Description of data processing to reproduce the observational results in Fan et al. 2018 (https://doi.org/10.1126/science.aan8461)

Yuwei Zhang and Jiwen Fan

The document describes the data processing details for reproducing the results from the observational part of Fan et al. Science, 2018 (Fan18). The data in Fan18 was analyzed by Yuwei Zhang who was a Ph.D. student in late 2016, and the record was lost due to the job change of Yuwei over the past 7-year period. In a recent study Öktem, Romps, and Varble, J. Atmso. Sci., 2023 (ORV23), they claimed to try to reproduce our results (but without working with us in data processing), and claimed different results. Here, we reanalyzed the 17 observational cases based on the information available in Fan18 and the scientific rationales and reproduced the results following the details below.

Step 1. Selection of 17 local convective system cases. We first selected the local occurring system (LOS) cases identified as occurring in the afternoon and characterized as scattered convections generated through solar heating at the surface based on Tang et al. (2016). Tang et al. (2016) identified from the surface radar and satellite images and did not describe the specific criteria used. While working for Fan18, we communicated with Dr. Shuaiqi Tang and documented the criteria used for identifying these cases: (1) convection occurred between 11:00-19:00 local time and (2) no convection occurred during the 3 hours preceding the locally occurring systems. The second criterion excludes cases with previous convective clouds, which are likely from the propagating systems. To ensure deep convective clouds, another criterion that the maximum echo height determined by reflectivity > 0 dBZ is higher than 10 km was applied. Convection is defined based on the Radar Wind Profilers (RWP) radar data (also see Step 2 below). To look at aerosol-cloud interaction, we also excluded cases having bad aerosol data based on the measurements from the Scanning-Mobility Particle Sizer (SMPS; available form the DOE ARM Data Archive https://adc.arm.gov/discovery). The 17 selected cases are shown in Table 1.

Step 2. Identification of the convection period and the analysis of the RWP data. The convection initiation was defined based on the RWP data (available from the DOE ARM Data Archive) and the following criteria: Ze > 0 dBZ above 4 km. The convection end time was defined as the time when there is no Ze > 0 dBZ between 1.5-4 km altitudes. From the convection initiation to the end of the convection we get the convective period shown in Table 1 for all 17 cases. For each case, we calculated the profile of the mean of top 10^{th} percentiles of updrafts > 0 m/s over the identified convection period. Only the updraft profiles at the time when echo classification is defined as "convection" or "weak convection" in the RWP dataset are considered. We excluded the times with data outside of the range from - 40 to + 40 m s^{-1} from the analysis based on the recommendations of the data developer, Dr. Scott Giangrande. The vertical profile of the top 10^{th} percentile updrafts is calculated first, then is averaged over the cases in every aerosol group. The moving average technique is applied to smooth values vertically. We noticed in a case located at the second aerosol group, 20140423, that there are many values larger than 30 m s^{-1} at the cloud bases and edges in the PBL before 18:36 UTC, thus these data were also excluded from the analysis.

Step 3. Identification of aerosol sampling period and analysis of aerosol data. For each case, we selected a rough 30-min aerosol sampling period. The aerosol data are from SMPS available from the ARM archive. The 30-min sampling window is within the 2-h period ahead of the convection initiation time. We need the environment aerosol concentration with the least cloud interference; thus, the 30-min aims to lie in the clear-sky condition if possible, to capture the high aerosol concentrations before the convective storms. If it is cloudy all the time over the 2-hr period, we selected a time window with the highest aerosol concentrations. Therefore, the 30-min time window is manually identified for each case based on this scientific rationale. The aerosol sample periods for all 17 cases are provided in Table 1.
After the 30-min aerosol sampling window was identified, the mean aerosol number concentration was obtained for each case. The four aerosol groups were obtained based on the mean aerosol concentrations of the 17 cases. For each aerosol group, the mean profile of updraft velocity is obtained by averaging of the profiles (mean of top 10^{th} percentiles, obtained in Step 2) of the cases in the group.

Table 1 Sampling time periods for aerosol and updraft velocity

Date of Case	Aerosol period (UTC)	Convection period (UTC)
20140311	1332-1401	1440-1852
20140317	1637-1706	1728-1906
20140321	1902-1931	1935-2028
20140322	1332-1401	1508-1654
20140323	1642-1711	1712-1930
20140326	1617-1646	1817-2235
20140401	1109-1138	1309-1905
20140412	1533-1602	1642-1807
20140418	1916-1945	2021-2136
20140420	1751-1820	1824-2015
20140421	1602-1631	1633-1802
20140423	1712-1741	1803-1910
20140516	1615-1644	1645-1815
20140519	1520-1549	1550-1835
20140520	1545-1614	1618-1742
20140530	1333-1402	1455-1809
20140531	1239-1308	1401-1503

With the above-described steps, the plot of Figure 2B for D > 15 nm in Fan18 is reproduced (Figure 1b vs. 1a here). The increase of vertical velocity from the low to high aerosol groups is clear. There are small quantitative differences (minor details in processing the data cannot be recalled).

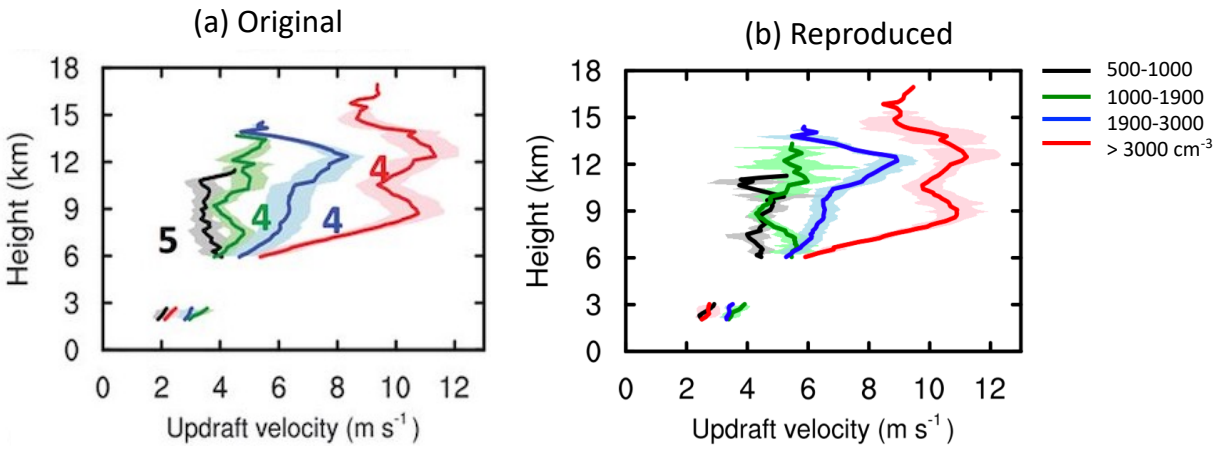

Fig.1 Vertical profiles of the updraft velocity averaged over the top 10 percentiles of updrafts for the four aerosol groups from lower to higher N_a for $D > 15$ nm: (a) the original plot in Fan18 and (b) the reproduced plot. The numbers on the lines of (a) denote the number of cases for each aerosol group which remains the same for (b).

The NCL code for processing the data is included in the file "code.ncl" of this repository. The input for running the code is "input_time.txt".