradient Boosting Machine, The Annals of Statistics, 29(5), 1189–1232, 2001.

Han, J., Kamber, M. and Pei, J.: Data Mining: concepts and techniques, Morgan Kaufmann., 2012.

He, Z. and Yu, W.: Stable feature selection for biomarker discovery, Comput Biol Chem, 34(4), 215–225, doi:10.1016/j.compbiolchem.2010.07.002, 2010.

Liu, Z. and Wimberly, M. C.: Climatic and Landscape Influences on Fire Regimes from 1984 to 2010 in the Western United States, PLOS ONE, 10(10), e0140839, doi:10.1371/journal.pone.0140839, 2015.

Steel, Z. L., Safford, H. D. and Viers, J. H.: The fire frequency-severity relationship and the legacy of fire suppression in California forests, Ecosphere, 6(1), 1–23, doi:10.1890/ES14-00224.1, 2015.

Urbieta, I. R., Zavala, G., Bedia, J., Gutierrez, J. M., San Miguel-Ayanz, J., Camia, A., Keeley, J. E. and Moreno, J. M.: Fire activity as a function of fire–weather seasonal severity and antecedent climate across spatial scales in southern Europe and Pacific western USA, Environmental Research Letters, 10(11), doi:10.1088/1748-9326/10/11/114013, 2015.

Yue, X., Mickley, L. J., Logan, J. A. and Kaplan, J. O.: Ensemble projections of wildfire activity and carbonaceous aerosol concentrations over the western United States in the mid-21st century, Atmos Environ (1994), 77, 767–780, doi:10.1016/j.atmosenv.2013.06.003, 2013.

Reviewer #2

Summary: This manuscript described a method for estimating burned area in the southern central region of the United States using three machine learning methods applied serially, with training derived from an existing dataset. The results show some skill in modeling total burned area over large areas. The work is focused mainly on the role of climatic variables in estimating burned area totals. While the methods in this paper might be of interest to the broader community, the manuscript is not well written (the structure is difficult to follow and it requires significant language editing throughout), the results cannot be reproduced because there is not enough information about the input variable processing, and the significance and limitations of the study are not explained well. Again, this method could prove to be useful to the broader community, but the manuscript needs significant work and for that reason I recommend rejecting this paper.

General Comments: While I think the methods presented in the manuscript have potential to produce useful results, the manuscript needs to be improved in order to create a more logical flow of information, better describe the data used, illustrate the output, provide a more complete literature review, and provide details about the usefulness and limitations of this study. Furthermore, it requires editing beyond the scope of scientific peer-review.
We thank the reviewer's comment. We have adjusted the sections of the manuscript to help readers better follow the article, and added more information about data, such as data source and the regridding method. With regards to the literature review, the focus of this study is on machine-learning-based prediction of wildfires and the relative importance of environmental controls of wildfires. Therefore, the literature review was mainly focused on this aspect. To follow the suggestions of the reviewer, we have added more literatures into the revised manuscript, including the ones mentioned by the reviewers. The details of the expected impact (line 70-73 and 558-567) and the limitation of this study (line 510-537) have been included in the manuscript.

(1)

The goals of this study are unclear – is the goal to predict wildfires in the future based on weather conditions, to support climate projections, or to simply estimate the amount of burned area?

The goal of this study is to develop a wildfire burned area prediction model that can be used to quantitatively estimate the contribution of different environmental factors that control wildfires at the grid level. We have stressed the goal of this study in line 69-73. To better represent this goal, we slightly revised the paper title to: "Quantifying the effects of environmental factors on wildfire burned area in South Central US using integrated machine learning techniques"

(2)

A related critique is that the structure of the paper makes it difficult for the reader to follow, there are effectively two methods sections with the results of the first set of methods in the middle.

As suggested, we have moved the validation method (original section 3.1) to Model section (new section 3.2).

(3)

Additionally, the authors never present a figure showing the modeled burned area, which should be the main output of this work and really needs to be emphasized in the main body of the manuscript.

We have moved the original Fig. S1 to the manuscript as Fig. 4 (Fig. R10 below). Fig. R10 shows the maps of monthly mean observed and predicted burned area averaged from 2002-2015 for both fire seasons. In addition to Fig. R10, Fig. 3 and Fig. 5 in the manuscript show the modeled burned area verses the observed burned area at the grid level and at the large-domain level. These results demonstrate that the model has the certain ability in predicting burned area at the grid-scale and at large-domain scale. We have rewritten and reorganized the corresponding paragraphs to emphasize the results in section 4.

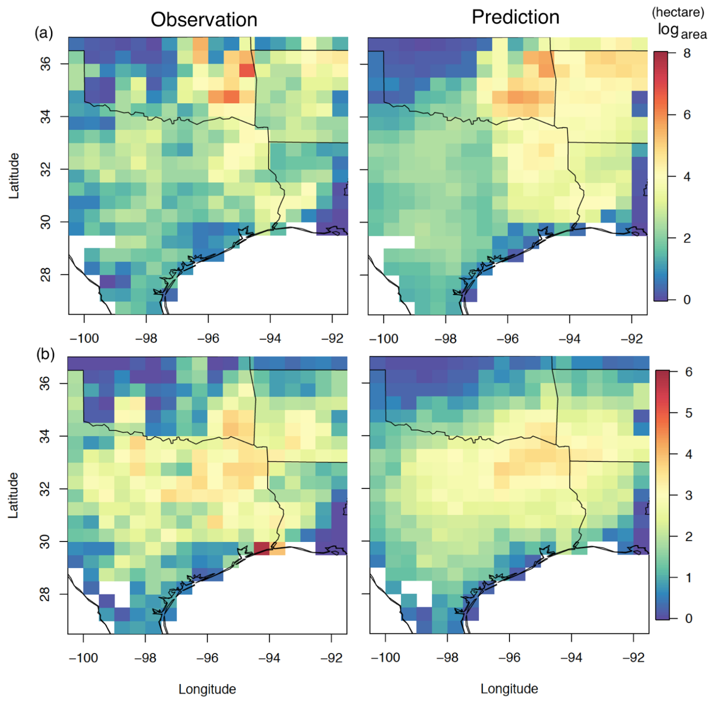

Fig. R10. Map of monthly mean observed and predicted burned area averaged from 2002 to 2015 for the (a) winter-spring and (b) summer fire season. (This figure is now Fig. 4. In the revised manuscript)

(4)

The authors have not considered a large body of wildfire research regarding satellite observations-driven modeling which is relevant to this work in the background research. Similar studies involving the effects of climate on total burned area should be noted by the authors, including Andela et al., 2017 and Zubkova et al., 2019.

The focus of this study is on machine-learning based prediction of wildfires and the relative importance of environmental controls of wildfires. Thus, the literature review was mainly focused on this aspect. However, we agreed reviewer's comments and added some references including the references mentioned by the reviewers.

(5)

Additionally, the methods section refers to aspects of the data which are not described until a later section.

As suggested, we have moved the data section (new section 2) before the model section (new section 3).

(6)

This organization is difficult for the reader to follow, and the description of the data used is insufficient, in part because the sources of the input data are not provided.

The sections have been rearranged as Data (section 2), Model (section 3), and Model validation and evaluation (section 4), Contributions of environmental factors to predicted wildfire burned area (section 5), and Discussion and Conclusion (section 6). To better clarify the data sources, we have included them into the manuscript in section 2.

(7)

The data preprocessing methods are unclear as well – how is a discrete thematic variable like land cover type represented at 0.5-degree resolution?

The land cover type data is represented as 30 m x 30 m pixels with an assigned value to represent a given class of land cover type. We used the nearest neighbor resampling method to regrid the land cover type data onto the lower resolution of 0.5° x 0.5°. The nearest neighbor resampling method is illustrated in Figure R11. This method does not change any value from the original dataset but assigns the value to the new grid according to the value of the grid closest to the center of the new grid. This method was chosen because it is the fastest of the interpolation methods. For instance, compared to the nearest neighbor resampling, the majority resampling is relatively slow and time-consuming, because it is sensitivity to the size of the filter window and thus more experiments are required to determine the filter window. Additionally, since the it does not change the values of the cells, it is widely used for discrete data and it can keep the extreme values that are highly related to large fires (Baboo and Devi, 2010).

We applied the nearest neighbor resampling method to both continuous and discrete thematic variables to the 0.5° resolution. We have added the data preprocessing methods into the Data section (line 109-113).

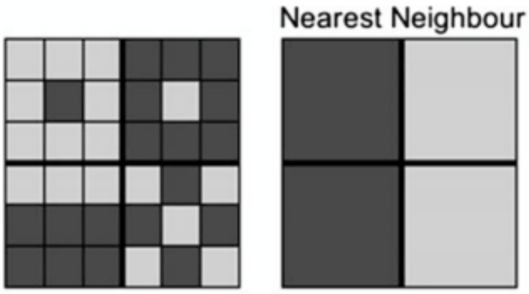

Fig. R11. Nearest Neighbor resampling technique. In this case, the nearest neighbor resampling is applied to grids with a resolution of 1x1 cell (left) to obtain grids with a resolution of 3x3 cell (right). (adopted from ESRI (2010)).

(8)

How are the translations between quantiles and area being made, given that the area of the grid cell varies with latitude?

The burned area of each grid cell was calculated by interpolating the fire data points into 0.5° grid cell based on their location. Although area of a grid decreases with latitude and higher latitude grids may contain fewer fire data points for interpolation, our interpolated burned area only depends on the magnitude but not the amount of the fire data points in a grid cell. Thus a higher latitude grid could have a large burned area despite its smaller grid area. For instance, we randomly sampled 10% of the grids from two groups of grids for ten times: grids with latitude ranging from 26.75° to 28.25° (representing larger grids in lower latitude) and grids with latitude ranging from 35.25° to 36.75° (representing smaller grids in higher latitude). As Figure R12 shows, the grids in lower latitude ranging from 26.75° to 28.25° overall have smaller burned area, with the mean log of burned area of 1.28 ± 2.35 ha for the sampled grids, while the grids in higher latitude ranging from 35.25° to 36.75° generally have larger burned area, with the mean log of burned area of 2.18 ± 2.67 ha.

The above analysis supports the argument that a higher latitude grid could have a large burned area despite its smaller grid area. Given that our model can successfully capture burned area in a grid cell across various latitudes and burned area may not be dependent on grid area, our interpolated burned area distributions therefore need not be normalized to cater grid cells with different grid areas. When the quantile of the burned area distribution is predicted by the model, we just need to use the predicted quantile to extract the final predicted burned area from the distribution.

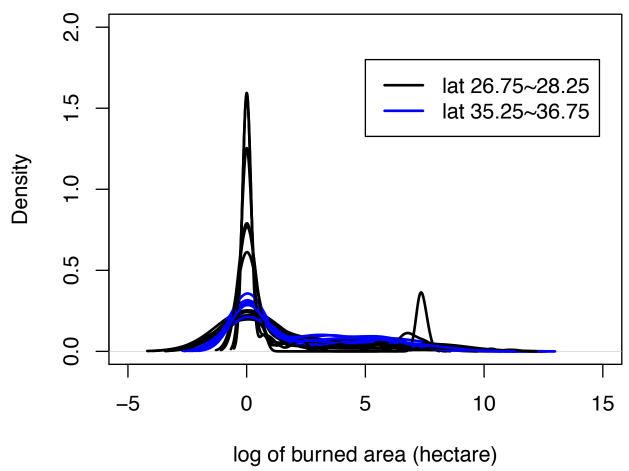

Fig. R12. Probability distribution of burned area for the randomly-selected grids in latitude range of 26.75° to 28.25° (black) and 35.25° to 36.75° (blue).

(9)
Also, the authors say the model predicts burned area at 50 km spatial resolution (with no indication of the map projection used), this is not the same as 0.5 degrees and this discrepancy needs to be resolved.
For this study, the point location of wildfire burned area was grouped into 0.5° grid cell based on their longitude and latitude. To avoid confusion, we have replaced the spatial resolution of '50km x 50 km' with '0.5° x 0.5°' throughout the manuscript.

(10) There are also questions about the fire data used to train the model – are prescribed fires included in the data (by definition, these are not wildfires in most cases)? Is there an estimate of the number of fires which are omitted?
The FPA-FOD fire data that we used excludes prescribed fires except for the prescribed fires that escaped their planned perimeters and became wildfires (Short, 2017). We have

clarified this in the manuscript (line 92-93).

Short (2014) compared FPA FOD data (1992-2011) with two national fire estimates from the US Department of Agriculture Forest Service (USFS) Wildfire Statistics and the National Interagency Coordination Center (NICC), which are available for 1992-1997 and 1998-2011, respectively. They showed that the annual number of fires estimated by FPA FOD is about 30% lower compared to that from the USFS estimation for the period of 1992-1997, as shown below (Fig. R13). The inconsistency of the fire number possibly may be caused by underestimation of small fires, as the fire burned area agrees well with USFS data. Our model will not be able to predict those small fires missing from the FDA-FOD as such information is not in the training dataset. The above discussions have been included in the text (line 93-98).

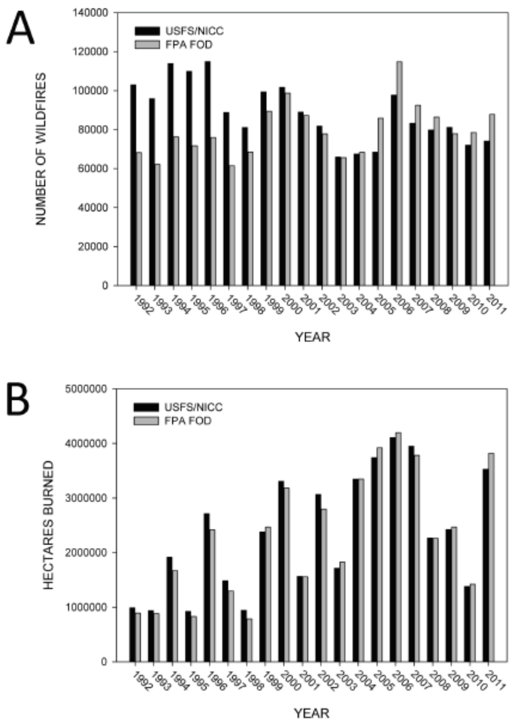

Fig. R13. Comparison of wildfire (a) numbers and (b) area burned area (hectares) in the US, 1992-2011, from published national estimates (USFS/NICC) and from FPA FOD. (adopted from Short et al. (2014); Fig. 4).

(11) Given the quantile-based approach, what happens if there is a fire or amount of burning which is greater than any in the training dataset (i.e. it would fall out of the range of the training data unless there was a training cell with 100% burned area)?
For a single grid, our four-step model can predict burned area greater than it had before based on its environmental conditions and by learning from other grids. However, random forest or quantile regression forest model cannot predict burned area larger than what it observed before from all the grids. For example, if the largest gridded burned area across the whole domain is 800 ha, the prediction for a single grid would never exceed 800 ha. Even though other methods such as MLR can predict burned area larger

than it observes before, there are some uncertainties in extrapolation (Amatulli et al., 2013; Mckenzie et al., 1996). We have included the above discussion in the manuscript (line 513-520).

(12)

Is the length of the training dataset long enough to capture all variability in fire activity as it relates to climatic conditions?

In 10-fold cross validation, the training dataset contains 16277 samples (It is derived from the total data length of 18085/10*9=~16277) for each fold for the winter-spring fire season. Assuming fire-climate relationships are unique for each individual grid, the large sample size is enough to capture all the variability in fire activity and its response to recent decadal climate. We have included this statement into the manuscript in line 429-432.

(13)

Why were remote sensing-derived datasets not considered?

We included satellite-derived monthly mean Leaf Area Index (LAI) obtained from MODIS instrument. In terms of fire data, since our focus is on wildfires and remote-sensing dataset does not separate prescribed fires from wildfires, we used FPA-FOD fire data instead of satellite-derived fire data such as GFED4 or FINN.

(14)

An important aspect of fire regimes which was not adequately considered in the manuscript is the role of human activity in the fire regime, especially in the United States where humans play an active role in the fire regime through suppression, ignitions, fuel load management, and landscape fragmentation in addition to being the source of ignition of approximately 85% of fires (according to the US Forest Service). These effects vary as function of not only population density, but sociopolitical norms which can vary from state to state. Recent papers such as the Andela et al. 2017 paper claim human activity is the major control on fire activity, and as such it cannot be ignored in a study region where the fire regime is likely human-driven. While the datasets describing human activity are certainly far from perfect, it is not possible to describe fire activity in a human-driven fire regime without considering human influences.

The focus of the paper is to quantify how environmental factors control wildfires in the study region under the present-day human management practices and human activities. Thus, we only included population density data of year 2010 to represent present-day human influence on wildfire activity. We acknowledge that population density is a rough estimate of effect of human activity on fires. We have included this statement into the manuscript in line 523-528.

(15)

Finally, there needs to be more effort in describing the expected impact of the work and the limitations of the method. For example, the abstract mentions that the work can be

used to assess fire management strategies but provides no details on how or why.
To clarify, the developed model aims to provide a broader impact on the community by accessing the quantitative contributions of the environmental controls of wildfires. An improved understanding of relative importance of the factors on wildfires would be useful for future fire prediction, fire management, as well as the linkage between wildfires and climate change. We removed the specific use of the model for fire management in the abstract and have restated the expected impact of the work and added the limitation of the method in the manuscript (line 510-537 and 558-567).

(16)
The quality of the input data is not discussed, which will propagate errors through the model, as well as the serial structure of the integrated model itself. At present, the manuscript is too focused on the machine learning exercise rather than on the scientific value of the work.
The discussion about the quality of the input data is added into the manuscript (line 520-523). Also, we have rewrote the manuscript to emphasize the scientific value of this work.

Specific Comments:
1. L75: Why were other months of the year excluded?
The total burned area of the two seasons accounts for 76% of the total annual burned area. As Fig. R14 shows, there are two peak seasons in South Central US: January to April and July to September. The dominance of wildfire occurrences in these months implies natural environmental conditions in these months are most conducive for wildfires. While wildfires do occur outside the fire seasons, their lower frequency implies that non-natural factors (e.g. human actions) can be relatively more important. As our study did not focus on human factors, we chose to exclude other months of the year. We have included the reason (line 74-79) and the Fig. R14 into the manuscript.

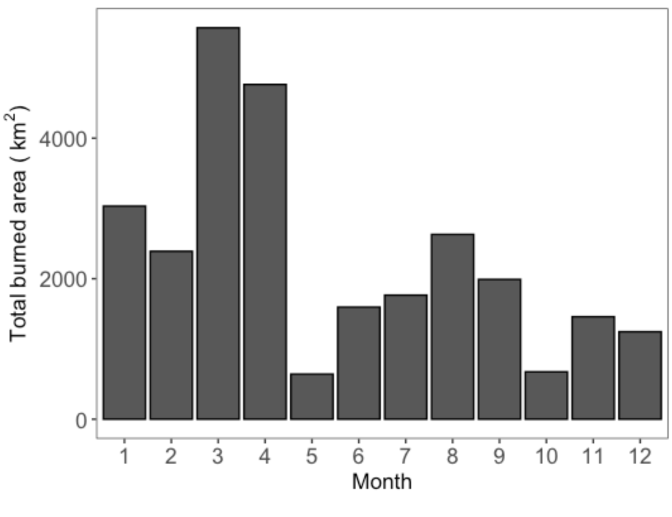

Fig. R14. Total burned area by month over South Central US. (This figure is now Fig. S1. In the revised manuscript)

2. L81: "Uneven data" is used throughout the paper but is not defined. Does this refer to unevenness spatially, temporally, or both?

It refers to both. The uneven distribution of burned area is defined as the situation where the number of grids with large burned areas is much smaller than the number of grids with small or zero burned areas. This situation exists for a single grid (temporal unevenness) and for all the grids within a given time period (spatially). The definition of unevenly-distributed burned area data has been included in the manuscript (line 57-59 and 182-183). An example of the uneven distribution of gridded burned area in winter-spring fire season is shown in Fig. R15. The Fig. R15 is included into the supplement to support the example of our case.

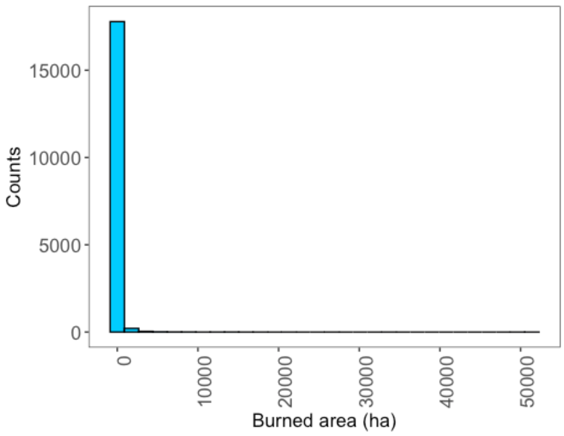

Fig. R15. Histogram of gridded burned area for the winter-spring fire season. (This figure is now Fig. S2a. In the revised manuscript)

3. L92-95: Given that the model compares the output to the quantile ranges, is it capable of estimating an amount of burning greater than has been observed in the training data?

See the response to general comment #11 above.

4. L155: Is there any concern about propagation of error through the model? What is the benefit of running three models in serial rather than one model alone or several models in an ensemble?

See the response to comment 6 for the reviewer 1, as shown below:

The reviewer is correct that biases from one step could be propagated to the subsequent steps, for example when the burned grids are predicted not to burn or when the unburned grids are predicted to burn. For the first case, when the burned grids are incorrectly predicted not to burn, the low bias is introduced because the burned grids would not proceed to step 3. For the second case, inclusion of unburned grids in step 3 may introduce a positive bias. We have included the discussions of the error propagation in section 6 (line 511-513).

To demonstrate our four-step model can achieve a higher accuracy and alleviate the issue of uneven-distributed dataset, we compare the prediction performance using random forest alone with that of the four-step model developed in this study, as shown in the Table R3 below:

Table R3. Comparison of MAE and skewness between the RF model and the developed model (The information of this table is now included into the Table S2 in the revised manuscript)

Model	**RF alone**	**Model developed in this study**
MAE (winter-spring)	1.34	1.13
Skewness (winter-spring)	37.40 (burned area)	0.70 (quantiles)
MAE (summer)	0.70	0.57
Skewness (summer)	33.83 (burned area)	0.96 (quantiles)

Skewness is a measure of the asymmetry of the probability distribution of a random variable about its mean. The skewness of a random variable X is the third standardized moment $\widetilde{\mu_3}$, defined as:

$$\widetilde{\mu_3} = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = \frac{\mu_3}{\sigma^3} = \frac{E[(X-\mu)^3]}{(E[(X-\mu)^2])^{3/2}} = \frac{\kappa_3}{\kappa_2^{3/2}}$$

where μ is the mean, σ is the standard deviation, E is the expectation operator, μ_3 is the third central moment, and κ_t are the *t*-th cumulants. If skewness is less than -1 or greater than +1, the distribution is highly skewed. If skewness is between -1 and -0.5 or between +0.5 and +1, the distribution is moderately skewed. If skewness is between -0.5 and 0.5, the distribution is approximately symmetric. The positive value indicates that the tail is on the right side of the distribution while negative value indicates that the tail is on the left.

Our model has a lower MAE, by 15% and 19% for the winter-spring and summer fire season, respectively, compared to the single RF model. The distribution of the quantiles in the developed model is more uniform than the distribution of the burned area, as shown in Figure R6 below and the skewness. We have added a discussion of this issue in the text (line 355-365) and the Table R3 in the supplementary information. The information of skewness calculation has been included in the supplementary information. Note that we have replaced 'quantile' with 'percentile' in the manuscript to better clarify the idea of quantiles.

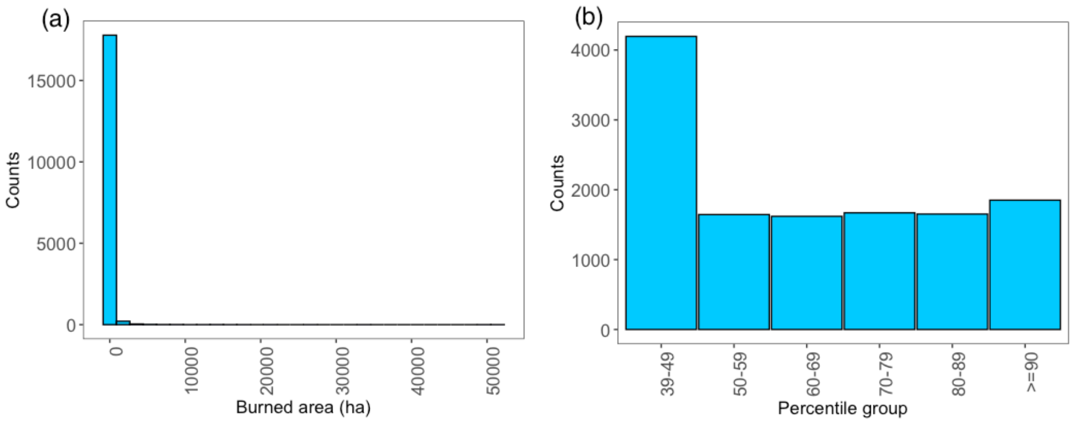

Fig. R6. (a) Histogram of burned area (b) Histogram of the percentile groups of burned area for the winter-spring fire season. (This figure is now Fig. S2. In the revised manuscript)

5. L164: Is there an estimate of the number of fires missed? Small fires constitute most of the fires by number, even though they add up to relatively little burned area (e.g. Malamud, Millington, and Perry 2005). It is noted that the dataset omits most small fires occurring on private land – these are not generally wildfires and such fires should be omitted anyways if the study is about wildfires.

See the response to general comment #(10) above.

6. L194-195: I don't think the climatic variables can be considered as fixed, especially since the assumption in later parts of the paper surround climate change scenarios which means their values do vary through time.

Climatic variables that are considered as fixed include only the mean and standard deviations of monthly meteorology over the past 22-years (1979-2000), because they do not vary by time over our study period (2002-2015) but characterize the spatial patterns of wildfire occurrence and intensity. The variables of climate anomaly are classified as climate variables (as opposed to fixed variables) since they are defined as the difference between monthly mean and the long-term average over 1979-2000 and their values vary by time. We have clarified this in the text (line 149-150 and line 435-440).

7. L209: Is any consideration given to preventing overfitting due the correlation between variables? For example, ecoregions and landcover types are likely to be related to one another.

Yes, we considered the collinearity of the variables when we designed the model. Thus, we chose logistic model and random forest model which work reasonably well under moderate collinearity (correlation coefficient < 0.7) (Dormann et al., 2013). We have added the concern of collinearity between variables into the text in line 213-215.

Although ecoregions and landcover types are likely to relate to one another, ecoregions represent large-scale areas comprised of similar biotic and abiotic phenomena while land cover types are able to provide more detailed land information within one ecoregion. For example, in the temperate prairies (one ecoregion of our study domain), it has complex land types including pasture, woody wetlands, evergreen forest, and cultivated crops. Inclusion of these two variables allows us to capture fire responses to large-scale ecoregions and small-scale land types.

8. L232: Please clarify the phrase "horizontal scale of around 700 x 700 km2" – the use of horizontal scale implies a one-dimensional unit (length) which does not match the unit specified. Also, as a suggestion, 700 x 700 km2 seems ambiguous and could be more clearly represented as "700 km x 700 km" or "490,000 km2"
Good point. To avoid confusion, we have changed 'horizontal scale of around 700 x 700 km2' to 'spatial scale of around 700 km x 700 km'. The similar changes were also made for the Table S1.

9. L252: SUS is never defined
We have defined 'SUS' as 'southern US' for clarification.

10. L268-270: One could argue that the model is in fact "hardwired" (editorially, the term is jargon and should be replaced) to the geographical features of the study domain – geography deals with the human components of space and time as well as the physical components. The tendency of the human population to ignite or suppress fires as a result relationship to sociopolitical factors (like local regulations) will influence the fire regime in ways which will not be captured by climatic variables and will change from location to location.
We agreed with reviewer's point but the main focus of this paper is on how the environmental factors control wildfires in South Central US under the present-day human management practices and human activities. Therefore, the geographical features in the manuscript refer to coordinate variables such as longitude and latitude. To clarify this, we have replaced the term 'hardwired' and rewrote the corresponding paragraph in the manuscript (line 348-351).

11. L286: Why were 14 variables chosen? This seems like an arbitrary cutoff, especially given the large number of variables which went into the model.
We chose the top 14 variables because they represent the top quarter (25%) of the selected predictor variables. %IncMSE represents the change of mean square error with and without permuting variables. A larger %IncMSE value represents a higher variable importance. To see the sensitivity of the importance to the variable rank, we calculated the ratio of %IncMSE at variable ranked as X^{th} percentile (~top Y) to the %IncMSE at variable ranked as top (Y+1). Larger ratio means larger drop-off of the %IncMSE between topY and top(Y+1), which indicates notable decrease of variable importance at the cut-off point (top Y+1). We compared the ratio at several cutoff points: 25^{th} percentile, 50^{th} percentile, and 75^{th} percentile:

Table R6. The ratio of %IncMSE at variable ranked as X^{th} percentile (Y^{th}) to the %IncMSE at variable ranked as $(Y+1)^{th}$ for the three selected percentiles

	25% (Y=14)	50% (Y=29)	75% (Y=43)
Winter-spring fire season	1.21	0.88	1.00
Summer fire season	1.06	1.01	1.00

As the table shows, 25-percentile cut-off point has largest ratio, indicating a large drop of variable importance at the variable ranked 15^{th} and the top 14 variables have significantly larger importance. Thus, the top 25% variables (the top 14 variables) were chosen to be further discussed. The reasons of choosing the top 14 variables and associated discussions have been included in the manuscript (line 380-382). Table R6 has been added into the supplementary information as Table S4.

12. L306: The fuel-related variables are among the least important presented in Figure 5 – how can the conclusion be drawn that fuel abundance is what determines the amount of burned area?
Although fuel-related variables are among the least important in the top 14 variables, they are the fifth and sixth most important variables when excluding the fixed variables. Our conclusion was mainly based on the importance of time-varying variables. Therefore, burned area in the winter-spring fire season is mainly controlled by RH anomaly that directly affects fuel moisture. Besides that, the antecedent fuel abundance and pre-fire-season drought conditions together determines the amount of dry fuel in the winter-spring fire season. To clarify this, we have rewritten the corresponding paragraph in line 407-410.

13. Table 1: The resolution of the data is presented, but it's not clear how the data are being re-gridded to the working resolution - if the fire analysis is being done at 0.5 degree and the climate data is at 32 km resolution then there are < 4 cells per burned area data point and the way which those 4 cells are represented has significant consequences.
See the response to general comment #(7) above.

References:

Amatulli, G., Camia, A. and San-Miguel-Ayanz, J.: Estimating future burned areas under changing climate in the EU-Mediterranean countries, Science of The Total Environment, 450–451, 209–222, doi:10.1016/j.scitotenv.2013.02.014, 2013.

Baboo, S. and Devi, R.: An Analysis of Different Resampling Methods in Coimbatore, District, Global Journal of Computer Science and Technology, 10(15), 61–66, 2010.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne,

P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D. and Lautenbach, S.: Collinearity: a review of methods to deal with it and a simulation study evaluating their performance, Ecography, 36(1), 27–46, doi:10.1111/j.1600-0587.2012.07348.x, 2013.

ESRI: Manual ArcGIS [software GIS] Version 10, Environmental System Research Institute, Inc., 2010.

Mckenzie, D., Peterson, D. L. and Alvarado, E.: Extrapolation Problems in Modeling Fire Effects at Large Spatial Scales: a Review, Int. J. Wildland Fire, 6(4), 165–176, doi:10.1071/wf9960165, 1996.

Short, K. C.: A spatial database of wildfires in the United States, 1992-2011, Earth System Science Data, 6, 1–27, doi:10.5194/essd-6-1-2014, 2014.

Short, K. C.: Spatial wildfire occurrence data for the United States, 1992-2015, Forest Service Research Data Archive (4th Edition), doi:10.2737/RDS-2013-0009.4, 2017.