Major comments (substantial doubts to address, which should not require enough extra work to classify the revision as "major")

45-47 ". In the second step, FFMC, DMC and DC are used to model the rate of fire spread (ISI) and the potential fuel available for surface fuel consumption (BUI)." The wind speed is missing (ISI = FFMC + WS10), and it would correlate to the POTENTIAL rate of spread (I mean, it is a variable involving the combustion of surface fuel, like dry leaves and such + the wind). Also, the BUI = DMC + DC, so it is not surface fuel consumption, but it involves the potential for a surface fire to burn the deeper fuel (Build Up) and become a much more persistent fire.

76-78 "Figure 1b shows the scatter plot of analysis and observations for all stations and every time step. While the FWI derived from the forecasted weather parameters seems to generally underestimate the FWI values compared to the values derived from the observations (slope ∼ 0.63)" AND Figure 1b: Many doubts arise from this scatterplot:

- The plot itself shows too many points to use a scatter. A density plot NEEDS to be used in this case, or two if you want to show separately the data from Finland.
- The regression which leads to the 0.63 slope seems to be off by looking at the scatterplot, which might be due to the lack of information about the point density. What causes the slope to be 0.63 (and not closer to 1, as the scatterplot would suggest)? Also, please specify the method used for the regression (I assume linear regression).
- I am aware that it is common practice to use the ECMWF analysis at minimal lead time in place of observations, but once you have found an important underestimate like you did, why did you dismiss it so fast? It seems like a very important matter that can have a huge impact on the paper's reliability. Please explain in depth why you can ignore this bias or what you did to correct it.

78-79 "a correlation is apparent. This good correlation can also be seen in the time series examples for a station in Finland and Greece": Please provide us with the necessary quantitative information (e.g., correlation coefficients for all the stations) to support this claim, especially coming right after the previous comment. Two sample stations (Figure 1c) are not enough to validate a claim on over 600 others; by this I do not mean that Figure 1c must be removed.

90-97 "Furthermore, data from all grid points in the training area is used to estimate a single set of coefficients for the given day (regional EMOS)": some additional explanation is then needed, how do you go from the μ_{kl} to the estimate used (I guess $\bar{\mu}_l$?)

Figure 2:

- It can be made clearer if the legend was more explicit (dotted line with triangles: spread (raw) / solid line with dots: RMSE (calibrated) / etc.)
- The legend, which is relative to all the three graphs must be outside the first graph. Consider also putting everything in a single column.

152-153 "In Northern Europe, the RMSE of the calibrated forecast is slightly above the RMSE of the raw forecast after 7 days of forecast," : please provide at least a hypothesis as to why this happens. The sentence on the subsequent lines "The regional differences could be explained with the generally higher FWI values in the more southern, fire prone regions compared to Northern Europe where FWI values are often very small" addresses the regional differences, but the difference from uncalibrated and calibrated NEU RMSE is not addressed.

Figure 3:

- The legend, which is relative to all the three graphs must be outside the first graph. Consider also putting everything in a single column.

Minor comments (typos and formalities)

13: "prevalent": word choice

16: "But not": cannot start a sentence with "but not"

18: unnecessary comma after "periods"

19 "heatwave 2018": either heatwave of 2018 or 2018 heatwave

24 missing (Oxford) comma after "during"

24-25: "Accurate and reliable weather forecasts ranging from a couple of days to 25 multiple weeks to identify high wildfire risk areas is an important part of SAFERS": Accurate and reliable weather forecasts, ranging from a couple of days to 25 multiple weeks, are an important part of SAFERS for identifying high wildfire risk areas. (Or equivalent paraphrasis)

25 "Here,": "in this paper,"

26 "short FWI": (FWI)

26: "Wagner, 1987": Author's last name is Van Wagner, throughout the paper

29: "One widely used": A widely used

36-39 "Although originally developed for Canadian weather and vegetation, it is used in many other regions, e.g., by the European Forest Fire Information System (EFFIS) to provide information on wildfires in the EU and neighboring counties (Giuseppe et al., 2020)": sentence needs to be more orderly and written better; also, author's last name is "Di Giuseppe"

39-40 "One advantage of using FWI is the relatively simple calculation only requiring four weather parameters in addition to information of the season (time of year) and geographical location": Rephrase, e.g. "The main advantage of using FWI is its relatively simple computation, only requiring four weather parameters and information about the season (time of year) and geographical location"

42-43 "the moisture content of three separate fuel layers of different depth and diameter": this is one interpretation of the three parameters (of course, the main one), meaning that -more or less- they contribute to the fire danger with the same time scale of a certain fuel layer. For example, the DC can also be an index of the lack of precipitation for a long time. I tend to be more cautious when interpreting these indices, but it is a relatively small issue.

48 "Often the FWI is classified": the FWI is often classified

50 "e.g. vegetation types": since this is the second time it appears, I have to point out that you cannot put "e.g." in the middle of a sentence without it being in parentheses or in a parenthetical expression (between commas). This is not formal enough for a paper, in my opinion.

52-53 "Fuel moisture codes (FFMC, DMC, DC) and consequently FWI values are dependent on preceding conditions. Thus, the preceding days noon values are used for FWI calculations and the calculations need to be initialized": To be put above, together with the input variables, and to be written more clearly.

62 "The resolution of the used TIGGE data": The resolution of the TIGGE data used in this paper

65 "can not": cannot

Figure 1 (caption): please provide a reference (IPCC) for the AR5 regions.

90 Formula (2): shouldn't it be $\log(\sigma_{kl})$?

91-92 "The logarithmic link log(sd)": define sd (ensamble standard deviation?). Besides, isn't it sd_{kl}?

119 "The bias of the forecast can be accessed by simply evaluating the difference between the average forecast and average 120 observation, which is defined as the mean error (ME)": The bias of the forecast can be accessed by simply evaluating the difference between the average forecast F_i and average observation O_i, which is defined as the mean error (ME)

137 "defined by the 6th IPCC Assessment Report (AR6 (Iturbide et al., 2020))": did you not show the AR5 regions before (Figure 1a)? It needs to be coherent, at least in the name in the caption of Figure 1a, if the regions did not change.

140 "Other regions can be selected as well, e.g. the calibration can also be done country-wise or at even smaller level.": colloquial, rephrase like "the calibration can also be performed over smaller areas (e.g., single countries)"

158 "the forecasted FWI is too low compared to observations": rephrase in a more formal "the forecasted FWI underestimates the observations[…]" or similar.