Review of “Impacts of aerosols on the tornado potential: A case study in Yangtze River Delta, China”

Recommendation: Major revision

General assessment

This manuscript examines how anthropogenic aerosols may affect the tornado potential of a Yangtze River Delta supercell using WRF-Chem ensemble simulations with radar and surface-data assimilation. The topic is important, and the use of a real severe-weather case, paired aerosol simulations, data assimilation, and detailed process diagnostics is potentially valuable.

However, the central causal chain—from aerosols to enhanced low-level condensation, stronger buoyancy and updraft, increased storm-relative inflow, and larger STP—is not yet demonstrated. Several major model and methodological uncertainties directly affect this interpretation. I recommend major revision.

Major comments

1. The approximately four- to fivefold precipitation overestimate directly compromises the proposed mechanism

Figure 3 indicates that the simulated precipitation rate during the main storm period is approximately four to five times the observed rate. The rain area is also overestimated, and the simulated cold pool is likely too cold due to the excessive rainfall.

This is not a peripheral model deficiency. The proposed mechanism depends on precipitation, evaporation, low-level moistening, cold-pool evolution, cloud-base height, updraft position, storm-relative inflow, SRH, and STP. Excessive precipitation can distort every link in this chain.

Moreover, aerosols do not appear to make a large difference in the overall precipitation shown in the manuscript. This raises a basic concern: if ALL and NoAero produce broadly similar precipitation amounts, the authors must demonstrate quantitatively how aerosol-induced spatial redistribution of precipitation produces the large differences in evaporation, cloud-base height, latent heating, and cold-pool–updraft configuration that underpin the proposed mechanism.

2. The simulations do not demonstrate the warm-rain suppression required by the proposed aerosol mechanism

The manuscript invokes aerosol suppression of warm-cloud precipitation efficiency, but it does not show that warm-rain production is significantly smaller in ALL than in NoAero. Indeed, the available precipitation results suggest little overall difference between the polluted and clean simulations.

The authors must quantify, for paired ensemble members, the ALL–NoAero differences in autoconversion, accretion, rain production below the freezing level, rain-initiation height and time, and warm-process surface precipitation. These differences should be evaluated across the paired ensemble members, including their magnitude, sign consistency, and uncertainty, rather than shown only as ensemble means.

They must also separate surface precipitation into rain produced entirely by warm-cloud collision–coalescence and rain originating from melting ice-phase hydrometeors.

If warm rain is significantly suppressed in ALL while melted precipitation increases and compensates for it, the proposed pathway may be physically consistent despite similar total rainfall. However, if significant warm-rain suppression cannot be demonstrated, the central aerosol-invigoration premise is unsupported in these simulations. In that case, showing similar total precipitation or merely different spatial distributions does not establish the claimed microphysical mechanism.

The approximately four- to fivefold rainfall overestimate makes this test especially important, because it raises the possibility that the Morrison two-moment scheme produces excessive warm rain in both simulations and is insufficiently sensitive to aerosol loading.

3. Cycling data assimilation may modify the aerosol-induced differences

Although the same assimilation procedure is applied to ALL and NoAero, their background states differ because of the aerosol perturbation. Because the ALL and NoAero background states differ, assimilating the same observations yields different observation-minus-background departures and therefore different analysis increments in the two experiments. Consequently, the ALL–NoAero differences present at 17:00 may reflect not only aerosol-driven model evolution but also the differential response of the assimilation system to the two backgrounds. The authors should show the analysis increments and ALL–NoAero differences throughout the assimilation period, particularly at 17:00. A cleaner sensitivity experiment would branch the two simulations from the same final meteorological analysis and alter only the aerosol conditions thereafter.

4. The 19:00 condensation-heating difference does not establish aerosol-induced low-level invigoration

Figure 5 implies a cloud base near 0.8–0.9 km AGL, so most of the claimed 0.3–1 km condensation-heating layer is normally subcloud. Yet Figure 9 shows substantial ALL–NoAero condensation-heating differences within this layer at 19:00.

This indicates that the simulations already differ in cloud-base height, cloudy fraction, updraft position, or all three. Earlier rainfall evaporation can moisten the low levels, lower the cloud base, and permit condensation at lower altitudes. The heating difference at 19:00 may therefore be a consequence of earlier nonlinear storm evolution rather than the initiating aerosol effect.

The abrupt reversal of the updraft response near 18:50 reinforces this concern. The aerosol perturbation does not suddenly change at that time, so the reversal likely reflects nonlinear differences in storm evolution, such as cell interaction, gust-front position, cold-pool development, or updraft relocation.

The interpretation also differs from Fan et al. (2018), where enhanced activation and condensation occurred well above the cloud base after coalescence had already reduced droplet surface area and created high supersaturation. It does not support direct aerosol-enhanced condensation throughout a largely subcloud 0.3–1 km layer.

The authors should show the time evolution of cloud-base height, low-level humidity, cloud fraction, cloud water, condensation, rain evaporation, buoyancy, and vertical velocity beginning well before tornadogenesis. They must demonstrate that the buoyancy anomaly precedes and quantitatively explains the updraft anomaly.

5. Independently tracked updraft cores may bias the comparison

ALL and NoAero updraft cores are tracked separately using local maxima and percentile thresholds. The averages may therefore represent different locations, cloudy fractions, core sizes, and stages of storm evolution.

This is especially important for Figure 9, where a larger heating rate may simply result from sampling more cloudy grid cells, a lower cloud base, or a more concentrated portion of the updraft.

The key diagnostics should be repeated using common storm-relative coordinates and identical masks applied to both simulations. The results should also be tested using union, intersection, and explicitly paired storm-object masks.

6. The individual microphysical heating terms are not independently closed

WRFlux closes the total potential-temperature budget but does not separate condensation, evaporation, deposition, and other subprocesses. The archived subprocess tendencies are rescaled to match the WRFlux total.

This guarantees closure of the sum, not correctness of the individual partitioning. Because the proposed mechanism depends specifically on condensation and rain evaporation, the authors should show the rescaling formula, correction magnitude, unscaled and scaled profiles, and sensitivity of the conclusions to this procedure.

7. The vertical momentum budget requires verification and demonstrated closure

Equation 6b appears to use second spatial derivatives of velocity for the extension-pressure source, whereas the standard formulation contains squared first velocity gradients. This is not only a notation issue: the printed right-hand side has dimensions of $M\ L^{-4}\ T^{-1}$, while the pressure-Laplacian term on the left has dimensions of $M\ L^{-3}\ T^{-2}$. The equation is therefore dimensionally inconsistent as written. The authors must clarify whether this is a typographical error or the expression actually implemented. If the printed formulation was used in the calculations, the pressure and acceleration analyses must be repeated.

They should also demonstrate complete momentum-budget closure, include a residual, define the base state and perturbations, identify omitted tendencies, and test sensitivity to the pressure-inversion domain and boundary conditions.

8. The ensemble analysis does not establish robust aerosol effects

The main value of the ensemble is to test whether the ALL–NoAero response persists across plausible perturbations of the modeled storm. The present ensemble analysis does not establish such robustness.

Eighty ensemble members were generated, but only the 70 judged to agree best with observations were retained. Because the reported ensemble-mean aerosol differences are relatively small, excluding 10 members —particularly extreme members or members with an opposite-sign response—could materially determine the mean, median, sign consistency, and statistical significance. The exact selection criterion must therefore be stated, together with whether identical paired members were retained in ALL and NoAero.

The reported Wilcoxon p-values are not adequate evidence of robustness. The signed-rank test assumes independence among the paired ALL–NoAero differences, but the manuscript does not establish that this condition is satisfied. The members share the same case, model, physics, boundary conditions, assimilated observations, and ensemble-generation framework, so substantial dependence among their responses is possible. Treating dependent members as 70 independent samples can greatly overestimate the effective sample size and yield artificially small p-values. Moreover, even a formally small p-value would not demonstrate that the effect is physically substantial, consistent across members, or insensitive to the exclusion of the 10 members.

For each main link in the proposed mechanism—precipitation, evaporation, cloud-base height, condensation heating, buoyancy, updraft strength, storm-relative inflow, SRH, and STP—the authors should show the paired ALL–NoAero distribution, median and effect magnitude, fraction of members with the same response sign, and results for both all 80 members and the selected 70-member subset. They should also show how the mean, median, confidence interval, sign consistency, and any statistical measure change when the 10 members are excluded.

Unless the inferred aerosol effects remain consistent for the full 80-member ensemble and are insensitive to the selection procedure, the ensemble cannot be claimed to demonstrate robust aerosol impacts in this case.

9. The model diagnoses tornado potential, not the tornado life cycle

At 3-km resolution, the tornado vortex is not resolved. The statement that the model reproduces the “whole life cycle of a supercell tornado” is therefore too strong.

The manuscript should consistently distinguish simulated STP and tornado-favorable conditions from actual tornado occurrence, intensity, and tornadogenesis.

Recommendation

The manuscript addresses an important problem, but the present analysis does not establish the proposed aerosol-to-tornado-potential mechanism. The precipitation bias, lack of demonstrated warm-rain suppression, assimilation design, nonlinear storm divergence, tracking method, heating attribution, momentum-budget formulation, and lack of demonstrated ensemble robustness must be addressed.

Recommendation: Major revision.