Interactive comment on "Modeling the smoky troposphere of the southeast Atlantic: a comparison to ORACLES airborne observations from September of 2016" *by* Yohei Shinozuka et al.

Anonymous Referee #2

This paper presents a statistical comparison of aircraft observations of smoke aerosols along repeated sampling tracks from the 2016 deployment of the ORACLES campaign against a variety of model simulated aerosols for grid cells along the same sampling tracks. Few field campaigns provide sufficient sampling to allow for such a comparison and the authors go to some lengths to demonstrate that the observations are indeed representative of the monthly-mean aerosols along the sampling tracks. There is no perfect way to perfume such a comparison. But for a minor comment on the screening of the data, I am satisfied with the approach.

The greater challenge for this paper is arriving at some generalized results that can

guide the modelers. At the root of the challenge is that models may have many deficiencies that contribute to errors in the representation of the aerosol plumes, from errors in emissions to errors in transport, and uncertainties in the appropriate aerosol particle sizes and optical properties. There are only a few clues as to which errors might be contributing to the biases documented in the paper, so the end result is an illustration that all of these sources of model error contribute to causing a wide spread in the resulting aerosol distributions and physical properties among the models. This information is certainly worth sharing with the community, and this is exactly the kind of effort we should hope to see when we have high-quality datasets such as that from ORACLES. I think this paper would be suitable for publication if the authors can draw a stronger connection between the general limitations of the models discussed in the introduction as motivation for the paper and the results that they found. Thus, the discussion at the end of the paper should state how the results relate to specific shortcomings in the models in the literature as summarized in the manuscript. In the absence of drawing this connection, the paper just seems like a list of various model-data differences with no coherent interpretation or generalized outcome that the reader can take away from the study.

Other comments:

The abstract claims a "new approach to utilizing airborne aerosol measurements", but is not explicit about what aspect of the study the authors are claiming is new.

Is there a citation or other evidence to support the use of "altitudes below (RH(%)-60)*40m to define the boundary layer depth?

The grey points in figure 2 are apparently observational values that could not be successfully placed in one of the three altitude classification. I presume these data are not included in the comparison with the model. Is there a sampling bias related to this? In particular I would think that the low altitude data points shown in grey, presumably corresponding to cases where the top of the boundary layer is too difficult to discrimi-

nate, do represent a condition that happens with some regularity. Shouldn't the models reproduce a similar condition occasionally?

Can the authors draw some connections between the systematic biases in the thickness of the aerosol layer and the extinction optical properties of the particles? Are there some known deficiencies in the aerosol radiative forcing or fluxes of any of these models that could be tied to the biases in plume thickness and optical properties reported by the authors? Do the biases the authors have found tend to reinforce on another in magnifying errors in the bulk radiative effect of aerosols, or perhaps are there some compensating errors? Answering these questions would help clarify what has been learned from quantifying all of these biases.

—————————————————