This manuscript discusses Na-layer lidar observations taken at four stations in China: Beijing, Hefei, Wuhan, and Haikou. All lidar stations are around 110E and span the latitudes from about 19N to 40N. These dataset offers a unique opportunity to study Na-layer and its relation to dynamics, chemistry and electrodynamics. The lidar measurements are then compared to satellite observations and model simulations with WACCM-Na, a version of a climate model that includes. The main conclusions of the study are a general agreement between observations but with significant discrepancies at times attributed to the lack of space-borne observations during norther winter in the darkness; and, a general disagreement between the model on the observations in terms of variability. In particular, the authors suggest that model variability can be improved with a better representation of the spectrum of subgrid gravity waves.

A few comments/queries below:

- Line 122: Please define what S4max (or S4 index) is for the non-experts.
- 4 & 6. Please use the same contour min/max for ease of comparison.
- Line 170. Could the horizontal resolution be the problem in resolving those fine structures/peaks?
- Line 181: Is it sub-sampling by 5 days or is it a 5-day average?
Figure 8: I seem to detect a quasi-biennial oscillation. Any thoughts?

Line 187: replace less with few

Figure 9c: I almost missed the vertical bars: maybe add the station name for clarity?

Line 203-206: I don’t understand what is implied here. Is the suggestion that lack of field line transport is a potential issue for WACCM-Na? How’s that so? I thought WACCM transports ions as part of the chemistry package.

Line 268: The sentences need some clarifications. I think the authors want to say that the limited resolution of sub-grid processes could explain the lack of variability. I buy that. However, the sentence seems to indicate that WACCM is missing all gravity wave and turbulence. And I don’t buy that.

Figure 14: Isn’t the seasonal variability also different between OSIRIS and CMP lidars?

Lines 349-351: The disagreement is more than slight. More importantly, how is it that fewer observations result in a high bias?