Dear Editor,

Thank you very much for the opportunity to revise the paper based on the reviewers' comments; they are much appreciated. We thank you for the thoughtful consideration and evaluation of our work. We will modify the MS according to some of the suggestions and we think the new version will be improved by incorporating clarifications to the methods used, as well as presenting the results and their implications in a more organized manner. Below you'll find our reply to each comment, including a description of changes to be written in the MS.
All the authors agree with the changes made.

Reviewer 1

While the writing is clear, there are a few unusual word choices and I recommend that the authors include another review for English grammar and word selection.

line 22 - what do you mean by "fire spots"? Fire locations?
R: We understand that some jargon may be too local, but in our MS we consider a **fire spot** as the ignition point of a wildfire (not in the wider sense of location) catalogued by the Chilean government agency, in charge of wildfire control and statistic CONAF. We translated it from the Spanish term "puntos de incendios". However, we understand the intention of the reviewer and it will be changed in the entire text for the term "fire hotspot".

line 52 - remove the comma after "windy"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 53 - suggest changing to "…slopes rapidly lead…"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 59 - "…categories of wildfire likelihood that correspond…"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 65 - suggest removing the word "trace"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 70 - change to "manages"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

lines 70-74 - not a complete sentence
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 78 - change to "sprawl"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 80 - change to "…2010) that lead to…"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 84 - "…changes is related to more…"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 98 - change to "…which are expanding rapidly, mostly for…"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 99 - "Also, the region contains a variety of …"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 101 - "characterizes"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

lines 107-109 - the change from 44,123 hectares to 46,697 does not seem like much of an expansion to me. I would suggest saying that it has maintained a steady amount of forestry plantations in spite of urban growth.
R: Yes, and thank you for pointing that out. We'll replace that sentence with your suggestion.

line 119 - change "portraits" to "portrays"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 121 - "frameworks"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 123 - change "ingested into" to "included in"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 126 - I suggest using the term "moderate" throughout the paper instead of "median," which has connotations of calculated statistical parameters
R: Thanks for the comment, you'll see this corrected in the corresponding lines of the new version of the MS.

lines 132 and following - wildfire spots term is unusual. Maybe use "wildfire locations" throughout the paper?
R: Thanks for reiterating this. Please see our reply to L22 because the rationale is the same.

line 145 - "…network, were retrieved…"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 158 - "NDII and NDVI data entered into the …"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 169 - change "ingested" to "included" (suggest changing this word throughout the paper)
R: Thanks for the comment, you'll see this corrected in the corresponding lines of the new version of the MS.

line 188 - change to "first"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

Section 3.1 - this is a lot of discussion of numbers without reference to any figures or tables showing the actual data. I think it would help to show the data in some way.
R: Yes, we agree that part is a bit convoluted. we propose a new table to summarize the numeric information. This will be added to section 3.1 in the MS

Climatic descriptors	NDVI index surface %	NDII index surface %	elevation ranges (m)	insolation ranges (W/m2)	land use/cover surface %	Slope (surface %)	Anthropic characteristic (distance in km from hotspots)
0.59ºC/year for Tx90	9.1% presented negative values (no vegetation)	7.4% negative values (no vegetation)	0 m to 910 m	882.06 W/m2 to 905.8 W/m2.	55% exotics forest plantations	high slope (>30%)	5.5 km Streams
0.51 days/year in CDD	35.7% positive intermediate values (grasslands, shrublands, young forest plantations)	42.8% positive values (shrublands and grasslands).	average elevation 212m	insolation mean 887.7 W/m2	7% native forests	low slope (<30%)	6.6 km urban areas
-0.52º/year in skin temperature	55.2% high values (adult forest plantations, native forest).	64.7% medium-high values (adult forest plantations and native forest).			15.2% agricultural zones		1.3 km major roads and highways
Maximum cooling - 4.2º/year					5.3% urban areas		
Minimum warming +4.6º/year							

SECTION 3.1 Table XX summary of the geodatabase components

line 234 - I am unfamiliar with the term "peri-urban". Is this a standard LULC classification?
R: Thanks for catching this up. Peri urban is not a standard LULC classification. The term "peri urban" was changed to "wildland urban interfaces", you'll see this corrected in the corresponding line of the new version of the MS.

line 243 - change to "pixels"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 250 - change to "hotspot"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 253 - change to "indicate"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 277 - what do you mean by "highest sectors"? Highest elevation, or largest area?
R: The term "highest sectors" was changed to "highest elevation". Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

Figure 2 - Are the NDVI and maybe some other categories self-determined? For example, if an area is counted as burned, because it is in the database, wouldn't the vegetation be a different type than the unburned land? That is, suppose pre-burn it was vegetation type A, but post-burn it is vegetation type B (since many locations revegetate with different species). Then your statistics would naturally associated vegetation type B with burned areas and assume they are more likely to burn than areas with vegetation type A.
R: Thank you very much for your question. We hope to be sufficiently clear with the answer.
In the study area replacement of vegetation is very unlikely, since that for more than 40 years, most of the land is owned by forestry companies (see studies by Heilmayr et al. 2016 and 2020), which after a forest fire, re-establish forest plantations with species such as *Pinus* sp. or *Eucalyptus* sp. In addition, the areas that are not planted with exotic species are also covered by them, because there is an invasion of seeds from nearby areas, producing a re-growth of exotic species. In conclusion, it is very unlikely that revegetation with species other than those existing prior to the forest fire will occur.

Figure 3 - use larger font to make it readable? Also, you use the term "Medium" for the middle category here - make sure you use a consistent term throughout the paper (I suggest "moderate").
R: Thanks. We'll enlarge the font, but it is also an issue of the size of the plot. We will make it bigger. As replied in previous comments, we will change to "moderate"

line 317 - replace "recurrently" with "frequently"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 334 - remove "for"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 390 - replace "causing" with "requiring"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 397 - use "2,1 million"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 404 - explain what you mean by "some degree". I would expect that to include the medium and higher categories, which is 40%. It seems like the 90% number would even include low risk, so I don't think this statement is as strong as simply saying "this work shows that at least 40% of the CMA is subject to at least medium probability of wildfire occurrence."
R: We agree with your assessment that mentioning the 40% is better as that percentage includes the areas that really show up as being in detectable risk. We will modify that sentence with the following: **"The machine learning model developed in this work shows that about 40% of the CMA is at least in a medium probability of fire recurrence".**

line 413 - "countries, along with other..."
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 428 - "areas) when there are more plantations."
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 431 - replace "pointing" with "indicating"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 435 - "two, not necessarily"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 443 - remove contraction
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 451 - "it becomes"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

line 462 - "improve or make"
R: Thanks for the comment, you'll see this corrected in the corresponding line of the new version of the MS.

Section 5 Conclusions - I suggest including a short paragraph summarizing the statistical results indicating influencing factors. The paragraph starting at line 480 belongs more in the discussion section, rather than introducing these new ideas in the conclusions.
R: We appreciate this suggestion that allows us to better summarize the results. We decided to delete the last paragraph of the conclusion, add a new paragraph at the beginning of the conclusion, and move the current first paragraph of the conclusion to the end, with some modifications. So, the new conclusion section will read like this:

"This study aimed to develop, implement and test a model of fire risk by combining natural and human factors are associated with wildfires' generation and spread. The combination of data using PCA and SOM allowed to ponder the relative importance of each factor, interpret how interweaved they are, and study the impact of landcover. Despite observed moderate to very high recurrence tend to cluster near urban areas and on plantations, the model presents a more complex interaction among factors, where climate (e.g. t-Tx90), elevation and human aspects (d-Urban and d-Roads, for instance) are able to predict observed hotspots densities, leaving land cover as a minor component. However, the comparison of the different land cover scenarios point to a detectable influence of plantations in increasing fire risk and the spatial distribution of recurrence.

Results indicated that 12.3% of the CMA's surface area has a high and very high risk of a forest fire, 29.4% has a moderate risk, and only 58.3% has a low and very low risk. This calls for reflection on the importance of spatial planning with a resilient focus on wildfires, according to the recurrence of these phenomena in these settings as they are increasingly more forced in the WUI, urban residential areas and industrial or port areas. These maps and this model are of vital importance for the Chilean government emergency agencies as well as for the city governments within the CMA. They are also relevant for understanding how these phenomena affect the Mediterranean ecosystems to which the CMA belongs, and therefore should be beneficial for researchers in other latitudes working on similar ecosystems."

Reviewer 2

I really appreciated the article, which concerns a very important issue (the fire risk in WUI) and which needs to be thoroughly investigated. The overall structure of the paper is well balanced in its parts, the methods used are clearly presented and the assumptions made are justified. I agree with the observations made previously, especially as regards the need to present the results in figures/maps to make reading easier.
I would like to suggest that you specify a few points:

Line141-142: What are the characteristics of this database? How many fires? Are the data punctual or spatial? What is the extent of the burned areas? How are the fires distributed over the observation period?
**R:** Thanks for catching this up. Although we believe the database is well described in the methods and in the results, we likewise think that we can further specify some additional features of the fires compiled. Thus, in the location mentioned (around L141 and 142 of the current version) we will add:

"**Detected fires are GPS georeferenced and subsequently added to a GIS using a predefined grid in which each cell represents 4 km^2 (2x2 km). The minimum area for a forest fire to be mapped is 10m^2. These data thus corresponds to a spatially explicit database where each cell centroid represents burned and/or burning areas**".
Also, at the beginning of section 3.1 we will add the following:

"**From the 5404 fires recorded in the database, spatial patterns of recurrence during the period under study were associated with peri-urban areas of the CMA, since the quadrants with a recurrence of more than 20 fires were found at less than 650 m from urban centers and highways. In fact, the pixel with the maximum number of fires between 2008 and 2019, 154 or 14 per year, is located at the middle of the study area near the city of Lota (~37.10ºS and 73.13ºW), and is surrounded by a number of locations with high recurrence, above 25% (see supplementary Figure 1). This suggests that the causes of these forest fires are mostly anthropogenic. On the contrary, the areas that did not record fire outbreaks during the study period were associated with remote locations with an average elevation of 250 m, distant 8.1 km from urban centers and 1.5 km from highways**."
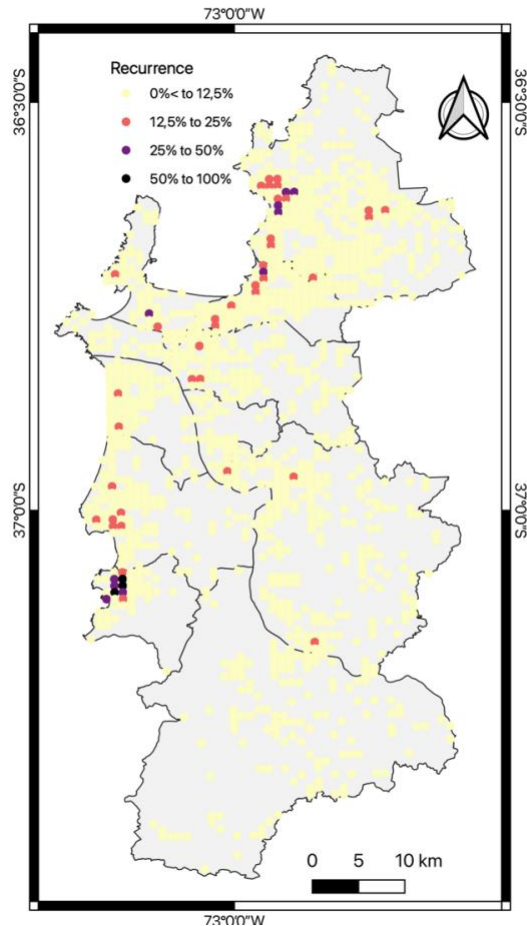
Supplementary Figure 1: Recorded fire recurrence according to the compiled database and using equation (1).

Line 234-239: has the reliability of the database been assessed? Could the increased occurrence of fires near urban areas and roads be due to a sampling bias, caused by easier detection of fires in busier areas? Do the three sub-databases (Fixed terrestrial, Mobile terrestrial and Aerial detection) have the same spatial extension and resolution?

R: These questions raise good points related to the quality of the input data and we appreciate because this allows us to provide more details in the manuscript. We respond to each question separately

(1) has the reliability of the database been assessed?: The databases with which the country's forestry services work are becoming more and more accurate and are continually being re-validated and refined. In recent years the minimum unit of detection of forest fires by private companies has reached thresholds of below the millimeter scale while the public forestry service reaches 0.001 meters. We will add this to the text proposed as response to the previous comment.

(2) Could the increased occurrence of fires near urban areas and roads be due to a sampling bias, caused by easier detection of fires in busier areas?: Although we cannot evaluate each location and thus it is impossible to rule out certain bias, we assess this situation does not affect the results considering the spatial scale of the work, because the sampling carried out in the

country by public institutions and private companies covers a large monitoring area. The use of three detection systems (fixed ground, mobile ground and aerial) leads to redundancy that minimizes false positives and negatives. Besides, the system is constantly monitoring and geo-referencing the sectors with presence of forest fires. The fixed terrestrial detection system is the most widely used in Chile, because it allows reaching vision radiuses of more than 20 km away. Here it is important to consider that no other database catches the total number of fires as this one. As it can be seen in the paper, we are aware it is important to have extra care in analyzing model results and thus we used MODIS MCD14ML as an additional source for validation. MODIS shows almost half the number of fires as the database used for the modeling, while showing a larger ratio between VH and H categories (see Table 3), that categories that the reviewer points out as related to a possible bias. While for the CONAF database this ratio is 3 (that is, density of VH is three times the density of H) for MCD14ML this ratio is 3.94. Thus, we consider the database is the least biased of the choices we had at hand. We are not adding changes to the manuscript but hope this clarifies our argument.

(3)Do the three sub-databases (Fixed terrestrial, Mobile terrestrial and Aerial detection) have the same spatial extension and resolution?: No. The mobile terrestrial detection only covers a predial scale (generally carried out by private companies) and is performed in sectors of better accessibility for different types of motorized vehicles. Fixed ground detection makes it possible to reach extensions of up to 20 km of vision and monitoring, since it uses large structures (greater than 20 meters high) in which a person is constantly watching with the help of binoculars. Finally, aerial detection allows reaching a large area per unit of time, since small airplanes are used to detect forest fires at a great distance. Fortunately, all monitoring, regardless of the detection method, is validated and entered a database with the same resolution using 2x2 km grids. We will add this extra information into the manuscript, section 2.1, near current L140.

Line 269-274: Did you also evaluate the contribution of the slope and aspect themselves, not embedded in the solar radiation calculation?

R: Thank you for the question. At first, we did think on slope and aspect as variables to include, but as you rightly point out these variables are used for the calculation of solar radiation. Both slope and aspect mainly determine the exposure of a given region to radiation that may exacerbate or inhibit fire spread. Therefore, we assessed that including them in the PCA model would only reveal a strong correlation among these variables while nothing about fire distribution. The PCA result regarding elevation and insolation further reinforces this (Figure 4), and helped us in deciding to not include these two variables. Although aspect can also be correlated with the exposure to prevailing winds that may facilitate fire spread, we have no way to get a database of wind fine enough to analyze this contribution.

References
Heilmayr, R., Echeverría, C. & Lambin, E.F. 2020. Impacts of Chilean forest subsidies on forest cover, carbon and biodiversity. *Nat Sustain* **3,** 701–709.
Heilmayr, R. Echeverría, C., Fuentes R., Lambin, E. 2016. A plantation-dominated forest transition in Chile. Applied Geography 75: 71-82.