Reviewer #2

The authors applied fire-climate-ecosystem model to study the effects on increasing wildfires on summer PM2.5 over the US. Although some good results obtained, I am more worried about the real accuracy of the model simulations, especially in the irregular wildfires occurred in the western United States, which needs to be well verified. In addition, additional analyses are needed to make the results more robust.

We are thankful to the reviewer for his thorough reading and thoughtful comments. We have provided our response to each of these points below (response in **blue** and revised text of manuscript in red font).

At the outset, we want to point out that as mentioned in our manuscript, Zou et al., 2019 have extensively evaluated the control run (2000_{ALL}) simulation of our fire model that is developed under the framework of the Community Earth System Model (CESM, http://www.cesm.ucar.edu/) against observations and previous fire modeling results. These evaluation results suggest that our fire modeling ensemble results (both burned area simulations and fire smoke impacts on air quality) under the 2000 climatological conditions (i.e., greenhouse gases/GHGs, sea surface temperature/SST, sea ice concentration/SIC, etc.) are within the uncertainty range among various satellite and ground-based datasets of the same time period.

Major comments:

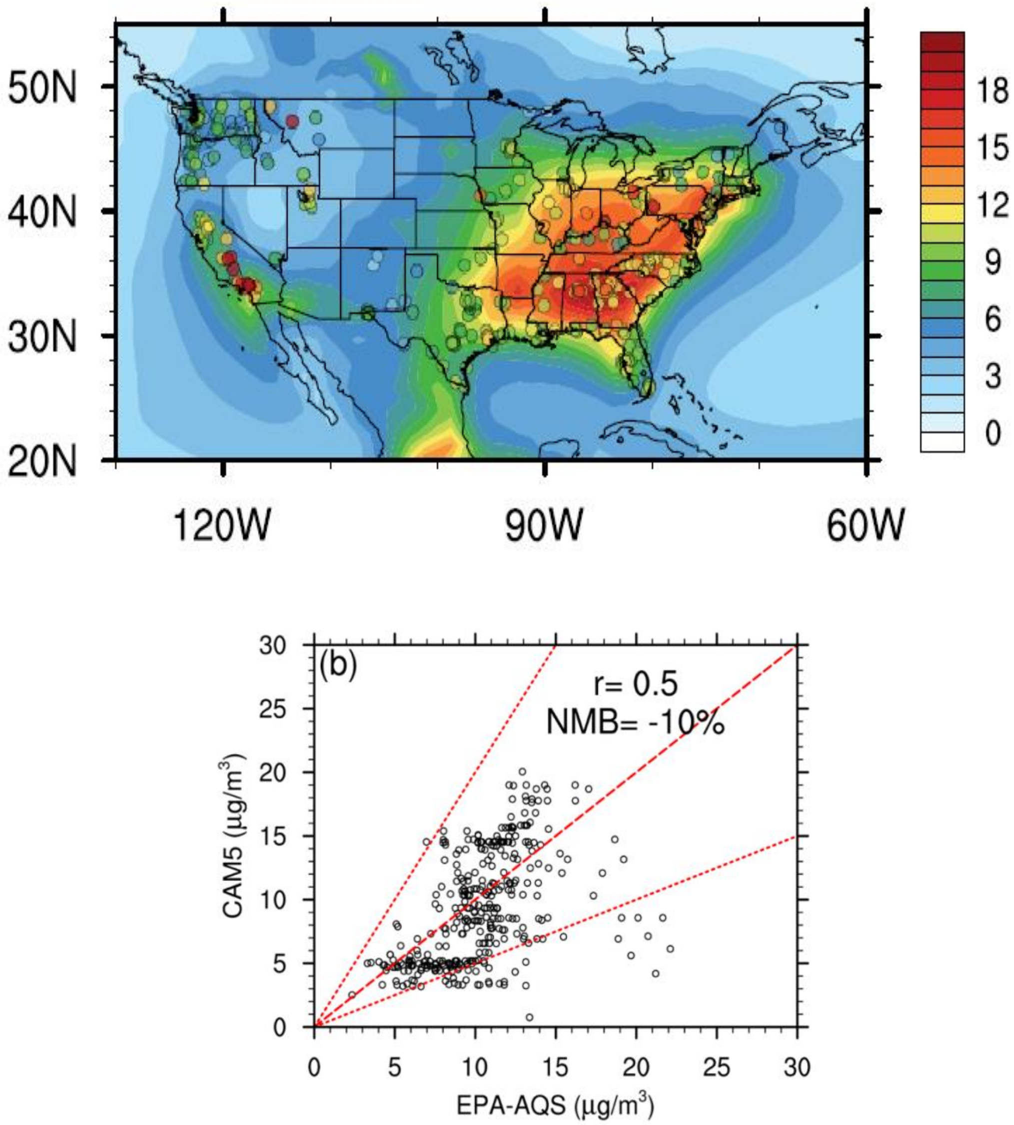
The authors are suggested to summarize the previous studies related to PM2.5 changes in the US, especially those focusing on wildfires in the Introduction.

My biggest concern is the accuracy of the model simulation results, especially for the western US, which is essential for the current study. The author simply compared the model results with a single satellite remote sensing product. Note that the sampling frequency of these two is different since there are a large number of missing values in satellite products under cloudy conditions. I suggest using ground-based observations to evaluate the model simulations from different spatiotemporal scales over North America since a rich ground-based observation network of PM2.5 concentrations and its components are available, e.g., EPA, and IMPROVE, etc.

In the revised manuscript, we have included the evaluation of simulated PM2.5 over the continental US against IMPROVE observations and the following figures and paragraph is added.

The simulated $PM_{2.5}$ has also been evaluated against the ground-based Interagency Monitoring of Protected Visual Environments (IMPROVE) data, showing similar spatial pattern and biases (10-25%) (Supplementary Figure 2). The biases are smaller over Eastern US and Southwestern US region. The simulated PM2.5 values over California matches quite well with the observed annual mean values. However, the biases over Northwestern US region are ~30-40%, a portion of which could be attributed to possible biases in model's meteorology in northwestern US region. Nonetheless, both satellite and in situ evaluation indicate that our simulation biases are largely within the uncertainty range among the various satellite and ground-based datasets, which have normalized mean biases ranging from -3.3%

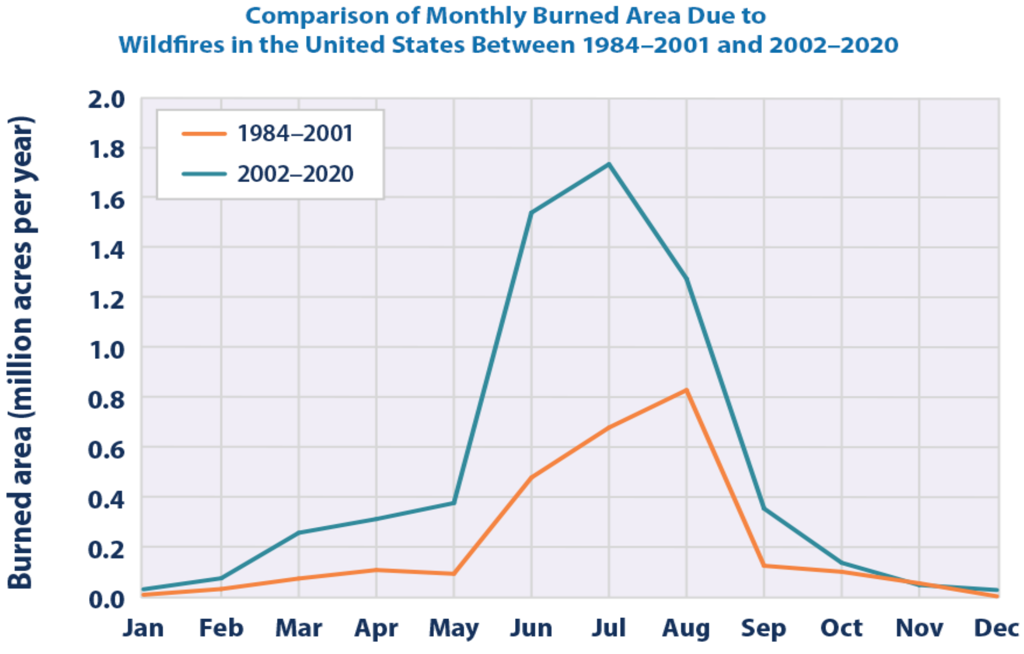
to 33.3% when benchmarked against the ground-based IMPROVE data over the contiguous US (Diao et al., 2019; Val Martin et al. (2015)).



Supplementary Figure 2: Evaluation of the CAM5 simulated surface *PM2.5* ($\mu g\ m^{-3}$) with fire emissions from RESFire. Comparison of decadal-averaged (2001-2010) annual surface *PM2.5* between the simulation (shading) and the Interagency Monitoring of Protected Visual Environments (IMPROVE) data from US EPA Air Quality System (AQS) in situ observations (colored circles) in the 2000s; (b) Comparison of simulated and in-situ measured annual mean surface *PM2.5*.

Another concern is that compared with summer (JJA), wildfires in recent years mostly occurred in dry autumn, e.g., California fire in 2020, which has been burning for nearly two months (September and October).

We agree that the wildfires in California in recent years are occurring in September and October, but at continental US scale, wildfire occurrence peak is actually shifting from August in 1990s to July month in 2000s. Monthly wildfire-induced burnt area statistics from EPA (Figure below) shows that JJA are the months of highest occurrence. Our simulations also replicated similar monthly variations. Hence, we used summer (JJA) months for focus of our analysis. Moreover, this monthly variation is dominated by fires in western US.



Data source: MTBS (Monitoring Trends in Burn Severity). 2022. Direct download. Accessed April 2022. www.mtbs.gov/direct-download.

For more information, visit U.S. EPA's "Climate Change Indicators in the United States" at www.epa.gov/climate-indicators.

Section 3.1: Suggest comparing different satellite PM2.5 and component datasets over the US since many open datasets are available. In addition, it is suggested to add a validation of a single wildfire year or month, e.g., a severe and continuous wildfire occurred in western US in September 2020.

In this experiment, our aim was to understand the changes at climatic scale, so we saved our simulated output values at a resolution of 10 year mean monthly values only, to reduce the huge expense associated with storage and file movement for analysis. Hence, we are not able to do monthly scale analysis for any specific year.

Nonetheless, as also mentioned in our manuscript at line number 185-190, the present day simulated wildfire burnt area and wildfire spatial pattern was comprehensively evaluated in the preliminary papers of CESM-RESfire simulations in Zou et al., 2019.

Zou et al. (2019) performed comprehensive evaluation of the RESFire simulated wildfire burnt area distribution, associated carbon emissions and terrestrial carbon balance to demonstrate reasonable model skill. Zou et al. (2020) compares global fire simulations by CESM-RESFire with modeling results reported in the literature to show better agreement with the GFED4.1s benchmark data and predicts more prominent changes in the future than those predicted by Kloster et al. (2010, 2012). These differences might come from differences in the climate sensitivities of the fire models and scenarios and other input data used to make future projections.

Section 3.2: It will be very interesting to take a look at the Fire Burnt Area, PM2.5 changes, and related contributions caused by wildfires in the past decade (2010-2020). Wildfires in the west of US have been burning more frequently in recent years, much higher than in the first decade.

We agree that wildfire burnt area and impact in the west of US has enhanced in recent years (2000-2020) compared to 1980-2000 and is projected to even more increase in the coming decades due to warming. However, the wildfire burnt area and impact over the US between the two recent decades of 2000-2010 and 2011-2020 is more or less same. In fact, the recent EPA report (https://www.epa.gov/climate-indicators/climate-change-indicators-wildfires) illustrate that the wildfire frequency over the US in 2010-2020 is a bit less compared to the same in the decade 2000-2010.

Here, we had taken two representative decades of present climate (2000-2010) and future climate (2050-60) and made sensitivity runs (with and without wildfire effect) for both the periods, so total eight 10-year global simulations. The comparison of these two representative periods provides us the qualitative understanding on how air quality in the eastern US be affected in future under climate change induced wildfire enhancement at climatology scale. Due to computational constrains, we are unable to simulate another decade 2010-2020. Also, we feel that these additional simulations will not significantly change our conclusions and understanding.

Minor comments:

Line 172: Please spell out the JJA and check such issue throughout the paper.

We have revised.

Line 193: Figure 1

Revised.

Lines 216-217: Again, sampling frequency is also a potential reason resulting in the differences that should be discussed.

We have included.

First, the satellite-derived data has a non-zero lower bound of $PM_{2.5}$ concentrations, so the ambient background concentrations for relatively cleaner regions such as the western US may be overestimated (Figure 1C), also the sampling frequency between these datasets are different.

Figure 5: Confusing. Are these temporal trends? If not, what the significant confidence level here used and how to calculate?

All the data shown in Fig.5 are fire-induced differences between the 2050 scenario and the 2000 scenario as indicated in the caption. They are not temporal trends. We subtract the fire effects in the 2000s from the fire effects in the 2050s [(2050ALL-2050WEF) – (2000ALL-2000WEF)] to evaluate the fire-induced changes from the 2000s to the 2050s. The significance of these changes is calculated using the Student's t-test between the two simulations.