

This revised paper is well written and well reasoned but in previous round of review, several key issues were pointed out. Specifically in the last round of review, the editor posed three key questions that the authors needed to address in the manuscript with which I agree.

- 1) What is the goal of the paper - to consider high air pollution conditions, and for what purpose?
- 2) Why were the episodes chosen, and what do we hope to learn more generally for other conditions?
- 3) How do the conclusions based on the CAPA episodes relate with other conditions during the years studied (2014-16) and to what extent are they relevant for other years before or after?

I believe that the authors have fully answered question 1 and the clear motivation has improved the manuscript. However, the current revision only partially answers questions 2 and 3. My major and minor comments are as follows:

Major Comments:

1a) The authors addressed the questions about other conditions by conducting 3 additional simulations with conditions other than complex air pollution episodes. These simulations are important and reveal some interesting findings. Particularly, the O₃ decrease in a high PM episode from API and ARF is close to double the impact in a complex episode and in cases without high PM pollution the impact is close to half the impact of a complex episode. However, the authors treat these cases as an afterthought with them being mentioned only in the last few lines of the conclusion/supplementary material. These cases need to be incorporated into the main text as another results section and they need more descriptive names such as: High PM_{2.5} (HI_PM), Low Pollution (LOW_POL), and High O₃ (HI_O3).

1b) The authors should also consider possibly simulating an additional case with each of these conditions to see if the differences between them and the CAPAs are robust, or alternatively they should demonstrate that the episodes they selected are representative of those conditions throughout the period of interest (2014-2017).

2) The authors have still not addressed historical conditions before 2014. The authors have stated this is because national observations are not available to pin-point complex air pollution episodes before 2013. However, the authors have also stated "Air pollution in China was characterized by high concentrations of PM_{2.5} before 2014 (Li et al., 2019a; Zhang et al., 2019) and by synchronous occurrence of high PM_{2.5} and O₃ or high levels of O₃ alone after 2017 (Dai et al., 2021; Li et al., 2019b; Li et al., 2020; Qin et al., 2021)." If there are no observations before 2013 to justify this statement, how is this statement supported? At a minimum, the authors should point out the lack of data before their period of interest in the manuscript as a reason they cannot explore the magnitudes of API and ARF before 2014. However, if the authors are at least aware of a high PM_{2.5} case from the 2001-2005 period it would be worth comparing that with the current HI_PM case to see if the magnitude of the API and ARF impacts have decreased in a significant way.

Minor comments:

Line 78: field not "filed"

Line 84: Should the unit be $\mu\text{g m}^{-3}$ or ppb?

Line 206: done instead of "did"

Line 291: should be "was caused"

Line 414: How can API contribute more than 100% to a reduction in O_3 ? The word contribution implies "percent of" not "percent change". If these values are a relative difference (i.e., percent change) then a word other than contribute needs to be used.

Line 430: during "the" warm season