

Point-by-point response to reviews

Referee #1

Minor comments

1. Many sentences in the manuscript are very long, which makes them hard to follow. Please, review the entire manuscript text and try to break long sentences into smaller pieces.

Reply: We agree with the reviewer, and we revised the manuscript to avoid long sentences throughout it

2. L132-134: According to the authors, WRF output was provided at 3h intervals and the weather variables used were averaged over the 12-20h time window. How can this affect the credibility of their analysis, particularly with respect to extreme fire weather? Are the authors certain that the 3h WRF output and subsequently the averaging allows for pinpointing extreme fire weather? A comment on that, also within the manuscript, would be highly appreciated.

Reply:

We thank the reviewer whose questions are relevant, and we need to clarify them within the simulation framework. There are two main reasons for averaging weather values over the 12-20h time window:

- 1) We do not have information for individual fire events regarding the active fire spread periods and on the corresponding driving weather conditions. Therefore, we calibrated the fire modelling system using the historical fire regime instead of individual fire-based calibration. Considering this data limitation, we run simulations with constant weather for spreading durations lower than 24h. Therefore, based on expert knowledge, we know that in the study area the main active fire spread frequently occurs during the 12-20h time window, when higher temperature and wind speed, and lower humidity conditions prevail.
- 2) Historical large fires have spreading durations larger than 3h, and as shown in the calibration matrix and considering that suppression is not being modelled, the most frequent duration in the calibrated model is 5h (60%), which in a real case scenario fire duration is longer.

Impact:

- The averaging effects that may decrease extreme values of the weather variables used will be compensated by the tuning of the duration of simulation, which is the output parameter that is extracted from the calibration and allows the reproduction of the historical fire regime. The duration parameter is not a real duration of a single fire but, different durations are used to cover the full range of fire sizes observed in historical data (Appendix E)
- We believe that we are not failing to estimate the fire spread under extreme weather conditions, as shown in the results of the accuracy assessment done for the simulation of the 9 largest fires (above 1000 ha, usually driven under extreme weather conditions), in Appendix E.

As suggested, we included the next sentence in the final version of the manuscript (Lines 130-134):

“This time window choice is because fire simulations are run for spread durations smaller than 24h with constant weather conditions. Averaged values of the weather variables needed to exclude milder weather conditions that typically occur during the evening and morning periods, when fire spread is reduced also due to effective fire suppression efforts. The eventual averaging effect of the extreme weather conditions is compensated by tuning the duration in model calibration (Sect. 2.4.4).”

Technical remarks

L22: improve --> improving

L28: increasing in suppression efforts --> increased suppression efforts

L34: rather than

L91: remove where (?)

L130: Weather Research and Forecasting model

L465: Table C4 --> Table C1

L466: run --> ran

L511: This value or These values

L409: on the landscape

L418: future instead of next future

All these remarks were corrected in the final version of the manuscript.

Referee #2

This is a well developed research methodology and should be considered for publication in NHESS. To further improve:

My only suggestion would be to include a small analysis of validation (accuracy assessment) of your predicted fire size or intensity/ hazard against forest loss due to fire dataset for Serra da Cabreira in 2018 (https://glad.umd.edu/dataset/Fire_GFL/).

Reply:

We would like to thank the reviewer's suggestion of including a small analysis of accuracy assessment to further improve the study. We did not understand why 2018 was the year selected for comparison since the dataset was updated to include 2020 and 2021.

It is relevant to compare fireline intensity (as measured by the CFL estimate) with the forest fire loss (FFL), particularly for the year of 2021, as the first fire season after the wildfire hazard assessment. We need to highlight some considerations regarding the fire regime in the study area and the simulation framework:

- Fuel model map used in the simulations includes burned areas from 2020, thus we compared it with forest fire loss map of 2021 (FFL21) only including in the analysis those pixels coded with medium and high certainty.
- Wildfire hazard was assessed using the 95th percentile of fire spread days to represent the extreme weather conditions that usually can produce large fires, specifically ≥ 100 ha
- In 2021 the largest fires have 118 and 124 ha, and are both winter fires (from January and March, respectively). According to the FFL21 none of these fires were highlighted with forest loss.
- According to the official Portuguese Fire Statistics database, there are some discrepancies when compared to the FFL dataset. Annual burned area values cannot be lower than the area identified as Forest Loss due to fires in 2021. As we can see in the histogram below (Figure 1), there are three years where this is true: 2008, 2014 and 2018 (in the last case the difference is ca. 1600 ha). This raises suspicions that at least in these years there might be a problem with the FFL dataset.

Besides these considerations, and assuming that fire simulations were run at 100m of spatial resolution (lower than the 30m from the FFL dataset) under extreme weather conditions, and ignoring the lack of correspondence between the reference burned area extent with the corresponding annual records of the FFL dataset, even though we compared the CFL estimates with the annual area of FFL from 2001 to 2021.

The hypothesis is that areas that were recorded with FFL more recently, should have lower estimated CFL values, thus being able to generate lower intensity fires. Areas that were in the past classified with FFL are expected to have higher intensities now, especially if they are not from recent fires.

Therefore, we summed the total area of FFL divided in three periods (before 2010; between 2010 and 2016; and after 2016) and quantified the area of the estimated CFL divided in two classes (below and above 2.5m, based on its relationship with FLI). We removed from

analysis the years 2008, 2014 and 2018 due to the reasons mentioned above. Results are shown in table A, and indicate that areas that burned more than 5 years ago were in 2021 able to generate high intensity fires, in more that 70% of the areas that were recorded in the past with FFL (only those with moderate and high confidence). Oppositely, in approximately 60% of the areas that burned after 2016 with assigned FFL, these have a lower potential to generate high intensity fires.

Concluding, with this small analysis of validation, we have some evidence that in the last five years, areas that were recorded with FFL, have in 2021 a lower chance of spreading high intensity fires, as measured by the CFL.

Table 2.

FFL	CFL (m)	
	<2.5	>=2.5
before 2010	21.0%	79.0%
2010 - 2016	29,7%	70,3%
after 2016	62,4%	37,6%

We included the next paragraphs in the final version of the manuscript, according with the previous results:

- As a simple validation exercise we evaluated if areas that were classified in the past with forest fire loss in the “Global forest loss due to fire” dataset (Tyukavina et al., 2022) are expected to have higher intensities in 2021, as estimated by the CFL. This dataset has a higher spatial resolution and we only used pixels coded with moderate and high certainty in the analysis. Comparison with the Forest Fire Loss (FFL) dataset was done by summing the total area of FFL divided in three periods (before 2010; between 2010 and 2016; and after 2016) and quantifying the area of the estimated CFL divided in two classes (below and above 2.5m, based on its relationship with FLI) (section 2.5.1)
- Comparison between the estimated CFL and the FFL dataset showed that the CFL is likely providing good results in the estimated areas of high intensity fires (CFL>2.5m), thus where it is expected higher fire impacts (Table 2). A large percentage of the areas that had FFL before 2016 (above 70%) are likely to have CFL values above 2.5m, thus high intensity wildfires in 2021. In recent burnt areas (latter than 2016), most of the areas that had FFL (62.4%) are prone to lower intensity fires (section 3.1)
- We also included Table 2 next to the previous paragraph from section 3.1.

Regarding other corrections made in the manuscript we only added some corrections in the reference section.