-Reviewer 1

This paper uses a recently developed metric, called the ‘stand-replacing decay coefficient’ (SDC), to summarize high severity fire for each of over 450 fires in NW California over the 1984-2015 time period. The SDC provides more information compared to metrics such as ‘high severity patch size’ in that it considers the interior distance from areas that are burned with a lower severity; areas that burned with lower severity likely have retained live trees that can supply seeds to areas that experienced stand-replacing fire. The authors look for differences in SDC among different managing agencies, over time, between fire management strategies (suppression vs. managed/WFU fires), and over varying weather conditions. This is a well-written paper and provides a valuable scientific contribution.

We thank the reviewer for their positive feedback and for their useful suggestions below.

Below are some concerns and suggestions:

One of my more major concerns pertains the how weather was characterized. Using Tmax as an example, different biophysical settings inherently have different Tmax. Lower elevations will, on average, have higher Tmax than higher elevations. Since CDF-managed fires are usually at a lower elevation than other fires, it is only natural that Tmax will be higher for CDF fires. Could this influence your interpretations? Also, native Tmax units (degrees C) are difficult to properly interpret when study areas cover broad geographic extents. For example, the mean Tmax for CDF-managed fires in 32.7 degrees C. Is this warmer than normal (for the specific geographic location) or cooler than normal (for the specific geographic location)? This value of 32.7 degrees C is much higher than the USFS-managed WFU fires (24.9 degrees C), but because these fires occurred in different biophysical settings, it is impossible to ascertain whether these Tmax values are hotter/cooler than average. Are WFU fires burning under ‘normal’ conditions? Are CDF-managed fires burning under ‘normal’ conditions? This is important because the species composition of any given site are usually acclimated to the weather and climate conditions of the site, and therefore, the fire regimes and the SDC would be expected to differ among biophysical settings. Perhaps the ‘counterintuitive result’ reported on line 352 is because the Tmax units are not normalized to the average temperature of any given site. Simply put, I think what is more interesting is the deviation from average. Consequently, a potentially more informative approach in characterizing Tmax is to characterize in terms of percentile or z-score Tmax. All this said, calculating percentiles or z-scores is time consuming. If the authors choose not to follow my suggestion, I think this particular caveat should be mentioned in the Discussion. For more info on the percentile idea, I believe Dillon et al. (2011; Ecosphere) and Birch et al. (2015; Ecosphere) used weather percentile data.

These are good and important points. We carefully weighed whether to relativize the weather data, and decided against it for the following reason: We believe that modeling SDC on absolute weather conditions indicates just how important fire weather is for influencing stand-replacement patterns, irrespective of biophysical settings. Fire behavior responds to absolute weather conditions, not relative conditions. To the extent that geographic differences in our different agency/class combinations are associated with differences in “normal” weather, we think it is important to highlight these differences as contributing ultimately to the size and configuration of stand-replacing fire effects. We have strived to be transparent about these geographic differences throughout the manuscript; that is one of the reasons why we included Table 2. For instance, the fact that the National Parks are generally located at higher elevations may explain the differences in TMax under suppression conditions (Table 2), but interestingly the average TMax under WFU conditions are almost identical between NPS and USFS, suggesting that each has opportunities to manage fire under less extreme conditions.

Regarding the acclimation of species composition to the regional climate and fire regime, while it is true that tree species in lower-elevation forests are generally more fire-resistant and might be expected to show less stand-replacement, these adaptations are largely overridden by changes in forest structure which, in concert with weather, are leading to increasingly large patches of stand-replacing fire, as our data show. It would be difficult to attribute regional differences in stand-replacing fire to deviations from normal climate as opposed to other land use changes that have rendered formerly heterogeneous, low-density stands more susceptible to stand-replacing fire. And regarding the “counterintuitive result”, we feel quite confident in the assertion that the fires that fall into this category are the result of the unique topography of the Klamath Mountains, particularly during the 1987 fire season, when inversions reduced fire activity despite high absolute temperatures, as previous work by Miller et al. (2012) has indicated.

We appreciate the reviewer’s larger point that the relationship between weather and our management variables is an important one, and we have made this connection more explicitly on **lines 338-348**.

I feel like some of the inferences regarding ‘management history’ and ‘fuels’ might be overstated. In my opinion, since you did not directly evaluate management history (e.g. last time thinned) or fuels, some of these inferences should be toned down. Also, because CDF-managed fires may be on either USFS or NPS land, ‘management history’ becomes more convoluted.

This is a good point, management agency is a coarse proxy for management history and does not represent fuels. We have revised the text and are now clearer about the limitations of this proxy, and have incorporated the reviewer’s specific comments on this topic below (e.g. **Lines 349-366**).

Unrelated to this paper, a potential avenue for future research may be to compare contemporary SDC values to historical SDC values to look for departures. One might expect that contemporary SDC values are smaller compared to some historic period. Obviously, getting ‘historic’ SDC is not easy, but maybe a simulation model or using historical aerial photos could be of use.

This is a good idea and one we have been thinking about, particularly with respect to historical aerial photos. We are in initial discussions with some collaborators about attempting this. We note that the addition to this paper of a new website app (**line 401**), which allows users to upload shapefiles of their own stand-replacing patches of interest, should open the possibility of this type of application being more widely attempted going forward.

Here are some specific comments

1. Line 23: ‘spatial scale’ might be a little esoteric for an abstract. How about ‘patch size’ or something to that effect?

Done. Now L23-24

2. Line 30: ‘past forest management’. This is a good example of overstating you inferences about fuels (second point above).

Agreed, as we describe below in response to the reviewer’s comment #8, we now distinguish between management class (WFU vs suppression) and management agency, and collectively refer to them as “fire management”. Without wanting to make this distinction in the abstract, we now simply refer to “fire management” here and are clear about management class, management agency, and the inferences that can be drawn from each variable, throughout the rest of the manuscript. Now L30

3. Line 53: is ‘top-killed’ the correct term? Why not simple killed.

Top-killed is generally used to describe a situation where the aboveground portion of the plant is killed, but resprouting may occur. In the mixed-conifer forests where we focus, this is generally not an issue as the dominant conifers do not re-sprout, but since the remote sensing techniques to assess burn severity (RdNBR) only measure aboveground “mortality” rather than permanent mortality, we believe this phrase is more technically accurate to the process at hand and wish to retain it. Now L54

4. Line 87: consider changing ‘scale’ to ‘resolution’.

Done. Now L89

5. Line 110: consider adding ‘, and therefore type conversion,’ (or something like that) between ‘conifers’ and ‘compared’.

Good suggestion, done. Now L112.

6. Line 154: I believe this is the first mention that this study was conducted in NW California. It should be mentioned in the Abstract and Introduction. On a similar note, a study area figure would be nice.

We do introduce California as the general study area in the Abstract but we now identify the NW CA/Sierra Nevada focus in the introduction on **Line 135**. At the reviewer’s suggestion we have included a study area figure as Figure 1.

7. Line 157: need a comma after ‘for our analysis’.

Done. (now L161)

8. Line 196: Again, ‘land management history’ is inferred. Better labeling as ‘managing agency’ or something like that. ‘Fire management’ refers to suppression vs. WFU, right? If so, I recommend labeling as ‘fire management’ throughout and not as ‘management’ as to avoid confusion.

Good points, we have distinguished “fire management agency” from “fire management class” throughout, and we use “management” to refer to the two together. Now L205

9. Line 199: I’m only guessing you used GLMs here (due to the R package). I think you need to explain this much better.

10. Line 208: It is not clear why you used CART in addition to GLMs. Please clearly explain your rationale.

Re: comments 9 and 10. we used linear models (because the transformed data were normally distributed) to compare alternative models, and having selected the best model, we used CART to visualize the model and identify important thresholds in the predictor variables. We now clarify this in the paragraph on **lines 205-225**.

11. Line 215: Is it necessary to do the five year averages? I think the CIs in the figures provide a clear enough illustration. Plus, due to temporal autocorrelation (line 216), the R2 values for the five-year averages are bogus since each observation is not independent.

This is a fair point, we removed the five-year averages.

12. Line 221: I’d call it ‘fire management class’ to avoid ambiguity.

Done. Now L233

13. Line 237 (and elsewhere): perhaps back-transforming the values back to native SDC units is pertinent. SDC itself is hard enough to understand, but reporting the log-transformed units complicates even further.

Good point; we are already presenting both transformed and untransformed SDC values in the discussion, and now we do that here as well (**Lines 255, 256**). We choose to report both rather than only the untransformed values because both our GLM and CART analyses were done on the transformed data to meet assumptions of normality. Reviewer 2 also had this suggestion.

14. Line 250: ‘the reduction’ – should this actually be ‘larger SDC values’?

Yes good catch; we have changed this and carefully checked the manuscript for similar mistakes (also changing “smaller” to “larger” on **Line 372 and 406**).

15. Line 271: ‘averaged across all fires within a given year’ – this appears to be the case for all plots. I recommend moving this up and perhaps including the figure legend.

Correct; we now specify that the SDC values in this section are averaged annual values, and that the weather parameters analyzed are average annual values of the maximum weather values during a particular burn window (**lines 278-286**), and we include this information in the Figure 3 caption.

16. Line 276: ‘Consistent with previous …’ – this statement is Discussion material.

Good point, we now raise this discussion in the context of the uniqueness of the Klamath Mountains, **lines 381-383**.

17. Line 289: In addition to reporting the total area, what about the proportion of area burned managed by each agency? USFS has by far the most area of ‘potential forest loss’ but USFS presumably has the most area burned. Is USFS doing better/worse proportionally than CDF or NPS. Would be good to add to figure 4 as well.

We have added this statistic to the paragraph in question (now **lines 296-297**), and also to the caption for Figure 4.

18. Comma after ‘previous work’.

Done (Assuming the reviewer was referring to the first paragraph of the discussion).

19. Line 314: Westerling is probably not the best cite here. They did not explicitly look at ‘anticipated changes in climate and fire frequency’. Maybe something by Jeremy Littell or Sean Parks’ in press article in Ecography.

Good suggestions, we added Jeremy’s and Sean’s papers to this statement. We also believe Westerling’s paper is appropriate insofar as area burned is concerned, because the regional focus (climate projection models downscaled to California, and chosen for their accuracy for that particular region) is actually the most appropriate to our study of the three papers, so we added area burned to the statement and retained the Westerling citation.

20. Line 317: ‘our results corroborate this’ – again, you did not explicitly evaluate fuels, so this is a bit overstated.

Fair point, we have toned down our language. Now L324.

21. Line 331: Did Abatzoglou and Williams even evaluate fuels? Similarly, I feel like most studies found a minimal to negligible influence of weather on fire severity. In my opinion, this is partially because it is terribly difficult to characterize weather.

Thanks, the reviewer is right that Abatzoglou and Williams looked at fuel moisture rather than fuel loads (which we implied). Based on this comment we decided to remove this sentence entirely. The point is not entirely pertinent to the results we present. Now ~L351.

22. Line 354: Do you mean ‘larger’ SDC values?

Yes, see above in response to reviewer’s comment #14.

23. Line 380: ‘desirable’ is a subjective term.

Good point, we have clarified the specific patch size associated with the thresholds we identified rather than calling them “desirable”.

24. I think a figure up front illustrating SDC would be useful since it is kind of difficult to understand. In theory, you could move figure 5 up.

The Collins et al. 2017 citation gives a detailed explanation of SDC and examples of how it is calculated; however we agree that more information in this paper would also be useful, so we have modified a figure from Collins et al. and included it in the Appendix as Figure A1.

25. Figure 2: I recommend making it more apparent in the legend that the x-axis in the left column is high severity fire and the right column is area burned. Took me longer than expected to figure it out.

We have added column headers to each side to further clarify the difference between them.

26. Figure 3: again, I’m not sure the five year mean column is necessary and I think the R2 values are inflated.

Agree, we have removed them.

27. Figure 5: inclusion of the fire perimeters would be nice.

We now display the fire perimeters underneath the stand-replacing area in Figure 5.

28. I think that Meyer (2015; Journal of Forestry), Chambers et al. (2017; FEM), and Cansler and McKenzie (2014; Ecological Applications) are relevant studies that should be cited.

These are great recommendations – we are familiar with all of them and have incorporated them where relevant.

-Reviewer 2

The manuscript, “Changing spatial patterns of stand-replacing fire in California mixed-conifer forests” applies a novel fire severity method to describe changes in high severity fire patterns wildfires over time and across different public land ownerships within California, USA. This manuscript represents a very thoughtful applied contribution of the SDC metric to understanding landscape-level fire patterns across fire prone ecosystems and is a natural extension of the original Landscape Ecology article. I suggest this article should be accepted with minor revisions. My comments are made to help in readability and accessibility to those unfamiliar with the methods presented here.

Thank you for these positive comments; we appreciate the suggestions below.

Minor edits

Page 4-6 lines 80 – 116: These two paragraphs provide a lengthy defense that 1) high-severity classified RdNBR values are representative of stand-replacing fires, and 2) the SDC concentrates on high-severity patches given that regeneration failures are most likely to occur in these patches, and 3) mixed severity patches include fine-scaled variability in tree mortality, but the ecological implications of this variability are different given their closer proximity to a replenishing seed source. This section could be revised and shortened, in my opinion, given that previous articles (cited on line 93) have shown the RdNBR cutoffs developed for this region are reliable and representative of the mortality levels cited on line 89, and the previous SDC Landscape Ecology article makes the argument for the grain of the analysis. It seems the authors are trying to preemptively take a rock out of the road where none exists.

We agree with the reviewer’s assessment of the intention of these two paragraphs,. We use the widely used terminology of low, moderate and high (or stand replacing) to describe fire severity at the level of the mapped pixel. We use the term “mixed-severity” to describe fires that contain stand-replacing (high severity) patches of varying size and shape (**lines 109-113**). We appreciate the reviewer’s suggestion to condense this material, and we acknowledge their point that some of these points are made in Collins et al. 2017. However we think that this detailed explanation of the importance of spatial scale in the definition of mixed-severity fires is critical to reiterate here, because this term is a “rock in the road” when it is used inappropriately: If most fires are mixed-severity fires based on the percentage thresholds, then “mixed-severity fire” becomes meaningless as a term to distinguish widely divergent post-fire environments that depend on stand-replacing patch sizes. Furthermore, we believe it is absolutely critical to remind readers that although basal area mortality greater than thresholds such as 70% or 90% may be used to classify RdNBR values into high severity classes, the vast majority of the area within these patches is true stand-replacing fire with 100% mortality, which the citations on **Line 99** make clear. If the editor believes length of the manuscript is a concern, we would be willing to condense some of this material.

Page 8, line 164: It should be made clear 1) what RdNBR cutoff value was used to define high-severity, 2) that the “>90% basal area mortality” is estimated from the RdNBR itself, 3) that the polygons were essentially patches of high-severity within a fire that were delineated by aggregating adjacent pixels of like severity class in a GIS.

This is a helpful set of suggestions that we have implemented on **lines 195-200**.

Page 11, line 231, Management class isn’t really a first-order control, it is the extinguishment of the flaming front that is the first-order control; management class should be referred to as A primary or THE primary predictor of SDC in your model.

Good point; the rpart analysis assigns the classification tree hierarchy based on importance values, so management class is statistically the most important predictor of SDC; we have made this change on **line 280**.

Page 11, line 237. The SDC scale itself is hard enough to decipher, and the ln(SDC) is pretty esoteric. I suggest using the original scale (put ln value in parentheses if it is required) and maybe it would be helpful to scale the scores that are being referenced to the min/max of all observed fires, to get a sense of where these scores line in relation to the sample of fires assessed here. So, small and complex shaped patches could score near the lower quartile, while larger and more simple shaped patches could score in the upper quartile.

Good point; we are already presenting both transformed and untransformed SDC values in the discussion, and now we do that here as well (**Lines 255, 259**). We choose to report both rather than only the untransformed values because both our GLM and CART analyses were done on the transformed data to meet assumptions of normality. Reviewer 1 also had this suggestion. With respect to a quartile analysis, we believe that Figure A1, referenced here, along with the distribution of ln(SDC) values present in in Figure 2 collectively give the reader a good idea of the range of possible SDC values. We also believe that reporting the actual value rather than a standardized version lets the user calculate the proportion of their stand-replacing area that is greater than a given distance from the patch edge using Eq. 1.

Page 11, line 231-246: this paragraph is fairly hard to follow. I would break it into comparisons between main effects. I.e., main differences between SDC among land ownerships, main influences of temperature, main influence of suppression, etc. After discussing the main effects, any nuisances can be discussed, but I would avoid merely running the readers down the tree.

Good point, we simplified our summary in this paragraph to focus on the main effects and trends (**now lines 281-297**).

Page 11. Did you consider putting wildfire size as a predictor in the CART model? If it were a main predictor then it may show different relationships further down the tree, and the cutoff value (s) in the tree may have important ecological or management implications.

Fire size is loosely correlated with SDC (Figure 2b, d), but we believe that many of the variables that we used in the CART model also predict fire size as well as SDC (as they relate to climate and fuels management), so fire size makes more sense as a response variable than a predictor variable. The fact that SDC is correlated with fire size is related to the fact that larger fires can have larger patches within them. A strength of SDC is that it is related to fire size but it more directly relates to the biological processes that drive the post-fire ecology.

Page 17, line 372. Not sure that low SDC is an indicator of mega-fire, but rather a consequence of mega-fires.

We agree that indicator might not be the best word; we have chosen to go with “characteristic of mega-fires”, as we think that checking the SDC may be a useful way to identify a “mega-fire”.

It was interesting that burning index was not a significant predictor of SDC given that it has been associated with high-severity fires in past studies (i.e., Lydersen et al. 2017 cited in the current document). This may be caused by 1) the methods used ascribe a single BI value across an entire fire area as was done here, or 2) there is a mismatch in scale from daily BI and the size of high-severity patches, or 3) the reliance on a remote measure of BI from a weather station not proximal to the high-severity patches themselves, or 4) a greater influence of fuels over fire weather. I suggest discussing the lack of an influence of BI on SDC in the discussion as it may prompt future work. For instance using SDC metrics for individual burn days using fire progression and severity maps (and potentially fuels or fuels surrogate maps) to help determine drivers of large-scale high severity fire patches.

Yes this is a good question, we now discuss this on **lines 410-421**.

Figure 1. The rpart figure is a bit tricky to read for those uninitiated in regression tree plots. For instance, it is not always intutitive to know which direction to walk down the tree give a continuous value at the break. 1) Please provide some instance in the caption, 2) it may be useful to mimic what the authors did for the first break (class = suppression) and have the labels on either side of the break. For instance, for max high temp, the label of the break could be Max high temp and to the left could be ‘> 24’ and to the right ‘< 24’. For agency it could be label ‘USFS’, and to the left ‘Yes’, and to the right ‘No’, etc. Furthermore, the values of the ln(SDC) are hard to interpret in a stand-alone graph, particularly without any explanation in the caption. At the very least untransform the variable, but it may also be helpful to color the boxes based on the quantiles of the SDC values across fires to get a sense of how these values compare to the empirical distribution of the observed fires. Ie.., is exp(-5.1) in the upper 90% of all fires? So, the coloring could go from blue to red with red indicating values in the upper quartile of observed fire values and blue the lower quartile. It would also be nice to know the sample size in each bin.

These are excellent points. We updated this figure to now make the breaks in the tree clearer, avoid the “Yes/No” confusion that the reviewer pointed out, and provided context for the ln(SDC) values by providing a histogram of all values for the 477 fires analyzed (New Figure 2). We adopt a color scheme along the lines of what the reviewer suggested. We chose not to untransform the variable because as we described earlier, the log transformation is necessary for the data to be normally distributed (as the histogram demonstrates), and all analyses were done on the log-transformed data. However we believe the new figure is much improved and we thank the reviewer for their comments.