Dependence of Deep Convective Cell Properties on Meteorological and Aerosol Conditions during TRACER

Dhwanit J. Mise[a,b,c], Yongjie Huang[a,c], Greg M. McFarquhar[b,c], Ming Xue[a,c], Alexander V. Ryzhkov[b,d], Jeffrey Snyder[d], Lulin Xue[e], Mariko Oue[f], Pavlos Kollias[f,g], Michael P. Jensen[g], Mateusz Taszarek[h]

[a] *Center for Analysis and Prediction of Storms, University of Oklahoma, Norman OK 73072, USA*

[b] *Cooperative Institute for Severe and High-Impact Weather Research and Operations, University of Oklahoma, Norman OK 73072, USA*

[c] *School of Meteorology, University of Oklahoma, Norman OK 73072, USA*

[d] *NOAA/OAR National Severe Storms Laboratory, Norman, OK 73072, USA*

[e] *NSF National Center for Atmospheric Research, Boulder, CO 80307, USA*

[f] *School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY, USA*

[g] *Environmental Science and Technologies Department, Brookhaven National Laboratory, Upton, NY, USA*

[h] *Adam Mickiewicz University, Poznan´, Poland*

Corresponding author: Yongjie Huang, Yongjie.Huang@ou.edu; huangynj@gmail.com

Early Online Release: This preliminary version has been accepted for publication in *Journal of the Atmospheric Sciences*, may be fully cited, and has been assigned DOI 10.1175/JAS-D-25-0130.1. The final typeset copyedited article will replace the EOR at the above DOI when it is published.

ABSTRACT

Deep convective cells significantly influence Earth's energy balance and water cycle. However, their accurate representation in numerical models remains challenging due to their small spatiotemporal scales and limited observational constraints. This study examines over ~400 deep convective cells near Houston, observed by a dual-polarization C-band radar during the Tracking Aerosol Convection Interactions Experiment (TRACER) Intensive Operational Period (June–September 2022). Cells are categorized by lifetime into short- (< 40 min), intermediate- (40–80 min), and long-lived (80+ min) groups. Long-lived cells were broader (~13.2 km at 2–4 km height) and deeper (~11.4 km) than short-lived cells (~6.4 km width, ~7.31 km height). Using Random Forest (RF) modeling and correlation analyses, precipitable water vapor (*PWV*), 2–6 km lapse rate, 0–8 km bulk shear, and fine aerosol mass concentration (*Mass_f*) are identified as key predictors of cell lifetime. Higher *PWV* is associated with significantly longer convective cell lifetimes compared to the low-*PWV* group, particularly within low *LR_26km*, moderate-to-higher *BS_08km* and low-to-moderate *Mass_f* environments. RF analysis also identifies low-level (0–2 km) equivalent potential temperature, *PWV*, *Mass_f* and surface latent heat flux as key predictors for cell width and height. Short-lived cells have higher aerosol number concentrations (500–1000 nm size range), linked to onshore wind conditions and marine aerosols, however, their low concentration suggests the sensitivity may reflect associated meteorological regimes rather than a direct aerosol effect. Long-lived cells have higher concentrations of organic and sulfate aerosols, while short-lived cells exhibit higher black carbon concentrations. These results highlight the intricate dependence of convective cell lifetimes and structure on environmental moisture, thermodynamics, wind shear, and aerosol characteristics.

SIGNIFICANCE STATEMENT

Accurately representing deep convective cells in numerical models remains challenging due to their small spatial scales and complex interactions with aerosols. This study examines over ~400 deep convective cells near Houston using C-band radar and other co-located instruments from the TRACER field campaign to investigate how meteorological and aerosol conditions influence cell lifetimes. Results indicate a strong dependence of convective cell lifetime on moisture, 2-6 km temperature lapse rate, 0–8 km wind shear, and fine aerosol mass concentration (*Mass_f*), with longer-lived cells being deeper and wider. Moisture-induced

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

increases in lifetime are evident across both low and moderate *Mass_f* conditions, with higher *PWV* consistently associated with longer-lived cells. Long-lived cells have higher concentrations of organic and sulfate aerosols, while short-lived cells exhibit higher black carbon concentrations. This work highlights the influence of both meteorological conditions and aerosol properties on convective cell properties.

1. Introduction

Deep convective cells are vital to Earth's hydrological cycle (Trapp et al. 2009) and energy budget via latent heat release (Tao et al. 2010; Koren et al. 2014). Through the vertical transport of heat, moisture, and momentum, they drive large-scale atmospheric circulations and influence atmospheric predictability (Sanderson et al. 2008; Del Genio 2012) highlighting their critical importance in Earth's atmospheric dynamics.

Convective cell properties are strongly influenced by environmental factors such as wind shear, atmospheric stability, and moisture content (e.g., Weisman and Klemp 1982; Doswell 1987; Markowski and Richardson 2010; Trapp 2013; Zöbisch et al. 2020; Wilhelm et al. 2023). The interactions among these environmental factors and aerosols further complicates the manner in which cloud microphysics and precipitation processes are modulated by aerosols (Fan et al. 2009, 2016). Aerosols, acting as cloud condensation nuclei (CCN) and ice nucleating particles, can alter droplet concentrations, sizes, and the freezing process (e.g., Rosenfeld 2000; Rosenfeld et al. 2008; Huang et al. 2025), thereby modify convective updraft strength through changes in latent heating (Rosenfeld et al. 2008; Fan et al. 2016; Fan et al. 2018; Igel and van den Heever 2021; Peters et al. 2023a; Varble et al. 2023). However, separating these aerosol effects from meteorological influences remains a major challenge (Tao et al. 2012; Grabowski 2018; Varble et al. 2023), as these factors often co-vary and frequently work together to influence convective cell dynamics in varying ways and amounts.

Cloud-aerosol interactions are fundamental to understanding the complexities of convective cloud processes and their impact on larger-scale weather systems. These interactions affect the lifecycle of convective clouds, including thermal evolution, hydrometeor size distributions, lightning activity, and precipitation efficiency, which are difficult to observe and accurately represent in models due to their rapid evolution and small spatial scales (Ladino et al. 2017; Fridlind et al. 2019; Lamer et al. 2023). Although advances in observational technologies have provided valuable insights about aerosol activation (Pruppacher and Klett,

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

1997; Seinfeld and Pandis,1998; Abdul-Razzak and Ghan 2000; Li et al. 2025), significant gaps remain in understanding the mechanisms of how aerosols interact with deep convective clouds, particularly due to their dynamic and thermodynamic complexity across warm, mixed, and ice-phase processes (Fan et al. 2016). Studies have shown that incorporating aerosol effects on deep convection can improve model performance, yet these effects are still underrepresented in global atmospheric models (Song and Zhang 2011; Tao et al. 2012; Fan et al. 2016). IPCC (2023) has highlighted the low confidence in simulating aerosol microphysical effects on precipitation, emphasizing the need for improved representations of aerosol-deep convection interactions to better predict precipitation, latent heating, and radiative impacts. Accurately representing these processes is essential for understanding regional and global climatic patterns, as well as for improving numerical weather prediction models (Davies et al. 2013; Bodas-Salcedo et al. 2019; Morrison et al. 2020). Despite advancements in modeling, challenges remain in linking convective cell properties to their controlling environmental parameters, including aerosols, which are a significant source of uncertainty in predicting anthropogenic influences on precipitation and atmospheric processes.

Previous studies have identified key synoptic-scale ingredients influencing deep convection and its upscale growth into mesoscale convective systems (MCSs) such as low-level convergence, upper-level divergence, baroclinic troughs and fronts, among others (Laing and Fritsch 2000; Rasmussen and Houze 2011; Song et al. 2019; Feng et al. 2022). However, how convective initiation (i.e., when radar first detects convection) couples with local environments and internal cloud processes remains poorly understood (Feng et al. 2022). Studies have also highlighted the complex interactions between aerosols and the lifecycle of convective systems, emphasizing both microphysical and thermodynamic effects (Yin et al. 2024; Chakraborty et al. 2016; Zang et al. 2023). Rosenfeld et al. (2008) hypothesized that aerosols have a stronger impact during the dissipating phase of deep convection. Based on cloud-resolving simulations, Fan et al. (2013) showed that thermodynamic effects can influence the developing stage, whereas aerosol driven microphysical effects dominate changes during the mature and dissipating stages. Using model simulations, Fan et al. (2007, 2009) showed that increasing aerosol concentrations up to an optimal threshold can enhance MCSs under weak shear and high humidity but inhibit them under strong shear and low humidity. Hu et al. (2019) found that increased CCN concentrations invigorate convection and enhance lightning activity up to an optimal level beyond which lightning rates decrease. Chakraborty et al. (2016) reported that a one standard deviation increase in aerosol optical depth (*AOD*) can

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

extend MCS lifetimes by 3 to 24 hours based on satellite observations. Hagos et al. (2013) linked convective system lifespans to boundary layer factors (e.g., cold pools, surface fluxes) and tropospheric processes (e.g., wind shear, moistening). Most of these studies focused on convective clouds within complex MCSs. Recent studies have also questioned the size and direction of aerosol effects on deep convection. Pollution tends to weaken rather than strengthen entraining parcels (Peters et al., 2023a). Theoretical calculations indicate that CCN-related changes to updraft speed are generally small and can even become negative (Igel and Van den Heever, 2021), consistent with the minor aerosol sensitivity shown in Grabowski and Morrison (2020). Observational analyses show a similar tendency; Veals et al. (2022) reported that higher aerosol concentrations were mostly associated with reduced deep convective cloud depth.

Building on this perspective, Varble (2018) emphasized that meteorological factors often have a stronger influence on convective dynamics than do aerosol effects, underscoring the need for targeted observations of isolated deep convective cells to better quantify aerosol impacts and disentangle the role that aerosols have from that of the "meteorology" (i.e., the kinematic and thermodynamic environment) in the structure and evolution of deep moist convection.

Observational studies, which are critical for benchmarking and evaluating model predictions, have been hindered by a lack of high-quality datasets that capture the complete lifecycle and variability of convective cells under diverse meteorological and aerosol conditions (Tuftedal et al. 2024). Coastal convective systems offer a unique perspective on these interactions, as their development is often driven by localized forcing mechanisms such as land-sea breezes rather than large-scale, often stronger forcings such as fronts (Bergemann and Jakob 2016; Birch et al. 2016; Tuftedal et al. 2024). While land-sea breeze environments do contain sharp thermodynamic and aerosol gradients between marine and continental air masses, these contrasts naturally produce a broad and repeatable range of environmental conditions over relatively small spatial and temporal scales. This variability frequently leads to the formation of numerous isolated convective cells whose evolution is strongly modulated by the surrounding thermodynamic and aerosol environment. Taken together, these features provide an effective testbed for assessing how convective cell properties respond across diverse meteorological and aerosol regimes. The Tracking Aerosol Convection Interactions Experiment (TRACER; Jensen et al. 2025) was conducted in the vicinity of Houston during

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

2021–2022, at times overlapping with the NSF (National Science Foundation)-funded Experiment of Sea-Breeze Convection, Aerosols, Precipitation, and Environment (ESCAPE; Kollias et al. 2025) to enable synergistic data collection and analysis. The Houston area presents a distinctive environment with frequent occurrence of isolated convective cells influenced by a wide range of aerosol conditions, from heavily polluted urban and industrial zones to nearby areas (both maritime and continental) with significantly lower aerosol concentrations. This campaign aimed to explore the interactions between aerosols and convective processes in a region characterized by a mix of urban pollution, natural aerosols, and diverse meteorological conditions. The evolution of numerous convective cells under various environmental scenarios was tracked using the second-generation C-Band Scanning ARM Precipitation Radar (CSAPR2) and other collocated instruments. This unique dataset allows for an in-depth investigation of convective cell properties, their lifetimes, and their dependence on environmental conditions. In this study, observational data collected during the TRACER Intensive Operational Period (IOP) from June to September 2022 are utilized. This study aims to understand the dependence of convective cell properties on environmental parameters, focusing on three specific research questions:

1. What are the key environmental parameters that control convective cell lifetime?
2. How do the convective cell properties change with key environmental parameters?
3. What is the relationship between the convective cell lifetime and surface aerosol conditions?

The structure of the paper is organized as follows: Section 2 describes the data and methodology employed in the study; Section 3 presents the detailed analysis of convective cell properties and their dependence on environmental parameters; conclusions are summarized in Section 4.

2. Data and Methods

a. Data

1) CSAPR2 AND KHGX RADAR DATA

The CSAPR2 is a C-band dual-polarization Doppler radar designed for high-resolution observations of reflectivity (Z), Doppler velocity, spectral width, and polarimetric variables. It

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

has a range of 110 km and a range resolution of 100 m (Lamer et al. 2023). The CSAPR2 was deployed at 29° 31′ 55.2″ N, 95° 17′ 2.4″ W during the TRACER IOP. The CSAPR2 optimized convective cell tracking dataset (Oue et al. 2023) offers a comprehensive collection of high-resolution radar observations of convective cells. The CSAPR2 employed the Multisensor Agile Adaptive Sampling (MAAS) framework (Kollias et al. 2020) for real-time storm tracking, enabling high-frequency scanning. Radar scans, including Plan Position Indicator (PPI) and Range Height Indicator (RHI) scans, were performed at intervals of less than two minutes, allowing the scans to capture rapid microphysical and dynamical changes in evolving convective cells. This very high temporal and spatial resolution makes the dataset particularly valuable for studying the microphysics and dynamics of convective systems. The MAAS framework used by CSAPR2 relies on the nearest Next-Generation Weather Radar (NEXRAD) KHGX, Geostationary Operational Environmental Satellites 16 (GOES-16) and Geostationary Lightning Mapper (GLM; Schmit et al. 2017) to select target cells (Lamer et al. 2023). NEXRAD KHGX is an S-band dual-polarization operational weather radar operated by the National Weather Service and located at 29°28′19″ N, 95°04′45″ W. The KHGX radar performs volume coverage patterns (VCPs), which involve 360° azimuthal scans typically at progressively increasing elevation angles (Brown et al. 2005; Kingfield and French 2022). Its maximum unambiguous range varies from 300 to 450 km, depending on the pulse repetition frequency and processing mode, with a range gate spacing of 250 m (Lamer et al. 2023). Although MAAS enables adaptive tracking, its reliance on KHGX guidance can occasionally lead CSAPR2 to miss parts of a cell's lifecycle, particularly the early development or dissipation stages, as it may shift focus to other cells. To address this limitation, NEXRAD KHGX observations were utilized to derive the observed radar-based complete lifetime of convective cells. Whereas NEXRAD radars can provide data useful for studies such as this (e.g., Pehl et al. 2025), the customizable scanning strategies and higher resolution provided by the CSAPR2 allows us to study the characteristics of convective cells with considerably greater detail. Radar data processing and analysis were conducted using the Python ARM Radar Toolkit (Py-ART; Helmus and Collis 2016), an open-source library widely adopted for the analysis and visualization of weather radar data.

2) Meteorological Conditions

In this study, hourly model-level data from the European Centre for Medium-Range Weather Forecasts (ECMWF) atmospheric reanalysis v5 (ERA5, Hersbach et al. 2023) at a

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

grid spacing of 0.25° were used to estimate meteorological parameters associated with convective cells. This approach addresses the limitations posed by the small number (~4–5/day and 5+ on enhanced IOP days) of balloon-borne soundings available at the TRACER sites. Huang et al. (2025) evaluated ERA5 profiles using all available TRACER IOP soundings, showing mean biases in temperature, relative humidity, and wind components (u and v) of approximately −0.10 °C, 1.57%, 0.01 m s^{-1}, and 0.02 m s^{-1}, respectively. These biases fall within the uncertainty ranges of the radiosonde sensors (Holdridge 2020). For each convective cell, ERA5 temperature, relative humidity, and wind profiles were extracted from the nearest ERA5 grid point to the cell's first detection location using CSAPR2. The profile time corresponds to the hour immediately preceding cell initiation, where cell initiation refers to the first time KHGX exhibits composite reflectivity exceeding 20 dBZ at any altitude. No horizontal interpolation was applied; instead, the nearest ERA5 grid point was used for all variables. Using the ERA5 vertical profiles, ~44 environmental variables (listed in Table A1, one is using MERRA-2) were calculated for each identified convective cell to characterize the surrounding environmental conditions influencing cell initiation and evolution. These variables are derived using the thundeR package (Taszarek et al. 2023). Parcel parameters involving entrainment were calculated following the methodology employed by Peters et al. (2022, 2023b). Variables include key thermodynamic and dynamic parameters such as convective available potential energy (CAPE), convective inhibition (CIN), and vertical wind shear across various levels (0–500m, 0–1 km, 0–3 km, 0–6 km, and 0–8 km). Precipitable water vapor (*PWV*), representing the total column-integrated water vapor, and lapse rates for different atmospheric layers were also calculated to assess vertical temperature gradients and atmospheric stability. These parameters are widely recognized as crucial to the initiation, evolution, and organization of deep convection (Weisman and Klemp 1982; Emanuel 1994; Stevens 2005; Giaiotti et al. 2007). To further assess the suitability of ERA5 for representing the environmental conditions relevant to this study, several key predictor variables including *PWV*, *BS_08km* and *LR_26km* were compared between ERA5 and those derived from the TRACER radiosondes (Fig. S1). These comparisons exhibit biases consistent with those documented by Taszarek et al. (2021), supporting the reliability of ERA5-derived environmental parameters used in the analysis. This comprehensive set of derived variables provides valuable insights into the environmental controls on convective processes and offers a robust basis for investigating the factors influencing convective cell lifetimes and evolution.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

3) Aerosol Conditions

Aerosol properties associated with each convective cell were determined using fine aerosol mass concentration (*Mass_f,* diameter <1 μm) values derived from the Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2) (Gelaro et al. 2017). The MERRA-2 Instantaneous Model-Level Assimilation Aerosol Mixing Ratio (M2I3NVAER) dataset, produced by the National Aeronautics and Space Administration (NASA) Global Modeling and Assimilation Office (GMAO) using the Goddard Earth Observing System (GEOS) version 5.12.4 atmospheric model, provides three-hourly, vertically resolved aerosol mass mixing ratios for individual aerosol species including dust, sea salt, sulfate, black carbon, and organic carbon across 72 model levels. The three-hourly dataset was selected rather than the hourly MERRA-2 aerosol product because only the former includes bin-resolved mass mixing ratios, which allow characterization of fine-mode aerosol components, including the submicron sea-salt bins that are not present in the hourly product. Aerosol fields were first bilinearly interpolated to each cell's latitude-longitude position, followed by linear temporal interpolation to construct an hourly time series consistent with the temporal resolution of the meteorological datasets. To verify that the temporal interpolation preserved the underlying aerosol variability, interpolated hourly surface mass concentrations were compared with the corresponding hourly MERRA-2 aerosol fields for species common to both datasets. They are extremely consistent, with correlation coefficients of r = ~0.99, and mean biases of 0.001 μg m^{-3} for black carbon, 0.02 μg m^{-3} for organic carbon, and 0.02 μg m^{-3} for sulfate. The high correlation and minimal biases indicate that the temporal interpolation introduces negligible distortion and maintains the temporal structure of the aerosol fields. Fine-mode aerosol mass for each convective cell was defined as the sum of sulfate, organic carbon, black carbon, and fine-mode sea-salt bins (dry diameter < 1 μm). Dust aerosol was excluded because of its predominantly hydrophobic nature and limited CCN activity. For each convective event, hourly fine-mode mass mixing ratios at the interpolated grid point were extracted at the lowest model level (~surface) immediately preceding convective initiation and converted to surface *Mass_f* using the corresponding air density. The use of aerosol mass concentrations to represent ambient aerosol conditions follows approaches adopted in recent studies (e.g., Pan et al. 2021; Block et al. 2024; Huang et al. 2025).

In addition to *Mass_f,* detailed aerosol measurements were obtained from ground-based instruments included in the Aerosol Observing System (AOS; Uin et al. 2019) deployed at the

Accepted for publication in *Journal of the Atmospheric Sciences.* DOI 10.1175/JAS-D-25-0130.1.

ARM mobile facility (AMF; Miller et al. 2016) main site (AMF1). Table 1 lists all aerosol instruments used in this study. The Aerosol Chemical Speciation Monitor (ACSM) provided aerosol composition data including organics, sulfate, nitrate, and ammonium prior to the initiation of each convective (Zawadowicz et al. 2021). Condensation Particle Counter (CPC) instruments measured total particle number concentrations for both ultrafine-type and fine-type (Koontz et al. 2021a, b). Scanning Mobility Particle Sizer (SMPS) and Ultra-High-Sensitivity Aerosol Spectrometer (UHSAS) together captured aerosol size distributions across a wide range of 10–1000nm (Shilling and Levin 2021; Uin et al. 2021). It should be noted that, as discussed by Uin et al. (2024), aerosol size distributions derived using optical techniques such as UHSAS can contain artifacts inherent to optical measurements; therefore, care must be taken when interpreting these data. The single-particle soot photometer device (SP2) quantified refractory black carbon (BC) mass using laser-induced incandescence (Jackson et al. 2021). This integrated dataset offers a comprehensive representation of aerosol composition, number, and size distributions prior to each convective initiation. Table 2 summarizes the aerosol variables included in this analysis.

The representativeness of the MERRA-2 aerosol fields over the TRACER domain was assessed through comparison with in-situ observations. Bin-resolved submicron aerosol mass concentrations from M2I3NVAER were compared with the 6-hour running-mean UHSAS number concentrations in the 100–1000 nm range (Figs. 1a-d). Despite differences between mass- and number-based quantities, both datasets showed co-varying temporal structure, with aligned peaks and troughs and a correlation coefficient of 0.75, demonstrating that MERRA-2 captures submicron aerosol variability consistent with UHSAS measurements. Additional evaluation was performed using co-located aerosol composition measurements from the AMF1 site (Fig. S2 a, b). Comparisons of MERRA-2 organic carbon (hydrophilic and hydrophobic) mass concentrations with ACSM-derived total organics and MERRA-2 black carbon (hydrophilic and hydrophobic) mass concentrations with SP2-derived refractory black carbon showed consistent temporal evolution, with 6-hour running-mean correlations of r= 0.77 for organics and r= 0.63 for black carbon. Although the MERRA-2 mass concentrations tend to be higher, this difference may reflect fundamental contrasts in representation between the MERRA-2 reanalysis and in situ measurements, including size cutoffs, sampling flow characteristics, and the treatment of refractory and non-refractory aerosol components in ACSM and SP2 observations. Taken together, the UHSAS, ACSM, and SP2 comparisons

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

indicate that the M2I3NVAER dataset realistically captures the temporal variability in fine aerosol loading during the TRACER period.

Further, to assess the consistency between SMPS and UHSAS measurements, aerosol number concentrations were calculated within the common size range of 60–500 nm by integrating each instrument's size distributions over this interval. This range was chosen to ensure overlap in detection capabilities and minimize bias from differing instrument cutoffs. The resulting number concentrations showed strong agreement (Fig. 1e), with a correlation coefficient of 0.94, a mean bias of –410.4 cm^{-3}, and an RMSE of 722.6 cm^{-3} across 2424 samples. The observed bias may be attributed to various factors such as differences in the working principles of the two instruments, variations in sampling frequency, particle refractive index, distinctions between mobility and optical diameters, and limitations in counting efficiency within the UHSAS optical system particularly affecting the detection of particles in the 60–100 nm size range (Ortega et al., 2019). Nevertheless, the high correlation coefficient (0.94) indicates a consistent relationship between the measurements, thereby ensuring the robustness and credibility of the data for subsequent analyses.

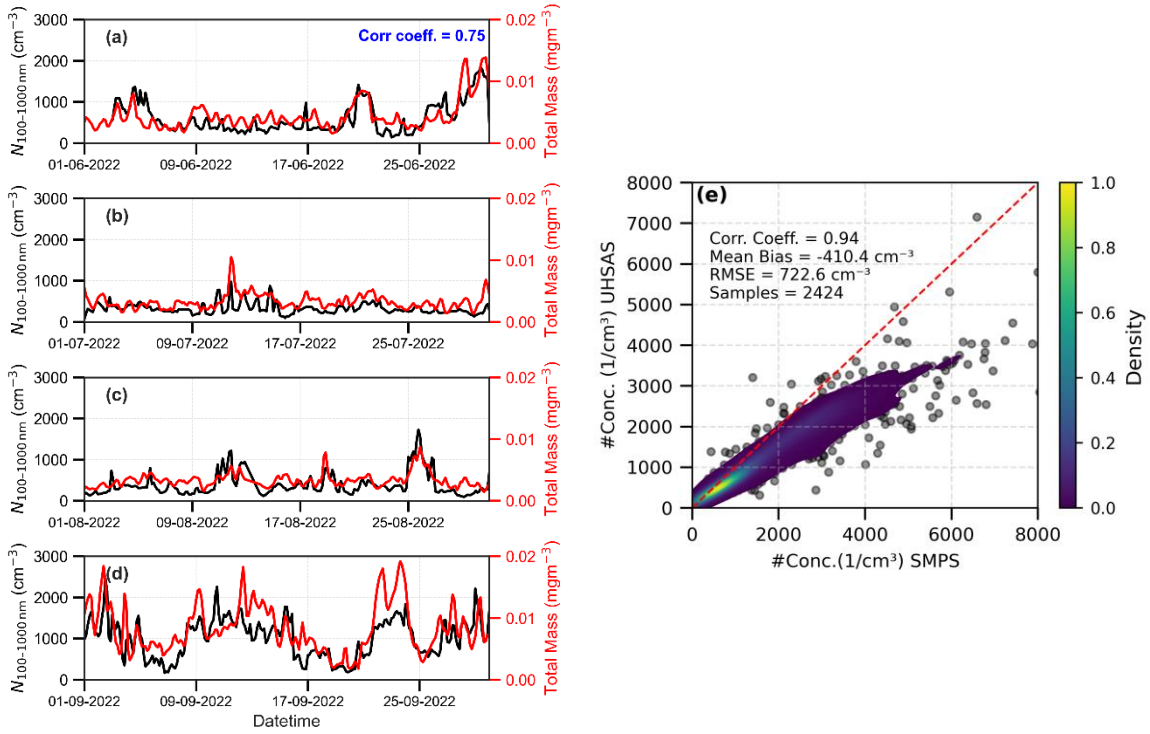


Fig. 1. (a–d) Monthly time series of UHSAS aerosol number concentration for diameter (D) in 100–1000 nm range (cm^{-3}) and MERRA-2 aerosol mass concentration from the M2I3NVAER

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

dataset for D in 100–1000 nm range (mg m^{-3}) from June to September 2022. Both time series are smoothed using a 6-hour running mean. (e) Comparison of aerosol number concentrations measured by SMPS and UHSAS in the 60–500 nm diameter range. The background shading represents the normalized kernel density estimate of the data distribution. The dashed red line indicates the 1:1 reference. Summary statistics including correlation coefficient, mean bias, and RMSE are annotated.

Table.1 Instruments from Aerosol Observing System (AOS) at AMF1 site.

Instrument	Size-Range	Accuracy	Reference
ACSM (Aerosol Chemical Speciation Monitor)	-	±30%	(Ng et al. 2011; Watson 2017)
CPCF (Condensation Particle Counter Fine)	10 - 3000 nm	2.5% (clean) 0.3% (polluted)	(Singh and Kuang 2024a)
CPCU (Condensation Particle Counter Ultrafine)	3 - 3000nm	2.5% (clean) 0.3% (polluted)	(Singh and Kuang 2024a)
SMPS (Scanning mobility Particle Sizer)	10 - 500nm	±2% sizing	(Singh and Kuang 2024b)
SP2BC (Single Particle Soot Photometer Black Carbon)	-	<10%	(Sedlacek 2017; May et al. 2014)
UHSAS (Ultra-High-Sensitivity Aerosol Spectrometer)	60 to 1000 nm	±5% sizing	(Uin 2024)

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

Table 2. Observed aerosol parameters associated with convective cells (for cells within 50-km radius around AMF1 site).

Variable Name	Description
$N_{3\text{-}3000}$	Aerosol number concentration for 3–3000 nm diameter range using CPCU
$N_{10\text{-}3000}$	Aerosol number concentration for 10–3000 nm diameter range using CPCF
$N_{10\text{-}100}$	Aerosol number concentration for 10–100 nm mobility diameter range using SMPS
$N_{100\text{-}1000}$	Aerosol number concentration for 100–1000 nm optical diameter range using UHSAS
$N_{500\text{-}1000}$	Aerosol number concentration for 500–1000 nm optical diameter range using UHSAS

b. Methods

1) DEEP CONVECTIVE CELL CLASSIFICATION

Using the CSAPR2 cell tracking dataset, a total of 413 deep convective cells were identified for analysis. For this study the deep convective cells are defined as those for which the 20-dBZ echo top height (ETH) exceeds the local freezing level at least once during the CSAPR2 scan. The freezing level height was estimated using the ERA5 temperature and geopotential height profiles. Across the analysis period, the freezing level varied between ~4.36 km and ~5.40 km with a mean of ~4.84 km. The restriction to the high ETHs effectively isolates actively precipitating deep convective structures while excluding shallower or predominantly stratiform clouds. The spatial distribution of these convective cells, based on their first tracked locations identified using CSAPR2, is shown in Fig. 2a, emphasizing their widespread occurrence that enables analysis across diverse environmental conditions.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

To capture the complete lifetime of each convective cell, the first CSAPR2-identified location was used as the anchor point, and the cell track was extended both backward and forward in time using the KHGX composite reflectivity. Because CSAPR2 frequently sampled only portions of a cell's lifetime, the corresponding KHGX volume scans were used to reconstruct the earlier and later stages of the same cell. During the study period, the KHGX volume-scan interval often varied between ~5–6 minutes; therefore, required time interpolation of composite reflectivity between adjacent scans, applied when required, to generate a 2–3-minute composite reflectivity time step suitable for continuous tracking. Manual verification was conducted to ensure that the same cell was consistently followed through time, particularly in situations where multiple isolated cells were present in proximity. In this study, cell lifetime is defined as the duration during which a convective cell maintains a composite reflectivity of at least 20 dBZ at any altitude. No minimum lifetime threshold was imposed. This 20-dBZ threshold is similar to those commonly used to determine storm ETH, e.g., 0 and 10 dBZ (Stein et al. 2020), 18 dBZ (Lakshmanan et al. 2013; Scovell and al-Sakka 2016; Miltenberger et al. 2018; Stein et al. 2020), and 20 dBZ (Feng et al. 2022). Convective cell lifetimes were comprehensively extracted using both the CSAPR2 and KHGX radar datasets. In addition, the maximum heights and widths of convective cells were determined using a series of CSAPR2 RHI scans and a reflectivity threshold of 20 dBZ. Cell height was defined as the highest altitude (AGL) of $Z \geq 20$ dBZ during the cell's tracked lifetime. To quantify horizontal structure, the maximum width in the 2–4 km and 4–6 km AGL layers was determined by identifying the largest horizontal extent of contiguous regions with $Z \geq 20$ dBZ associated with the target cell, based on all RHI scans intersecting the cell throughout its lifetime. This approach ensures that the measured width represents the broadest cross-sectional span of the cell within each layer across multiple RHI azimuths. To assess the sensitivity of width estimates to reflectivity thresholds, calculations were repeated using Z thresholds of 10, 30, 40, and 50 dBZ, which showed consistent variation patterns across all thresholds. Widths below 2 km were excluded due to limited low-level radar coverage for distant cells, and layers above 6 km were omitted because not all cells reach those heights. The 2–6 (2–4 and 4–6) km range was therefore selected to provide consistent and reliable width estimates across all cells.

To characterize the height distribution of the convective cells, echo-top heights were examined using both 5 dBZ and 20 dBZ thresholds. Based on the full dataset, 95.5% of cells exceeded 6 km using a 5 dBZ threshold, and 88.8% exceeded 6 km using the 20 dBZ threshold, indicating that the population is strongly dominated by cells reaching well into higher levels.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

Additional sensitivity tests were conducted by applying progressively stricter ETH thresholds of > 6 km, > 7 km, and > 8 km using the 20 dBZ height. The key relationships (aerosol-effects) presented in this study remained consistent across all height-based subsamples, indicating that the main conclusions are insensitive to the cell height cutoff, likely because the vast majority of cells already exceed a 5 dBZ height of 6 km. Two caveats apply when interpreting ETH-based classifications. First, the 20-dBZ ETH may often lower than the actual cloud top and can be influenced by vertical reflectivity gradients, so it may underestimate the full cloud depth even though it is used here for consistency with the lifetime definition. Second, satellite-retrieved ETH was not incorporated, and the analysis therefore relies solely on radar-based metrics. The distribution of cell lifetimes (Fig. 2b) indicates that most (~70%) convective cells fall within the lifetime of 31–75-min range. Fewer cells exhibit lifetimes shorter than 30 min (~11%) and longer than 75 min (~18%), highlighting a skewed distribution where the majority of cells tend to have lifetimes shorter than 75 min. The convective cells were classified into three categories based on their lifetimes: short-lived (0–40 min), intermediate-lived (41–80 min), and long-lived (80+ min), with 119 (~29%), 236 (57%), and 58 (~14%) cells in each respective category. This three-tiered classification captures the full range of observed lifetimes and allows evaluation of lifetime sensitivity to environmental conditions while distinguishing short-, intermediate-, and long-lived cells. The 40-min threshold separates brief events from more sustained cells, while the 80-min cutoff isolates the longest-lived ~14%, which are of particular interest due to their extended duration and potential for greater vertical development. Although the thresholds are discrete and somewhat arbitrary, they are informed by the observed lifetime distribution and enable physically meaningful comparisons across regimes. Environmental parameters, including meteorological and aerosol conditions, were analyzed for the hour preceding the initiation of each cell. Sensitivity tests using a two-hours window yielded consistent results. To ensure representativeness of surface in-situ aerosol observations, only cells within a 50-km radius of the AMF1 site were considered for the in-situ aerosol analysis. This 50-km threshold was determined based on a sensitivity analysis showing the results (i.e. variation of $N_{10\text{-}100}$ and $N_{500\text{-}1000}$ with cell lifetime discussed in Section 3c) associated with the SMPS (10–100nm) measurements remain consistent for cell within 50 km, while the results associated with UHSAS (500–1000nm) measurements show consistency up to cells within 80 km. Although Fig. 1e shows good agreement between SMPS and UHSAS, it reflects only their overlapping size range of 60–500 nm, whereas the sensitivity analysis considers the different size-range used from each instrument. However, the primary rationale for adopting

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

the 50-km threshold was to avoid including convective cells that originated over the Gulf of America, which is also consistent with the methodological approach of Wang et al. (2024).

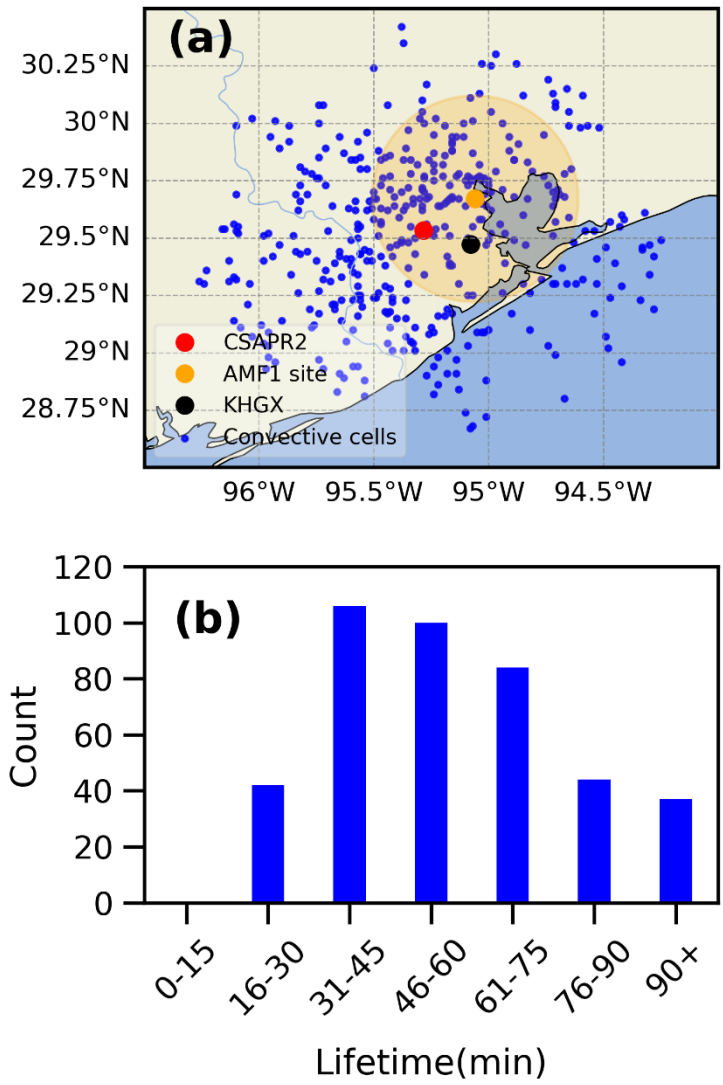


Fig. 2. (a) Geographical distribution of convective cells identified during the TRACER IOP. The markers represent the CSAPR2 radar (red), the AMF1 site (orange), and the KHGX radar (black), and the blue dots represent the convective cell locations where the CSAPR2 tracked it for the first time. The shaded region (light orange) represents a 50 km radius around the AMF1 site, highlighting the cells considered in the surface aerosol analysis. (b) Bar plot of convective cell lifetimes for all the cells (~413), categorized into duration bins. The x-axis indicates lifetime intervals (min), and the y-axis shows the number of convective cells within each bin.

Furthermore, convective cells were classified based on prevailing onshore or offshore winds. This classification was conducted using the mean *u* and *v* wind components between

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

1000 and 850 hPa, derived from ERA5 reanalysis data. Cases with mean southeasterly or southwesterly winds were categorized as onshore events, while those with northeasterly or northwesterly winds were classified as offshore events (Fig. 3). Based on this criterion, 329 cells occurred under onshore wind conditions and 84 under offshore wind conditions. This classification, based on onshore and offshore wind conditions distinguishes between the relatively clean southerly winds originating from the Gulf and the potentially more polluted northerly winds flowing from land, was adopted to separate marine and continental flow regimes relevant to aerosol conditions without attempting a detailed characterization of coastal-flow dynamics.

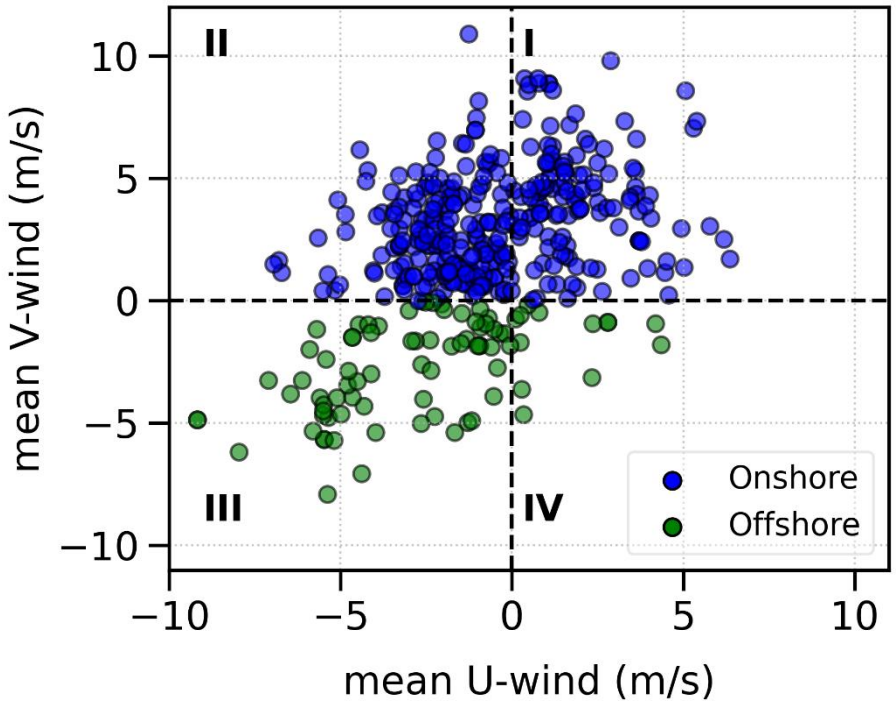


Fig. 3. Scatterplot of the mean *u*- and *v*-wind components (m s^{-1}) between 1000 and 850 hPa from ERA5 reanalysis. Data points in quadrants I and II, corresponding to southwesterly and southeasterly flows (onshore wind conditions) respectively, are shown in blue, and points in quadrants III and IV, corresponding to northeasterly and northwesterly flows (offshore wind conditions) respectively, are shown in green.

2) RANDOM FOREST

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

The Random Forest (RF) algorithm is a group of decision trees, where each tree is built using a random selection of training data and predictors. These trees collectively contribute to the final prediction (Breiman 2001). A notable advantage of the RF algorithm is its ability to quantify predictor importance, which is particularly valuable for identifying the environmental factors most strongly influencing convective cell properties. Due to its ability in managing noisy data and outliers, it is particularly well-suited for scenarios involving complex interactions and non-linear relationships (Biau and Scornet 2016; Genuer et al. 2017). RF has been widely applied in atmospheric sciences, including air quality forecasting (Lei et al. 2022; Vovk et al. 2024), cloud liquid water retrievals (Schulte et al. 2024) cloud microphysical analysis (Huang et al. 2025), etc.

In this study the RF algorithm implemented in the Scikit-learn Python package (Pedregosa et al. 2011) is used, incorporating Monte Carlo cross-validation with 100 iterations to assess the model uncertainty. Environmental parameters including meteorological and aerosol conditions associated with each convective cell were used as predictors, while convective cell properties served as target variables, enabling effective model training and prediction.

3) BOOTSTRAPPING

Bootstrapping is a non-parametric resampling approach that involves repeatedly sampling the original dataset with replacement to create numerous simulated datasets of the same size (Efron 1979; Efron and Tibshirani 1994). For each resample, the mean is calculated, capturing inherent variability in the data. After a large number of iterations, the bootstrapped means are aggregated to compute the overall average and confidence intervals. This method is particularly advantageous when the underlying distribution of the data is unknown or the dataset is small, as it does not rely on strict parametric assumptions. Bootstrapping is widely applied in atmospheric science for diverse purposes, including confidence band construction, radar profile analysis, statistical comparisons, model enhancement, and uncertainty quantification in both observations and reanalysis (Mudelsee 2011; Marchand et al. 2006; Xu 2006; Rao 2000; Cuesta-Valero et al. 2022; McFarquhar 2004; McFarquhar and Heymsfield 1997; Liu et al. 2024).

In this study, the bootstrapping method is used to estimate the mean aerosol number size distribution and its associated confidence intervals across different convective cell categories, including classifications by lifetime and on-offshore wind direction. Because some categories contained relatively small or uneven sample sizes, bootstrapping was applied to address this

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

sample imbalance and improve the statistical robustness of the analysis. Additionally, this method was employed to assess the statistical significance of difference in aerosol size distribution between convective cell categories, allowing for a more reliable comparison between them.

4) K-MEANS CLUSTERING

K-Means clustering (Hartigan and Wong 1979) is an unsupervised machine learning algorithm that partitions a multidimensional data space into a predefined number of clusters. Each cluster is represented by a centroid, and data points are assigned to the nearest centroid based on a distance metric, typically Euclidian distance. This clustering method has been extensively applied in atmospheric sciences for classifying atmospheric sounding profiles (Sarmadi et al. 2017; Lang et al. 2018; Truong et al. 2020). In this study, K-means clustering is applied to atmospheric soundings derived from ERA5, using temperature, specific humidity, and *u*- and *v*-wind components at multiple pressure levels (1000, 925, 850, 800, 700, 650, 600, 550, 500, 450, 400, 350, 300, 250, and 200 hPa) as input variables. Each variable at each pressure level is standardized prior to clustering, so all variables contribute equally to the Euclidean distance metric used by the K-means algorithm; no additional weighting is applied. The optimal number of clusters was determined by analyzing the percent change in total spatial variance (TSV) as a function of the cluster number, selecting the appropriate number of clusters just before a sharp increase in the percent change in TSV.

3. Results

a. Key environmental parameters that control cell lifetime

To identify the key environmental parameters controlling convective cell lifetime, an RF model was trained using the meteorological variables listed in Table A1. Prior to model training, highly collinear predictors were removed through correlation-based screening using multiple thresholds ($|r| \geq 0.90$, 0.85 and 0.80; see Fig. A1 for the correlation matrix of some of the key variables). Depending on the correlation threshold applied, the number of variables retained for model training is lower than the total number listed in Table A1. Monte Carlo cross-validation was performed to verify the stability of variable importance rankings across data subsets. Since these rankings reflect the combined outcome of multiple sensitivity and collinearity-based filtering, individual importance scores are not reported. The relative importance of environmental parameters in predicting convective cell lifetime indicates that

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

precipitable water vapor (*PWV*) was the most influential among the ~45 variables assessed over 100 iterations, followed by the 2–6 km temperature lapse rate (*LR_26km*), 0–8 km bulk shear (*BS_08km*) and fine aerosol mass concentration (*Mass_f*). Sensitivity experiments, systematically removing top-ranked variables, consistently show the remaining original top variables retained their high importance, further reinforcing the robustness of these relationships. As an additional step to ensure reliability of importance estimates, we further implemented permutation importance, which is model-agnostic and better suited for handling feature redundancy and found consistent rankings, with *PWV* and *BS_08km* remaining top predictors. Additionally, partial dependence plots (PDPs, Fig. S3) were generated to visualize the marginal effects of key variables. The PDPs reveal a nonlinear increase in predicted lifetime with increasing *PWV* and a gradual positive influence from *BS_08km*.

To further evaluate these findings and enhance the robustness of the model, a bootstrap resampling approach was employed to generate 1,000 samples while preserving the original variable distributions. The RF model was retrained using Monte Carlo cross-validation on the resampled datasets. The resulting variable importance rankings were similar, with *PWV* continuing to exhibit the highest mean importance followed by *LR_26km*, *Mass_f*, surface sensible heat flux (*SSHF*) *and BS_08km*. also consistently ranked among the top ten predictors. Expanding the analysis to 2,000 bootstrap iterations further reinforced the dominant influence of *PWV*, *LR_26km*, *BS_08km*, and *Mass_f*.

Overall, the application of bootstrapping led to improved model performance and strengthened the reliability of the identified predictor hierarchy. Complementary statistical analyses on the original dataset using Spearman's rank correlation (Kokoska and Zwillinger 2000; Kendall and Stuart 1973), Kendall's tau (Kendall 1938, 1945; Noether 1967; Fenwick 1994), and distance correlation (Székely et al. 2007) consistently identified *PWV*, *LR_26km*, and *BS_08km* (distance correlation) as the top three parameters. While rank-based correlation coefficients are modest (Spearman's ρ: ~0.20 for *PWV*, ~0.10 for *LR_26km*), these correlation coefficients are statistically significant. Notably, *Mass_f* does not rank within the top 10 variables in these correlation tests, possibly due to its relatively limited variability within the dataset. These bivariate correlation results provide support for the multivariate RF-derived rankings, with general consistency observed in the strength of associations.

Considering these comprehensive analyses, *PWV*, *LR_26km*, *BS_08km*, and *Mass_f* were selected as the four primary environmental parameters governing cell lifetime for further

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

analyses. *PWV* consistently demonstrates the highest importance across all above tests, emphasizing the critical role of atmospheric moisture in sustaining convection. *PWV* is also strongly correlated with mid-level relative humidity (*RH*), particularly mean *RH* at 2–5 km (*RH_25km*, r = 0.84) and *RH* at 3–6 km (*RH_36km*, r = 0.87). High *PWV* values imply abundant moisture availability for condensation and latent heat release, fueling updrafts and supporting the longevity of convective cells. Another important variable is the temperature lapse rate for 2–6 km levels (*LR_26km*), a key metric of atmospheric stability. A higher lapse rate (i.e., greater temperature decrease with height) indicates increased atmospheric instability, which enhances buoyancy and accelerates convective development. The importance of *BS_08km* highlights the role of vertical wind shear in storm organization. Higher wind shear can promote the separation of updrafts and downdrafts, reducing precipitation loading within the updraft and delaying cell dissipation. Another noteworthy parameter is *Mass_f*, suggesting a potential role of aerosols in influencing convective processes.

While this analysis provides the importance of individual variables in predicting convective cell lifetime, the independent and combined effects of moisture, instability, dynamics, and *Mass_f* are investigated in subsequent sections to further explore how interactions among them to modulate convective cell lifetime.

b. Cell properties as a function of key environmental parameters

After identifying the most important parameters affecting convective cell lifetime, the independent effects of the top four variables *PWV*, *LR_26km*, *BS_08km*, and *Mass_f* on cell lifetime were further examined (Figs. 4a–d). Stratifying the dataset into percentile-based bins (low: 0–33rd, moderate: 33–66th, high: 66–100th) for each variable, the variation in mean cell lifetime across these groups was analyzed. The results indicate that cell lifetimes are generally longer in high *PWV* conditions (Fig. 4a), with notably longer median lifetimes (~62.5 minutes) in the high *PWV* group compared to the low group (~49.0 minutes) and this difference is statistically significant at $p < 0.05$ by the Mann-Whitney U test (Mann and Whitney 1947; Wilcoxon 1945; McKnight and Najab 2010). This nonparametric approach was selected due to the skewed distribution of lifetime measurements. In contrast, *LR_26km* does not exhibit a clear relationship with lifetime, with the difference between median lifetime (~55 minutes) for the low group and (~52.5 minutes) high group (Fig. 4b) not statistically significant. For *BS_08km*, median lifetimes were ~52, ~53, and ~54 minutes for the low, moderate, and high groups, respectively (Fig. 4c), with very minor and not statistically significant differences across

File generated with AMS Word template 2.0

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

groups, indicating that *BS_08km* did not exhibit a strong influence alone on cell lifetime. In the case of *Mass_f* (Fig. 4d), there is no consistent variation in convective cell lifetimes across the *Mass_f* percentiles having median lifetime of ~53.0 minutes in low, ~56.0 minutes in high and ~52.0 minutes in moderate, with the differences again minor, not statistically significant, and reflecting a non-uniform independent influence on convective cell lifetime in this dataset.

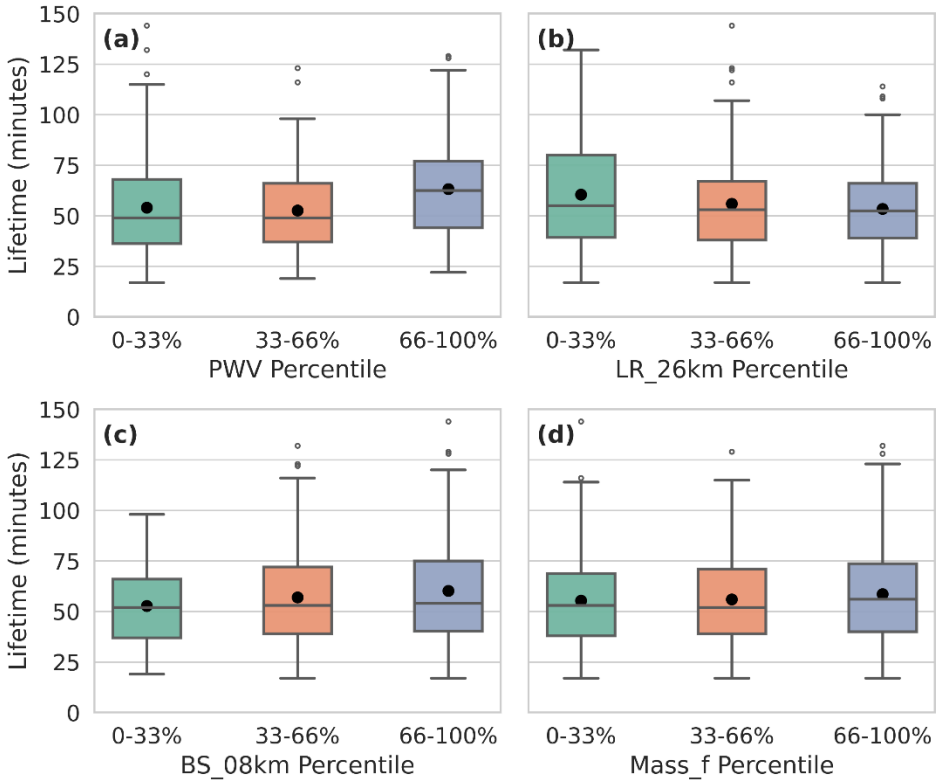


Fig. 4. Boxplots of convective cell lifetime (in min) as a function of different variables, stratified by percentiles (0–33rd, 33–66th, 66–100th) of (a) *PWV*, (b) *LR_26km*, (c) *BS_08km* and (d) *Mass_f*. Box represents the interquartile range (IQR; 25th–75th percentile), the horizontal line inside the box indicates the median, the whiskers extend to 1.5×IQR, and the black dots denote the mean values. Each group contains approximately 135–138 samples.

The combined effects of these parameters are further examined. Figure 5a shows convective cell lifetimes separated by *LR_26km* percentiles and further categorized by *PWV* percentiles. In the low *LR_26km* group, the median cell lifetimes increase steadily with rising *PWV*, and the difference in cell lifetime between low (~44 minutes) and high (~72 minutes) *PWV* as well as moderate (~48 minutes) and high *PWV* conditions is statistically significant at $p < 0.05$. This

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

highlights the substantial role of moisture in influencing cell lifetimes under weaker lapse rate conditions and could be associated with enhanced latent heat release during condensation, leading to deeper convection as moist parcels ascend. In the moderate *LR_26km* group, a notably longer median cell lifetime (~58 minutes) is associated with high *PWV*, but this is lower than the corresponding high *PWV* lifetime observed in the low lapse rate group. However, there is no statistically significant difference in lifetime across the *PWV* categories within this moderate lapse rate group. In the high *LR_26km* group, median cell lifetimes also show strong sensitivity to *PWV*, with a statistically significant difference between low (~49 minutes) and high (~60 minutes) *PWV* as well as moderate (~49.5 minutes) and high *PWV*.

Similarly, Fig. 5b displays the variation in convective cell lifetimes across *BS_08km* percentiles, corresponding to low, moderate, and high *PWV* categories. In low *BS_08km* environments, cell lifetimes exhibit minimal variation across the different *PWV* categories, suggesting that under weak shear (*BS_08km*), moisture alone is insufficient to substantially impact convection duration. This highlights the limited role of *PWV* in sustaining convection when *BS_08km* is weak. However, under moderate and high shear conditions, cell lifetimes (median values) vary significantly with *PWV*. Specifically, higher *PWV* is consistently associated with longer-lived cells having median cell lifetime ~72 minutes for moderate and ~62 minutes for higher *BS_08km*. There are also statistically significant ($p < 0.05$) differences in cell lifetimes between low and high *PWV* categories for both moderate and high *BS_08km*. This indicates that as *BS_08km* strengthens, the thermodynamic contribution of moisture becomes increasingly critical in sustaining convection. In high shear environments, median cell lifetimes increase systematically with *PWV* showing statistically significant difference in cell lifetime between low and high *PWV*, emphasizing the synergistic effect of strong vertical wind shear and abundant moisture. This relationship highlights the dual importance of wind shear and moisture in supporting sustained convective activity, and points to their combined influence as a key factor in promoting longer-lived convective cells.

Figure 5c illustrates the variation in convective cell lifetimes across *Mass_f* percentiles, further stratified by *PWV* categories (low, moderate, and high). This analysis examines the combined impact of aerosol loading and moisture availability on the lifetime of convective cells. In low *Mass_f* conditions, convective cell lifetimes significantly increase with *PWV* having median lifetime of ~48 minutes in low, ~52 minutes in moderate and ~61 minutes in high *PWV*, suggesting that in relatively clean environments, thermodynamic support from

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

moisture is the primary driver of convective persistence. Under moderate *Mass_f* conditions, high *PWV* is associated with significantly longer cell lifetimes with a median lifetime of ~68 minutes compared to ~49 minutes (low *PWV*) and ~50 minutes (moderate *PWV*). The moderate *Mass_f*–high *PWV* group exhibits the longest and statistically significant lifetimes among all *Mass_f*–*PWV* combinations. In high *Mass_f* conditions, the median cell lifetimes are ~52 minutes in low *PWV*, ~40 minutes in moderate *PWV* and ~61 minutes in high *PWV*, with the high–*PWV* group having the longest median lifetime. However, these differences are not statistically significant in high *Mass_f* conditions, indicating that no robust *PWV*-dependent lifetime signal is detected under high *Mass_f* conditions within the present dataset. Across low and moderate *Mass_f* percentiles, the longest cell lifetimes (median values) are consistently associated with high *PWV* groups. While *PWV* emerges as the dominant control across low and moderate *Mass_f* groups, the independent role of *Mass_f* is further examined using RF-based perturbation sensitivity analyses presented later in this section. In addition to examining how convective cell lifetimes vary across *PWV* percentiles within fixed groups of *LR_26km*, *BS_08km*, and *Mass_f*, the inverse relationship was also evaluated: how cell lifetimes vary across *LR_26km*, *BS_08km*, and *Mass_f* percentiles within fixed *PWV* groups (Figs. 5d–f). The observations indicate that under high *PWV* conditions, convective cell lifetimes are significantly longer in the low *LR_26km* group compared to the moderate and high lapse rate groups (Fig. 5d), whereas, for *BS_08km* (Fig. 5e) under high *PWV* conditions a statistically significant difference is observed between the low-moderate and low-high *BS_08km* groups with higher median lifetime observed in moderate *BS_08km* (~71 minutes) followed by high *BS_08km* (~67 minutes) and low *BS_08km* (~55 minutes). In contrast, under low and moderate *PWV* conditions, no significant differences in cell lifetimes are found across the *LR_26km*, *BS_08km*, and *Mass_f* groups, except for a notable case at low *PWV*, where a significant difference in lifetime is observed between the low and high *Mass_f* groups, with the high *Mass_f* group exhibiting a higher median lifetime (Fig. 5f).This systematic trend underscores the dominant role of atmospheric moisture in extending convective cell lifetime, with *LR_26km*, *BS_08km*, and *Mass_f* contributing to systematic variations in lifetime. It is important to note that the aerosol-related sensitivities discussed above remain consistent when the analysis is repeated using alternative cell height (Z > 20 dBZ) thresholds of > 6, > 7, and > 8 km for defining deep convective cells, suggesting that the results are not sensitive to the specific height criterion used to define deep convective cells within the given sample size. Although low-level convergence forcing the convection is not directly observable in the present

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

dataset, potential spatial and diurnal modulation of convective forcing is indirectly assessed using convective cell location and local time. The spatial distribution (Fig. S6a) of convective cell lifetime shows no systematic geographic organization across the study region, and key environmental predictors exhibit similar homogeneous spatial patterns (Fig. S6 b-f). In addition, convective cell lifetime does not display a pronounced diurnal cycle. Stratifying the analysis by time-of-day, 06–12 LT (Local Time) and 12–18 LT, yields aerosol-lifetime relationships that are qualitatively consistent with those discussed above for the complete dataset, indicating that the observed aerosol effects are unlikely to be artifacts of spatial location or diurnal timing.

To further verify these relationships and recognize potential limitations of the percentile-based approach, particularly the inability to fully control for remaining predictors and within-bin correlations, RF-based perturbation sensitivity analyses are conducted following the approach of Borque et al. (2022). The same RF configuration described in Section 3a (Monte Carlo cross validation, 80/20 train-test split, collinearity threshold $|r| \geq 0.90$) is used while the specific percentage changes reported are associated with these hyperparameter choices. Repeated testing across varying configurations yields qualitatively consistent results. Two-variable perturbations are performed by simultaneously perturbing pairs of the four important predictor variables (*PWV*, *LR_26km*, *BS_08km*, and *Mass_f*) by −50%, −30%, +30%, and +50% while holding all other inputs the same (Fig. S4), designed to evaluate the same two-variable combinations as those examined in Figs. 5a–5f using observational data. The results confirm that the highest predicted lifetime increases arise from the combination of higher *PWV* (+30%, +50%) paired with lower *LR_26km* (−50%, −30%), yielding an ~15% increase in median predicted lifetime relative to the baseline predicted lifetime from the unperturbed test set. This is followed by higher *PWV* (+30%, +50%) with higher BS_08km (+30%, +50%) yielding an ~8% increase, consistent with the observational patterns discussed above. Across all *Mass_f* perturbation scenarios (−50%, −30%, +30%, and +50%), predicted lifetime remains largely governed by *PWV*, with *Mass_f* contributing negligible modulation (~±1%) regardless of the direction or magnitude of *Mass_f* perturbation, further corroborating the dominant role of *PWV* seen across all *Mass_f* groups in Figs. 5c and 5f. To complement these two-variable results, independent perturbations of each predictor (−50%, −30%, +30%, and +50%) while holding all other inputs the same (Fig. S5 a–d), further confirm the dominant role of *PWV*, with up to an ~11% increase in median predicted lifetime under positive perturbations (+30%, +50%) and a modest decrease (< 5%) at −30%, whereas the −50% perturbation yields a

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

marginal positive deviation (~2%), potentially attributable to the limited training data. For *LR_26km*, negative perturbations (−50%, −30%) are associated with ~10% higher predicted lifetime, whereas positive perturbations (+30%, +50%) yield modest reductions (< 5%). *BS_08km* shows a positive but limited independent effect, with predicted lifetime increasing under positive perturbations (+30%, +50%) and decreasing under negative perturbations (−50%, −30%), though the magnitude remains limited (< ±5%) in both directions. *Mass_f* exhibits no systematic or consistent response in either direction, consistent with its near-zero partial dependence (Fig. S5d). Taken together, the percentile-based analyses and RF perturbation sensitivity tests consistently indicate that *PWV* exerts a dominant and robust control on convective cell lifetime, with *LR_26km* and *BS_08km* providing meaningful but secondary contributions, particularly in combination with high *PWV*. *Mass_f* shows no consistent or systematic independent influence on predicted cell lifetime, with the exception of a statistically significant lifetime difference between low and high *Mass_f* groups under the lowest third of *PWV* values (Fig. 5f). Within this lowest *PWV* group, the negative correlation between *Mass_f* and *PWV* ($r = -0.33$) suggests that *PWV* is not the primary driver of the observed *Mass_f*–lifetime relationship in that bin; however, modulation by other atmospheric variables cannot be fully excluded. Overall, any aerosol-related signal in the present dataset appears contingent on the prevailing moisture regime, and the limited and non-systematic nature of the *Mass_f* influence does not provide robust evidence for an independent aerosol effect on convective cell lifetime.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

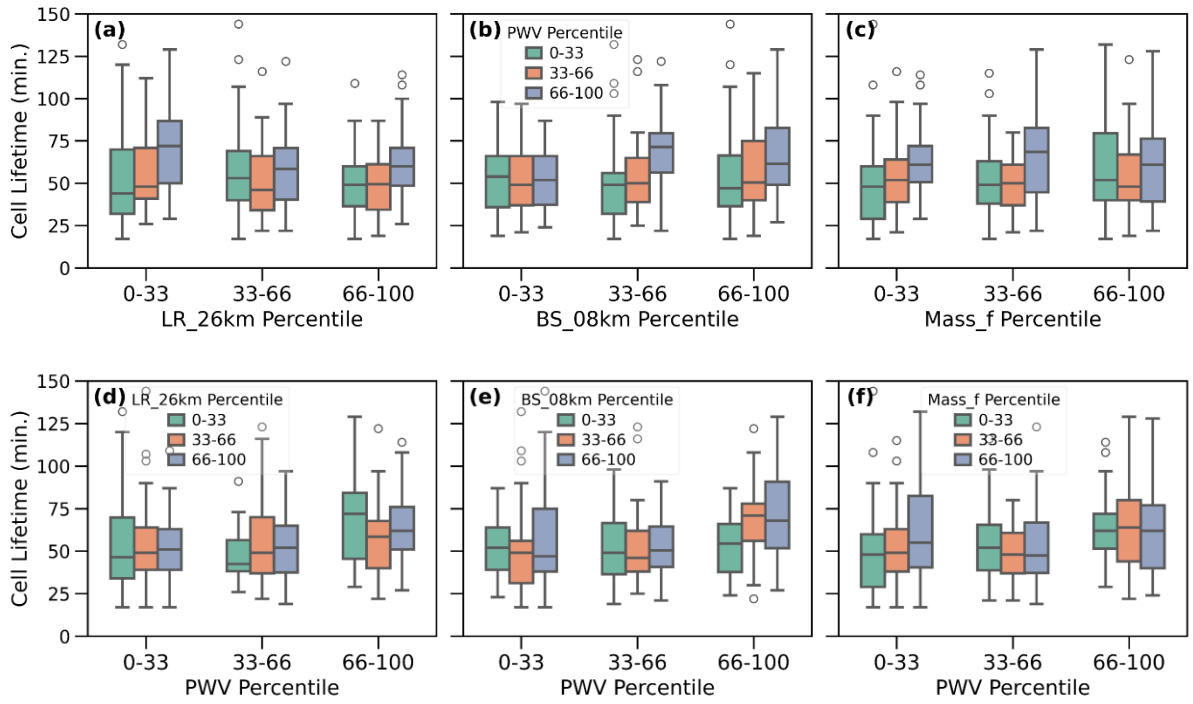


Fig. 5 (a–c) Boxplots of convective cell lifetime (in min) as a function of different variables, stratified by *PWV* percentiles (0–33rd, 33–66th, 66–100th). (a) Cell lifetime as a function of *PWV* and *LR_26km*. (b) Cell lifetime as a function of *PWV* and *BS_08km* (c) Cell lifetime as a function of *PWV* and *Mass_f* (d–f) Boxplots of convective cell lifetime (in min) as a function of *PWV* percentiles, stratified by percentiles of (d) *LR_26km* (e) *BS_08km* (f) *Mass_f*, within each *PWV* group. Box represents the interquartile range (IQR; 25th–75th percentile), the horizontal line inside the box indicates the median, the whiskers extend to 1.5×IQR Each of the nine subgroups consists of ~44–45 samples.

To better understand convective cell lifetime, it is essential to examine cell properties such as cell height and width, as these characteristics are closely linked to the physical processes governing cell development and lifetime. It should be noted that “cell” is defined by the extent of the precipitation field, which may or may not be an indicator of the extent and intensity of the updraft that produces the precipitation. The vertical extent of storms influences the redistribution of heat, moisture, momentum, aerosols, and trace gases (Gagin, 1985), which in turn influences cell lifetime. The width of convective cells is primarily influenced by available instability, whereas the depth of narrow cells is more sensitive to midlevel relative humidity (Varble et al. 2024), suggesting that broader cells may be more resilient to environmental variability. Therefore, examining cell height and width provides critical insight into the

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

mechanisms controlling convective cell lifetime around the Houston region. The frequency distributions of maximum cell width at 2–4 and 4–6 km AGL as well as the height, defined using CSAPR2 radar reflectivity thresholds of $Z > 20$ dBZ (Section 2b1) are analyzed over the lifetime of each cell. The distributions of convective cell widths and heights exhibit clear stratification by lifetime category. For both the 2–4 km and 4–6 km AGL layers (Figs. 6a–b), short-lived cells (0–40 min) are concentrated at narrower widths, with peak frequencies occurring between ~2–7 km. In the 2–4 km layer, long-lived cells (80+ min) show a broader width distribution extending well beyond 20 km, with increased frequencies observed in the ~7–17 km range. A similar pattern is evident in the 4–6 km layer, where long-lived cells exhibit a wide spread of widths, peaking at the ~7–15 km range. Intermediate-lived cells show transitional behavior in both layers, with width distributions that lie between those of short- and long-lived cells. Quantitatively, the mean width of short-lived cells was ~6.41 km (2–4 km) and ~5.85 km (4–6 km), while long-lived cells reached mean widths of ~13.2 km and ~12.9 km, respectively. A similar pattern is evident in the cell height distribution (Fig. 6c). Short-lived cells predominantly peaked between ~5-7 km AGL, while long-lived cells exhibited a broader distribution with elevated frequencies extending to ~15 km. The mean maximum height of short-lived cells was ~7.32 km, increasing to ~8.81 km for intermediate-lived cells, and reaching ~11.4 km for long-lived cells.

Cell heights and widths were also examined using other Z thresholds (e.g., 10, 30, 40, 50 dBZ), and the overall patterns are consistent. These results indicate that longer-lived cells tend to be both wider and deeper, supporting the idea that wider updrafts are less susceptible to entrainment mixing (Morrison 2017; Peters et al. 2020; Lecoanet and Jeevanjee 2019; Varble et al. 2024). Entrainment of dry environmental air into clouds leads to buoyancy decreases through mixing and evaporative cooling, promoting weakening and dissipation. Wider cells mitigate this effect by maintaining stronger updraft cores, thereby sustaining convective development for longer durations. Further, deeper cells can also produce a wider precipitation footprint due to the longer time taken by precipitation to reach the surface from higher altitudes, indicating a bidirectional relationship between cell width and depth. In contrast, narrower cells are more vulnerable to entrainment and dissipate more rapidly. However, an intermediate group of cells with moderate height may also exhibit long lifetimes, potentially due to favorable environmental conditions such as shear and cold pool interactions. This highlights the

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

importance of cell structural properties such as height and width in regulating convective cell lifetime.

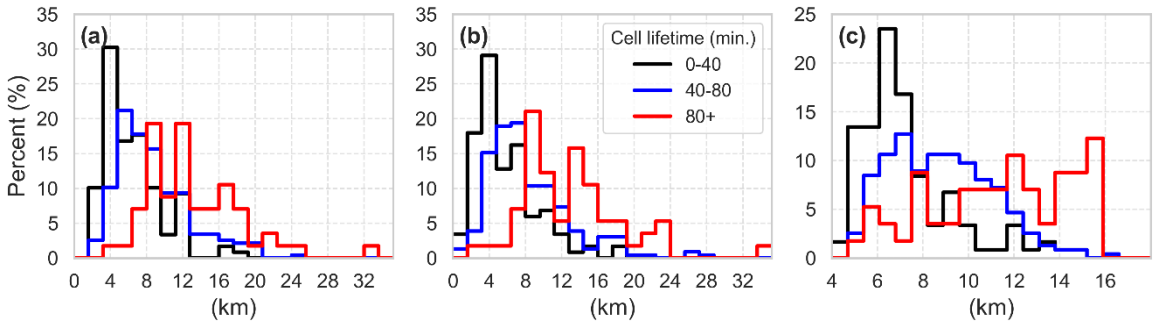


Fig. 6. Distribution of maximum convective cell (a) width at 2–4 km AGL, (b) width at 4–6 km AGL, and (c) height. for three cell lifetime categories: 0–40 min (black), 40–80 min (blue), and 80+ min (red), based on a radar reflectivity threshold of $Z > 20$ dBZ.

To further understand the factors influencing the observed variations in convective cell height and width, a similar analysis was conducted using the RF model. This analysis aimed to predict the maximum cell height and width at two height layers (2–4 and 4–6 km) based on the same environmental parameters. Monte Carlo cross-validation was employed to confirm the consistency of the results. The results highlight the mean equivalent potential temperature between the surface and 2 km (*THETAE_02km*) as the key parameter influencing maximum cell height, underscoring the critical role of low-level thermodynamic conditions in deep convection. *Mass_f* followed closely, suggesting a potential microphysical modulation of convection by aerosols, possibly through impacts on cloud droplet nucleation and latent heat release. Other influential factors include the *PWV*, *SLHF* and *BS_06km*. For maximum cell width at 2–4 km, *PWV* remains an important factor, closely followed by *LR_06km*, underscoring the role of instability and thermodynamic conditions in determining the horizontal dimensions of a storm. *Mass_f* and *BS_06km* are also prominent variables. At 4–6 km, *PWV* again showed highest importance, followed by *Mass_f* , *THETAE_02km*, and *SLHF*.

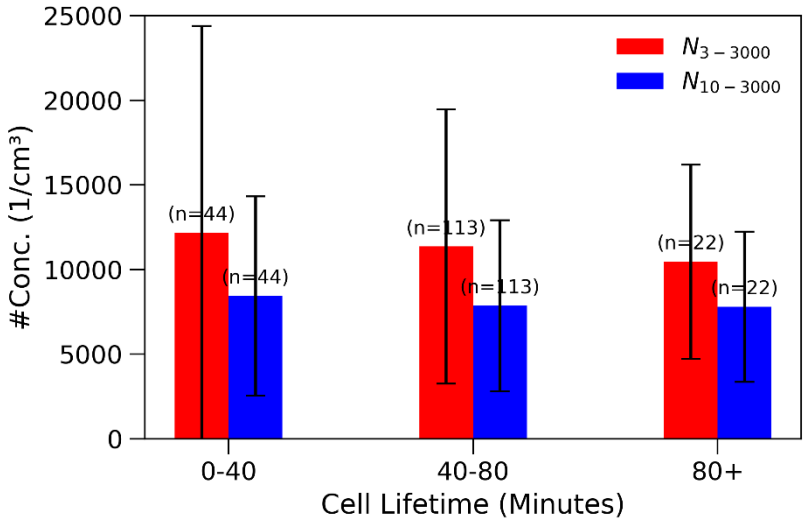
These results underscore the important role of *THETAE_02km*, *PWV*, *Mass_f* and *SLHF* in influencing both convective cell height and width. To assess the robustness of these findings, a sensitivity analysis was performed by systematically excluding the top-ranked variables from the input feature set. The RF model retained consistent performance, reinforcing the reliability

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

of the identified relationships. To further validate these findings, a bootstrap resampling approach was employed to generate 1,000 samples while preserving the original variable distributions. The RF model was retrained using Monte Carlo cross-validation on the resampled datasets. The resulting variable importance rankings remained largely stable, consistently identifying the same key variables among the top predictors. Overall, the application of bootstrapping also led to improved model performance.

c. Relationship between convective cell lifetime and surface aerosol conditions

In earlier sections of this study, *Mass_f* was used as a proxy for aerosol loading when examining convective cell properties. The TRACER campaign provided the added advantage of surface-based aerosol measurements, enabling a more detailed investigation of how aerosol size and number concentrations relate to convective cell lifetime. To investigate this particle number concentrations from the CPCU ($N_{3\text{-}3000}$) and CPCF ($N_{10\text{-}3000}$) (Fig. 7) were analyzed. While $N_{3\text{-}3000}$ exhibits a decreasing trend with increasing convective cell lifetime, this trend is not statistically significant. $N_{10\text{-}3000}$ exhibits minimal variation across lifetime categories, suggesting that the observed decrease in $N_{3\text{-}3000}$ is likely driven by changes in the 3–10 nm particle size range. Theoretical studies indicate that particles in the 3–10 nm size range are prone to rapid coagulation (Cai et al. 2022; Kulmala et al. 2004), potentially growing into larger particles capable of acting as CCN. Increased concentrations of these ultrafine particles may enhance CCN availability, which could potentially shorten convective cell lifetimes by altering cloud microphysical processes.



File generated with AMS Word template 2.0

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

Fig. 7. Comparison of particle number concentrations (cm^{-3}) across convective cell lifetime categories (i.e., 0–40, 40–80, and 80+ min) measured by CPCU ($N_{3\text{-}3000}$, red) and CPCF ($N_{10\text{-}3000}$, blue), n denotes the number of samples in each category. Error bars represent one standard deviation.

To better understand potential size-dependent effects in more detail, additional analysis separating particles into discrete size bins of 10–100 nm (ultrafine mode) and 100–1000 nm (fine mode) was conducted. This allows for a more nuanced examination of how different aerosol populations interact with convective processes. The bootstrapped aerosol particle size distribution and corresponding number concentrations for the 10–100 nm mobility diameter range ($N_{10\text{-}100}$), derived from SMPS measurements, are shown in Figs. 8a and 8b. The normalized aerosol number concentration (Fig. 8a) reveals a peak near 20 nm and shows a gradual decrease in number concentration in the 30–80 nm range with increasing cell lifetime. However, these differences are not statistically significant and do not indicate a strong influence of particles in this size range on convective cell lifetimes. Correspondingly, the $N_{10\text{-}100}$ (Fig. 8b) is highest for short-lived cells and decreases with increasing cell lifetime, but the differences between intermediate and long-lived cells are relatively small and not statistically significant. Figures 8c and 8d show the bootstrapped normalized aerosol number concentration and total number concentrations for the 100–1000 nm optical diameter range ($N_{100\text{-}1000}$) using the same cell lifetime classifications. The most notable feature is a gradual decrease in number concentration within the 500-1000 nm optical diameter range ($N_{500\text{-}1000}$) as cell lifetime increases (Fig. 8c and Fig 10c). This difference is statistically significant between short- and long-lived cells at the 95% confidence level, and between short- and intermediate-lived cells at the 90% confidence level, as determined through bootstrapped confidence intervals. Additionally, $N_{100\text{-}1000}$ is very similar between short- and long-lived cells, whereas intermediate-lived cells exhibit comparatively lower number concentrations (Fig. 8d). The overall increase in $N_{100\text{-}1000}$ for long-lived cells compared to intermediate-lived cells appears to be associated with particles in the 100–500 nm range, as the number concentration in the 500–1000 nm range is lower for long-lived cells relative to those for short-lived cells (Fig. 10c). The potential sources of these particles are further discussed later in this section.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

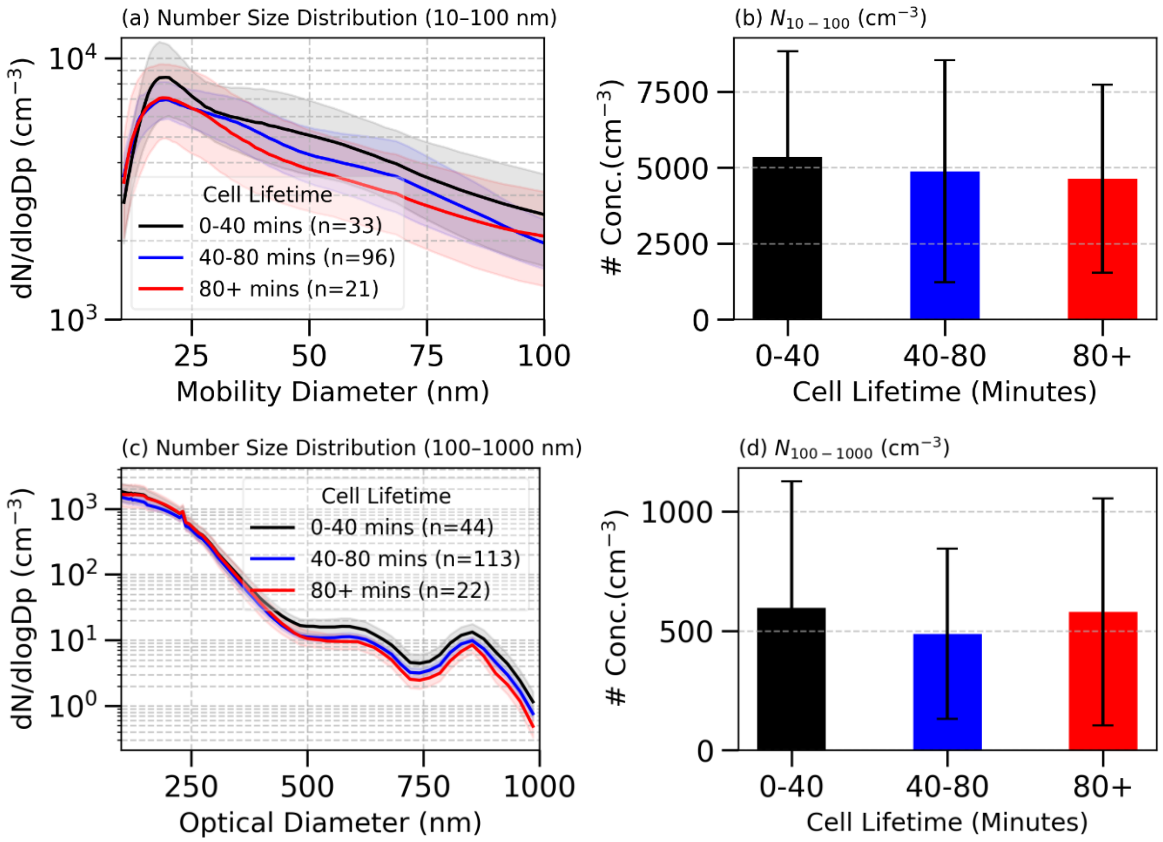


Fig. 8. Bootstrapped (a and c) aerosol size distributions and (b and d) number concentrations (b: $N_{10\text{-}100}$ and d: $N_{100\text{-}1000}$) for two particle size ranges: 10–100 nm mobility diameter measured by the SMPS (a and b) and 100–1000 nm optical diameter measured by the UHSAS (c and d). Shaded regions in (a) and (c) represent 95% confidence intervals obtained from the bootstrapping procedure, while error bars in (b) and (d) represent the standard deviation of the number concentrations. The n in the legends in (a) and (c) represent number of samples in in each category.

To investigate the potential sources of $N_{500\text{-}1000}$ which is observed to be inversely related to convective cell lifetime, convective cells were classified according to the presence of either onshore or offshore wind conditions. Although the focus is on $N_{500\text{-}1000}$, an initial look at the 10–100 nm range ($N_{10\text{-}100}$; Fig. 9b) from SMPS measurements shows consistently higher concentrations under offshore wind conditions, pointing to dominant land-based sources. The bootstrapped aerosol particle size distribution (Fig. 9c) and corresponding number concentrations ($N_{100\text{-}1000}$; Fig. 9d) derived from UHSAS measurements in the 100–1000 nm optical diameter range were analyzed using the same classification. The most notable feature

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

in this size distribution is the gradual increase in number concentration in the ~500–1000 nm optical diameter range under onshore wind conditions (Fig. 9c). This indicates the sources of aerosols in this size range are likely linked to transport by onshore winds from the Gulf. Additionally, higher $N_{100\text{-}1000}$ under offshore wind conditions predominantly originate from particles within the 100–500 nm optical diameter range.

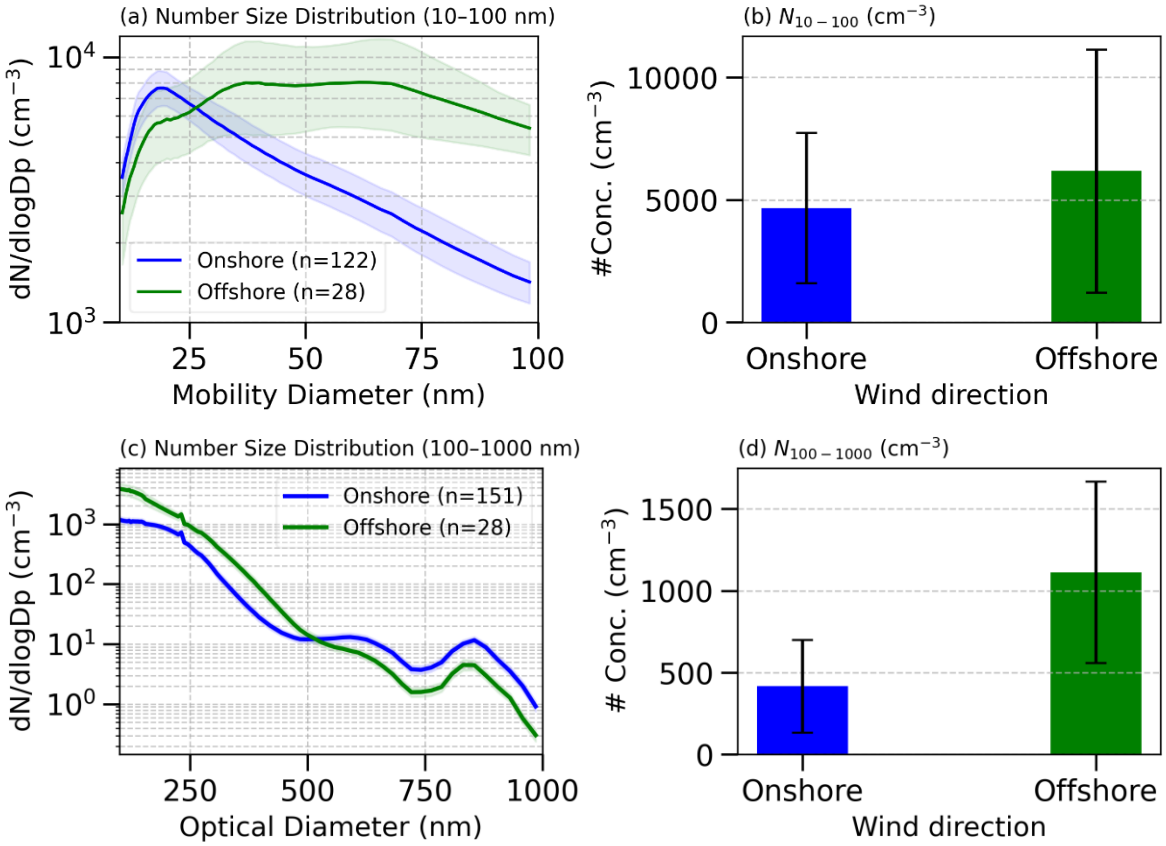


Fig. 9. Bootstrapped aerosol (a and c) size distributions and (b and d) number concentrations (b: $N_{10\text{-}100}$ and d: $N_{100\text{-}1000}$) associated with prevailing onshore and offshore winds for the (a and b) 10–100 nm mobility diameter range measured by SMPS and (c and d) 100–1000nm optical diameter range measured by UHSAS. Shaded regions in (a) and (c) represent 95% confidence intervals obtained from the bootstrapping procedure, and error bars in (b) and (d) represent the standard deviation of the concentrations. The n in the legends in (a) and (c) represent the number of samples in each category.

To further investigate how the convective cell lifetime varies under different flow regimes, cell lifetimes were examined for both onshore and offshore wind conditions. The box-and-

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

whisker plot (Fig. 10a) illustrates the variation in cell lifetimes, where convective cells associated with onshore wind conditions have slightly shorter lifetimes (mean: ~55 mean; median: ~53 min) compared to those under offshore wind conditions (mean: ~61 min; median: ~57 min). The overall variability in lifetimes remains similar for the onshore and offshore wind conditions. $N_{500\text{-}1000}$ gradually decreases with increasing cell lifetime, a trend apparent for both onshore and offshore wind conditions, although this difference is not statistically significant. Notably, $N_{500\text{-}1000}$ is consistently higher under onshore wind conditions, highlighting a stronger association of aerosols in this size range (500–1000 nm) with onshore wind conditions. Thus, aerosols in this optical diameter range (500–1000 nm) or the meteorological regimes characterized by elevated concentrations of these aerosols may be linked to shorter convective cell lifetimes. Additionally, the observed minor variations in lifetime could also result from the inability to fully isolate the effects contributed by aerosols smaller than ~500 nm.

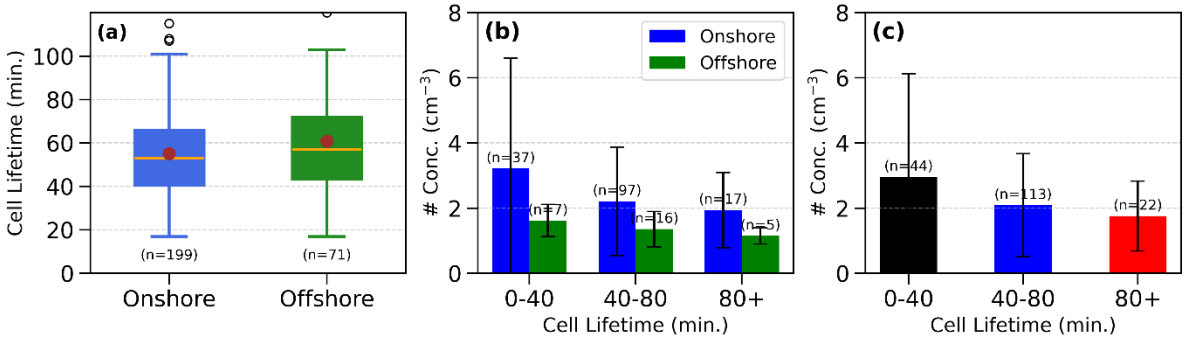


Fig. 10. (a) Box-and-whisker plot of convective cell lifetimes classified by onshore and offshore wind conditions, where dots represent means, and the horizontal orange lines represent medians. (b) Aerosol number concentration within the optical diameter range of 500–1000 nm ($N_{500\text{-}1000}$) as a function of convective cell lifetime for both onshore and offshore wind conditions, (c) Aerosol number concentration within the optical diameter range of 500–1000 nm ($N_{500\text{-}1000}$) as a function of convective cell lifetime for all the cells. n denotes the number of samples in each category.

Finally, to gain a more comprehensive understanding of aerosol influences, the aerosol composition associated with convective cell lifetimes and onshore and offshore wind conditions was examined. The aerosol composition analyzed using ACSM (Fig. 11a), reveals that total organic aerosol concentrations are highest for long-lived convective cells (mean of ~0.9 μg m^{-3}), intermediate for short-lived cells (mean of ~0.7 μg m^{-3}), and the lowest for cells with intermediate lifetime (mean of ~0.5 μg m^{-3}). In contrast, sulfate aerosol concentrations show a decreasing trend with shorter cell lifetimes, with the highest levels observed in long-

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

lived cells. Other aerosol species, including ammonium, nitrate, and chloride, do not exhibit notable variations across different cell lifetimes. Black carbon (BC), primarily produced through incomplete biomass and fossil fuel combustion, plays a critical role in atmospheric processes due to its ability to serve as CCN, influencing cloud microphysics depending on hygroscopic coatings acquired during aging (Tian et al. 2024). Laboratory studies (Dalirian et al. 2018; Gohil et al. 2023) indicate that BC's CCN activity increases with inorganic coatings and is modulated by the presence of organic fraction. Analysis of BC concentrations across cell lifetime categories (Fig. 11b) shows that short-lived cells have the highest mean mass concentration (~47 ng m^{-3}), followed by long-lived cells (~33 ng m^{-3}), and intermediate-lifetime cells the lowest (~25 ng m^{-3}). These results suggest that short-lived convective cells are associated with more polluted environments characterized by elevated BC levels. A comparison of aerosol composition between onshore and offshore wind conditions (Fig. 11c) further indicates higher concentrations of organics, sulfate, ammonium, and nitrate under offshore conditions. This reflects enhanced pollution characteristics associated with offshore environments. These observations suggest that aerosol composition and pollutant levels play a role in modulating convective cell lifetimes. However, this association may arise from underlying meteorological differences (e.g., surface winds and CAPE) between oceanic (onshore) and continental (offshore) regimes, rather than from a direct aerosol influence.

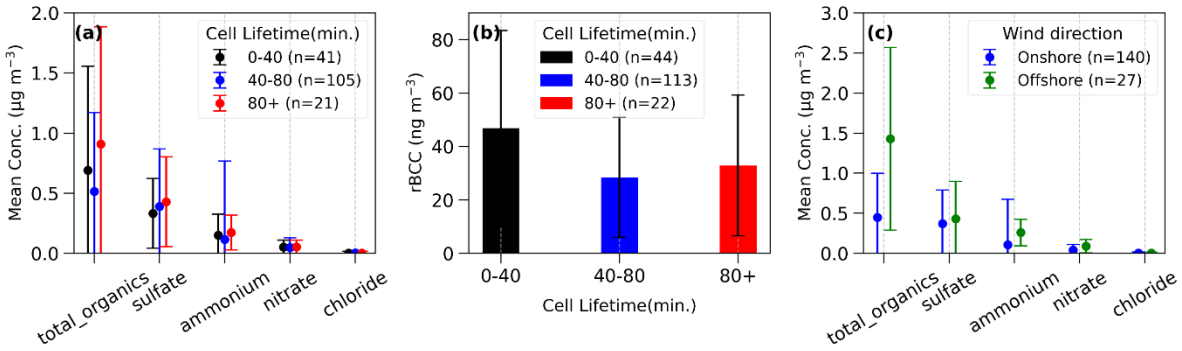


Fig. 11. (a) Mean mass concentrations (μg m^{-3}) of aerosol species (total organics, sulfate, ammonium, nitrate, and chloride) for different convective cell lifetime categories (0–40 min, 40–80 min, and >80 min). (b) Black carbon mass concentrations (ng m^{-3}) for different convective cell lifetime categories. (c) Mean mass concentrations (μg m^{-3}) of aerosol species (total organics, sulfate, ammonium, nitrate, and chloride) for onshore and offshore conditions.

d. Case study

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

In the previous section, the role of $N_{500\text{-}1000}$ in relation to shorter convective cell lifetime under higher concentration was noted. These observations suggest a potential link between $N_{500\text{-}1000}$ and convective evolution, motivating further investigation into the underlying processes. To deepen the analysis, K-means clustering was applied to atmospheric soundings associated with the convective cells within the 50 km radius around the AMF1 site, aiming to identify groups characterized by similar meteorological conditions. Based on the change in TSV, the optimal number of clusters was determined to be eight. This spatial constraint increases the probability that the aerosol observations will be representative of the air mass that the convection occurs in, resulting in four clusters (clusters 2, 3, 4, and 5) retaining sufficient samples for further analysis.

By focusing on convective cells within these four selected clusters (Fig. 12, wind profile not shown), it was ensured that meteorological conditions within each cluster are nearly similar, with differences in temperature, moisture, and wind profiles throughout the vertical column for varying clusters. The distributions of key environmental variables for each cluster are shown in Fig. S6. This clustering allowed isolation of the potential influence of $N_{500\text{-}1000}$ while minimizing the confounding effects arising from varying meteorological environments. Nevertheless, disentangling the specific impact of aerosols on convection remains challenging due to the inherently complex interaction among atmospheric processes. To optimally leverage the available aerosol data and maximize contrast, a percentile-based approach was adopted, selecting the lowest 0–20th percentile of $N_{500\text{-}1000}$ to represent the LOW $N_{500\text{-}1000}$ regime, and the highest 80–100th percentile to represent the HIGH $N_{500\text{-}1000}$ regime. By comparing only these distinctly contrasting aerosol environments, the identified aerosol-driven differences in convective structure are more robust. To assess whether meteorological variability was adequately controlled within each cluster, we compared the mean values of key parameters *PWV*, *BS_08km*, and *LR_26km* between the LOW and HIGH $N_{500\text{–}1000}$ regimes (Table S1). Sample sizes within each cluster–percentile subset is very small (3–6 samples), preventing robust statistical testing and limiting the usefulness of distributional plots. Even so, the mean comparison showed that, within each cluster, *PWV* and *LR_26km* are broadly consistent across aerosol groups (LOW and HIGH), with only a modest *BS_08km* difference noted in Cluster 4, where the LOW-aerosol group exhibited notably higher *BS_08km* values. Using this framework, the vertical profiles of Z, differential reflectivity (Z_{DR}), and specific differential phase (K_{DP}) derived from CSAPR2 RHI scans corresponding to the cross section with maximum Z within each cell's tracked lifetime were examined. These vertical profiles

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

represent the mean and standard deviation of radar variables (Z, Z_{DR} and K_{DP}) within convective cells across all events in each aerosol regime (LOW and HIGH). For each case, the 1st–99th percentiles of each radar variable were computed at each height level, and the mean across these percentiles was taken as the representative profile for that case. The group mean and standard deviation were then computed across all case profiles within each aerosol regime (Fig. 13). Despite nearly similar meteorological conditions within each cluster, notable differences emerged between the LOW and HIGH $N_{500\text{-}1000}$ regimes, providing valuable insight into how $N_{500\text{-}1000}$ variations independent of meteorological state can influence convective microphysical structures.

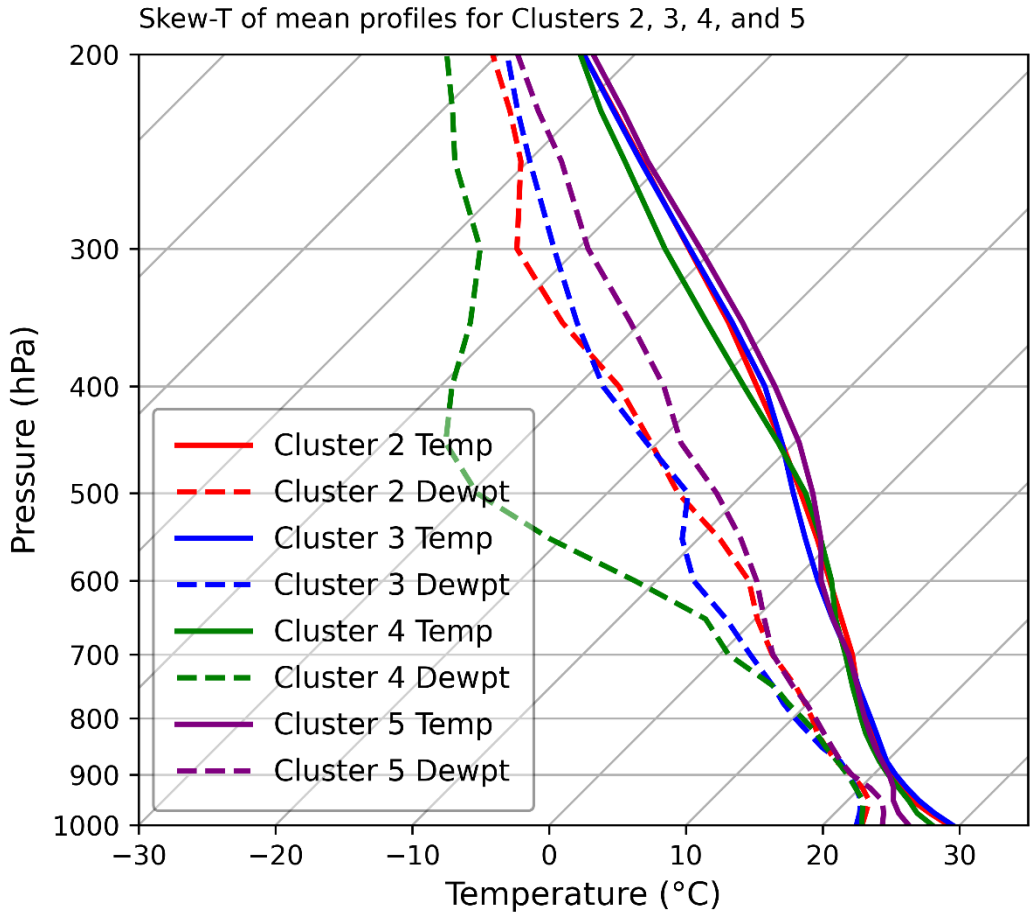


Fig. 12 Skew-T diagram showing mean temperature (solid lines) and dewpoint (dashed lines) profiles for Clusters 2–5, derived from K-means clustering of soundings associated with convective cells within a 50-km radius of the AMF1 site. These four clusters are selected from a total of eight identified clusters.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

In Cluster 2 (Figs. 13a–c), Z profiles show slight suppression in high $N_{500\text{-}1000}$ from 1–6 km, with LOW peaking at ~36 dBZ (median) near ~2 km and HIGH peaking at ~33 dBZ near the surface. The Z_{DR} and K_{DP} differences are modest, with elevated Z_{DR} in LOW up to ~4 km and elevated K_{DP} in HIGH below 3 km. Cluster 3 (Figs. 13 d–f) reveals large differences: in LOW, Z shows stronger reflectivity, with values exceeding ~38 dBZ between ~1 and 3.5 km while Z in HIGH exhibits a relatively weak reflectivity profile with values remaining below ~30 dBZ throughout the column. Z_{DR} is elevated between 1–5 km in LOW (>~1.5 dB at 1–3 km), while in HIGH it peaks near ~1 km (~1.4 dB), whereas K_{DP} is higher in LOW from the surface to ~4 km (>~1.5° km^{-1} at 1–2 km) with peak values at 1 km exceeding 2° km^{-1} compared to <~1.2° km^{-1} throughout the column in HIGH. The large variations in all three variables between LOW and HIGH, particularly below ~4 km, likely reflect microphysical differences associated with change in $N_{500\text{-}1000}$. In HIGH the vertical profiles show systematically weaker Z, Z_{DR}, *and* K_{DP} throughout the column. In contrast, LOW group shows stronger Z, Z_{DR}, *and* K_{DP} profiles specifically below 4 km (Figs. 13d–f). In Cluster 4 (Figs. 13g–i), LOW and HIGH show comparable Z and K_{DP} profiles up to ~4 km, while the primary difference is observed in Z_{DR}. Z values in both cases remain near ~35–37 dBZ below ~3 km, although LOW is ~2–3 dBZ stronger at ~4–5 km. Likewise, K_{DP} profiles are broadly comparable, ranging from ~0.5–1.2 ° km^{-1} up to 5 km, with LOW exhibiting slightly enhanced values below ~1.5 km and HIGH showing marginally larger values at 3–4 and 5–6 km. In contrast, HIGH exhibits substantially larger Z_{DR} values up to ~6 km with peak values crossing ~3 dB at ~1 km, whereas LOW remains closer to ~1.4 dB at the same level (Figs. 13g–i). Cluster 5 (Figs. 13j–l) exhibits deeper Z > ~40 dBZ up to ~4 km and > 20 dBZ to ~7.5 km in LOW, while HIGH decreases sharply to < ~20 dBZ by ~5 km. Z_{DR} in LOW is elevated (> ~1.2 dB up to 3.5 km) while in HIGH it peaks near the surface (>~1 dB). K_{DP} is generally higher in LOW, while HIGH exhibits suppressed K_{DP} throughout the column. In LOW, peak K_{DP} of ~1.8 and ~1.9 ° km^{-1} are observed near ~2 and ~3 km, respectively (Fig. 13l). The most apparent signature of a classical warm rain process, marked by a rapid increase in Z, Z_{DR}, and K_{DP} toward the surface due to collision-coalescence below a cloud base observed in Cluster 5 under HIGH.

Overall, Clusters 2, 3 and 5 exhibit consistent differences in Z and Z_{DR} profiles between high and low $N_{500\text{-}1000}$ regimes under nearly similar meteorological conditions within each cluster. In general, high $N_{500\text{-}1000}$ loading is associated with suppressed low-to-mid-level Z (Clusters 2, 3 and 5), while low $N_{500\text{-}1000}$ cases exhibit larger Z indicative of more efficient growth of precipitation-sized hydrometeors. Z_{DR} profiles under low $N_{500\text{-}1000}$ conditions show

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

elevated values at ~2–4 km (Clusters 2, 3 and 5), suggesting active coalescence. K_{DP} shows clear separation (Clusters 3 and 5), with low $N_{500\text{-}1000}$ regimes supporting higher K_{DP} peaks (~1–4 km), while high $N_{500\text{-}1000}$ cases display lower and generally weaker K_{DP} (Clusters 3 and 5), consistent with reduced rainwater mass and droplet concentration. Collectively, the results show differences in the warm-rain pathway under high $N_{500\text{-}1000}$ environments, potentially sourced from the onshore winds (Fig. 9c). Cluster 4 does not follow this consistent pattern, likely due to imperfect meteorological separation within this cluster, as reflected by the notably higher *BS_08km* values in LOW cases compared to HIGH cases (Table S1). Marine aerosols are generally considered efficient CCN, but at high concentrations, they may delay droplet coalescence resulting in suppressed warm rain development. In addition, $N_{500\text{-}1000}$ is very low (up to ~8 cm^{-3}), and the sample size in these categories is limited, so part of the observed sensitivity may instead reflect the distinct meteorological regimes associated with episodes of elevated $N_{500\text{-}1000}$ rather than $N_{500\text{-}1000}$ itself. Although the best attempt was made to effectively separate such regimes using the clustering approach, the number of cases considered is small, limiting a strong conclusion. A more detailed, process-focused analysis using model sensitivity experiments would be required to clearly separate these influences.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

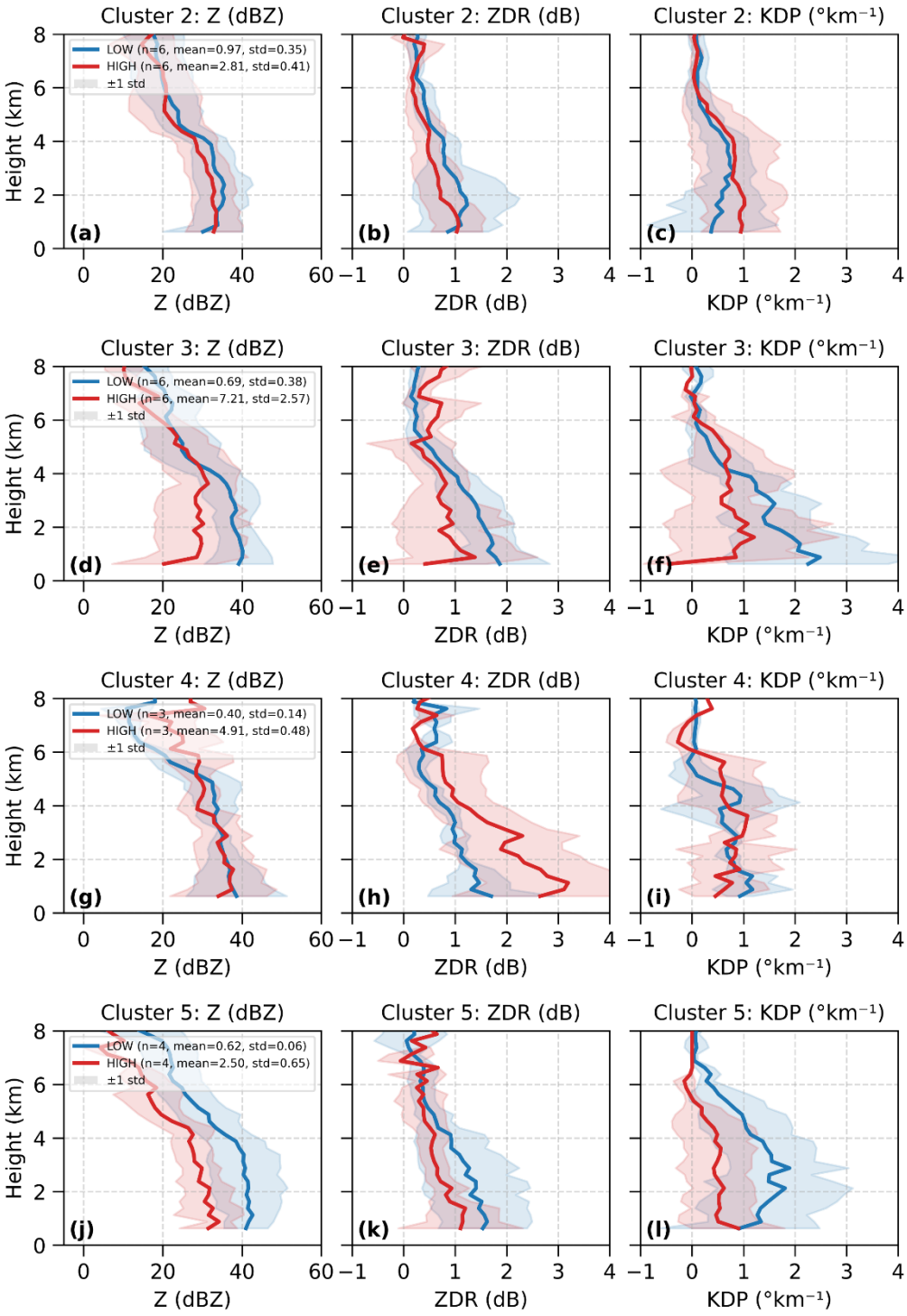


Fig. 13. Vertical profiles of radar reflectivity (Z, dBZ; left column), differential reflectivity (Z_{DR}, dB; middle column), and specific differential phase (K_{DP}, ° km^{-1}; right column) for convective cells within different sounding-based k-means clusters (rows: Cluster 2, a–c; Cluster 3, d–f; Cluster 4, g–i; Cluster 5, j–l). For each case, the 1st–99th percentiles of each radar variable were computed at each height level, and the mean across these percentiles was taken as the representative profile for that case. Profiles shown represent the mean and standard

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

deviation computed across all case profiles within each regime, stratified by aerosol concentration regimes defined by $N_{500\text{-}1000}$ percentiles: LOW (0–20th percentile; blue) and HIGH (80–100th percentile; red). Shaded regions represent ±1 standard deviation across cases. The number of convective cells (*n*), along with the mean and standard deviation (std) of $N_{500\text{-}1000}$, is annotated in each *Z* panel.

Figure 14 presents convective cell properties stratified by $N_{500\text{-}1000}$ across four sounding-based K-means clusters (Clusters 2, 3, 4, and 5). A consistent pattern is observed in cell lifetime across Clusters 3, 4, and 5 (Figs. 14d, g, and j), with shorter mean lifetimes associated with HIGH $N_{500\text{-}1000}$. Mean lifetimes decrease from 61.5 to 50.5 minutes in Cluster 3, from 74.3 to 43.0 minutes in Cluster 4, and from 64.0 to 44.8 minutes in Cluster 5, aligning with the observations discussed in the previous section. However, Cluster 2 (Fig. 14a) shows a slightly higher mean lifetime under high $N_{500\text{-}1000}$ conditions, which may reflect classification artifacts, as the mean $N_{500\text{-}1000}$ concentrations for HIGH and LOW regimes in this cluster are not well-separated (low ~0.97±0.35 and high ~2.81±0.41; Fig. 13a), suggesting potential overlap. For Cluster 3, this difference is large and is reflected in Z, Z_{DR}, K_{DP} profiles (Figs 13 d–f), lifetime, height and width (Figs. 14 d–f) showing larger suppression with HIGH $N_{500\text{-}1000}$. However, the relationship between $N_{500\text{-}1000}$ and maximum cell height varies by cluster (Figs. 14 b, e, h, and k): convective cells in Cluster 2 are taller under high $N_{500\text{-}1000}$ conditions, whereas Clusters 3 and 5 exhibit taller cells under low $N_{500\text{-}1000}$ conditions and minor variations observed in cluster 4. In contrast, the maximum cell width at 2–4 km shows a consistent trend across all clusters, with wider cells consistently associated with low $N_{500\text{-}1000}$ and narrower cells under high $N_{500\text{-}1000}$ conditions (Figs. 14c, f, i, and l). This result aligns with earlier analysis in Section 3b that *Mass_f* is among the top 4 important predictors for cell width, highlighting the sensitivity of horizontal storm structure to aerosol loading. Conversely, cell height was primarily influenced by low-level thermodynamics (*THETAE_02km*), followed by *Mass_f*, suggesting that differences in meteorological state among clusters modulate the aerosol impact on convective vertical structure.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

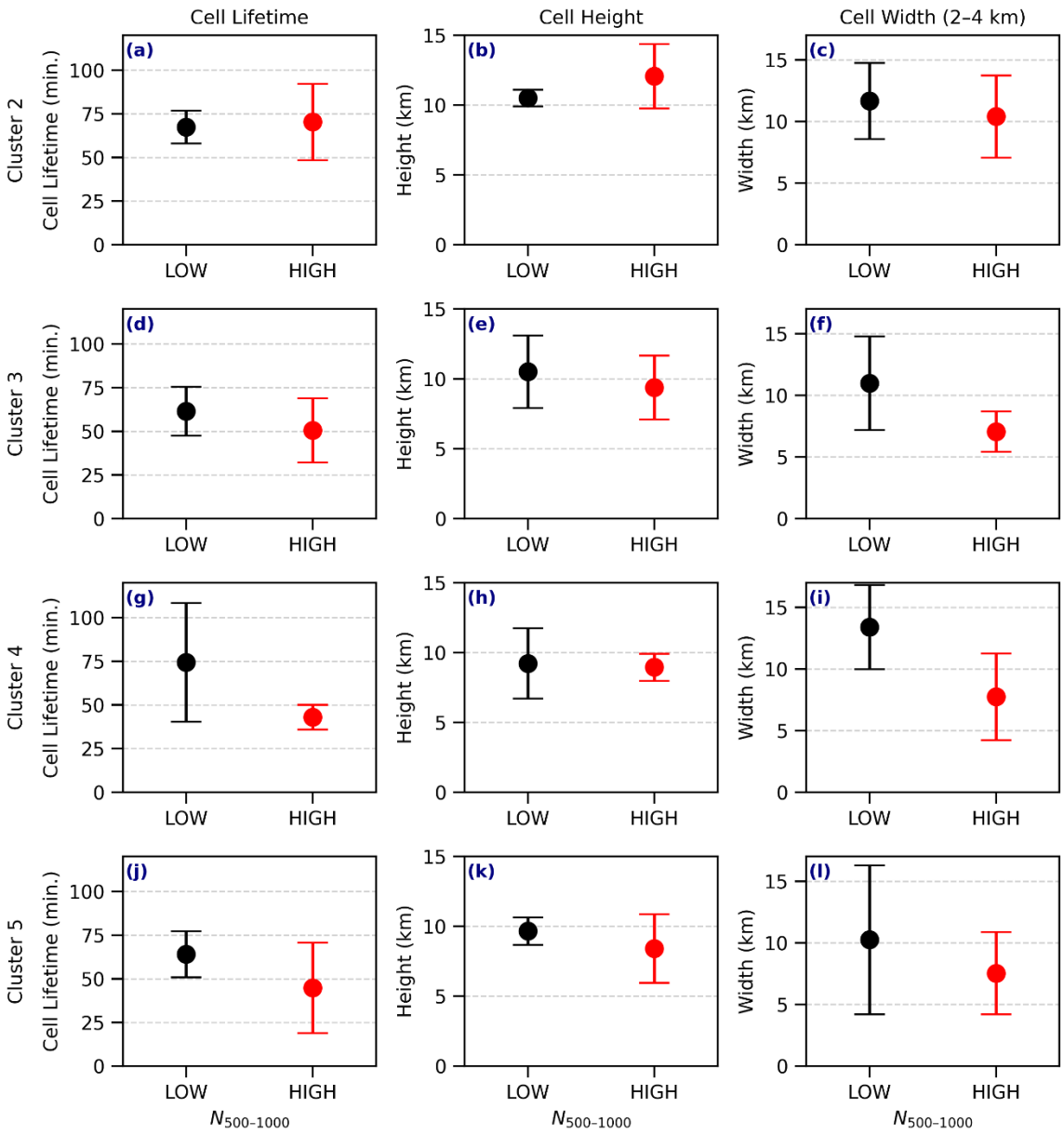


Fig. 14 Convective cell properties under different $N_{500\text{-}1000}$ conditions for four sounding-based k-means clusters: Cluster 2 (a–c), Cluster 3 (d–f), Cluster 4 (g–i), and Cluster 5 (j–l). Each row represents one cluster, with columns showing (from left to right): (1) cell lifetime (min), (2) maximum cell height (km) where reflectivity > 20 dBZ, and (3) maximum cell width at 2–4 km (km) where reflectivity > 20 dBZ. Bar colors indicate aerosol concentration regimes defined by $N_{500\text{-}1000}$: LOW (0–20th percentile, black) and HIGH (80–100th percentile, red). Error bars indicate variability (±1 standard deviation) within each aerosol group.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

4. Conclusions

New insights into the complex interactions between aerosols, meteorological conditions, and convective cell properties were investigated by analyzing data from the TRACER IOP conducted in the vicinity of Houston. By integrating radar data from CSAPR2 and KHGX, ERA5 reanalysis, MERRA2 aerosol reanalysis, and ground-based aerosol measurements, over ~400 deep convective cells under diverse environmental conditions were analyzed. The key findings are summarized as follows:

1) A random forest model identified precipitable water vapor (*PWV*), 2–6 km temperature lapse rate (*LR_26km*), 0–8 km bulk shear (*BS_08km*), and fine aerosol mass concentration (*Mass_f*) as the most important predictors of cell lifetime. *PWV* and *LR_26km* were also identified by various correlation tests (Section 3a). In low-*Mass_f* (clean; 0–33rd percentile) environments, increasing *PWV* was significantly associated with longer convective cell lifetimes, with median values of ~48 minutes in low *PWV* and ~61 minutes in high *PWV*. In moderate *Mass_f* (33–66th percentile) conditions, a significant increase in lifetime was observed between low and high *PWV* (median values of ~49 and ~68 minutes, respectively), indicating that moisture remains an important control even under moderate aerosol loading. Overall, *PWV* consistently emerges as the dominant factor influencing convective cell lifetime across aerosol regimes, with significantly longer cell lifetimes observed in the Moderate *Mass_f*–high *PWV* environments.

2) Deeper and wider convective cells were associated with longer lifetimes, while shallower and narrower cells exhibited shorter lifespans. At 2–4 km AGL, short-lived cells (0–40 min) peaked between ~2–7 km in width (Z > 20 dBZ, mean: ~6.4 km), while long-lived cells (80+ min) peaked at ~7-17 km. A similar pattern was observed at 4–6 km AGL, with mean widths of ~5.8 km for short-lived and ~12.9 km for long-lived cells. Cell heights (Z > 20 dBZ) exhibited the same trend, with mean maximum values increasing from ~7.3 km for short-lived cells to ~11.4 km for long-lived cells, indicating deeper vertical development associated with longer-lived convection.

3) Ground-based aerosol measurements highlight statistically significant differences in aerosol number concentrations within the 500–1000 nm range ($N_{500\text{-}1000}$) between short- and long-lived ($p = 0.05$) as well as intermediate- and long-lived ($p = 0.10$) cells based on bootstrap analysis, indicating that larger $N_{500\text{-}1000}$ loadings were associated with shorter convective cell lifetime. The onshore wind was identified as the primary source of these larger particles.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

Furthermore, aerosol composition analysis shows that long-lived cells were characterized by elevated concentrations of organic aerosols and sulfate, while short-lived cells exhibited higher concentrations of black carbon.

4) A detailed analysis using sounding-based K-means clusters reveals that variations in $N_{500\text{-}1000}$ are associated with systematic differences in convective microphysical structures under meteorological conditions that are broadly comparable within each cluster (Fig. 12). High $N_{500\text{-}1000}$ is generally associated with suppressed radar reflectivity (Z) profiles, characterized by reduced magnitudes (Clusters 2, 3, and 5). Polarimetric variables such as Z_{DR} and K_{DP} further highlight these aerosol impacts. Low-altitude (~2–4 km) regions in low $N_{500\text{-}1000}$ regimes show higher Z_{DR} values (Clusters 3, and 5), together with higher K_{DP} in mid-levels (~2–4 km; Clusters 3 and 5), consistent with enhanced rain water content and drop growth relative to high $N_{500–1000}$ cases. These microphysical distinctions are reflected in observable differences in storm morphology, with high $N_{500\text{-}1000}$ conditions consistently associated with shorter-lived (Clusters 3, 4 and 5) and narrower convective cells. Although cell-height responses to aerosol variability differ among clusters, the consistent narrowing of cell width at 2–4 km under high $N_{500\text{-}1000}$ conditions points to a systematic aerosol influence on horizontal convective structure. Overall, these patterns suggest that high $N_{500\text{-}1000}$, potentially driven by aerosols associated with onshore winds (Fig. 9c).

Marine aerosols are generally considered efficient CCN, but at high concentrations they may delay droplet coalescence. However, $N_{500–1000}$ concentrations are relatively low, and the number of cases within these categories is limited. As a result, part of the observed sensitivity may instead reflect distinct meteorological regimes that co-occur with elevated $N_{500–1000}$ rather than $N_{500–1000}$ itself. Although the clustering approach was designed to separate such regimes, the available sample is small, limiting the strength of any causal interpretation. A more detailed, process-level modeling analysis would be required to clearly disentangle these influences. It is also noted that the upper size bound of the 500–1000 nm aerosol category reflects the measurement limit of the UHSAS instrument (optical instrument); number concentrations of particles larger than 1000 nm were not available, and therefore the effective upper bound of this size range may extend beyond 1000 nm.

This study highlights the value of combining observational datasets, reanalysis products, and advanced analytical methods to investigate factors influencing convective cell lifetimes. By using the random forest model, complex relationships between environmental factors and

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

convective cell properties were effectively captured. Statistical techniques such as bootstrapping addressed challenges from uneven and small sample sizes, ensuring robust assessment of aerosol size distributions and significance testing. In conclusion, the results deepen the understanding of cloud-aerosol interactions in deep convection, emphasizing the interplay between environmental conditions and convective cell properties. Aerosols in the 500–1000 nm diameter range are associated with shorter cell lifetimes over the Houston region. However, disentangling the effects of aerosols from the complementary impacts arising from the other aerosol size ranges remains challenging. For instance, smaller aerosols may act as CCN, promoting cloud growth and indirectly extending convective cell lifetimes, whereas larger aerosols may alter precipitation processes and affect cell dynamics differently. Although these approaches improved the understanding of the relationship between aerosols and convective cell lifetimes, uncertainties persist due to the overlapping roles of different aerosol size ranges and their covariability with meteorological conditions. The limited sample size (413 convective cells) also constrains the generalizability of the findings. These complexities highlight the need for further research to disentangle the intricate interactions among aerosol compositions, particle size distributions, and cloud microphysical processes. Future research should evaluate whether cloud-resolving models can reproduce these observed patterns and better characterize convective cloud-environment interactions using data from convective cells observed during the TRACER field campaign.

Acknowledgments.

This research was supported primarily by the U.S. Department of Energy's (DOE) Atmospheric System Research (ASR), an Office of Science Biological and Environmental Research program, under Grant No. DE-SC0024317 and under contract DE-SC0012704. The computing for this project was performed at the OU Supercomputing Center for Education & Research (OSCER) at the University of Oklahoma (OU). We thank two anonymous reviewers for their constructive comments and suggestions, which improved the clarity and quality of this manuscript. We also thank Prof. Jason Furtado for his help with a statistical question.

Data Availability Statement.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

All Atmospheric Radiation Measurement (ARM) TRACER campaign data can be accessed through the ARM Data Center at https://www.arm.gov/data/. NEXRAD radar data is publicly available via the NOAA Open Data Registry at https://registry.opendata.aws/noaa-nexrad/ . The ERA5 reanalysis dataset is available at https://doi.org/10.5065/BH6N-5N20. MERRA-2 M2I3NVAER aerosol reanalysis product is available at https://doi.org/10.5067/LTVB4GPCOTK2 (last access: 13 February 2025).

APPENDIX

Environmental parameters

Table A1. Environmental parameters associated with convective cells.

Variable Name	Description
SB_CAPE	Surface based parcel-derived convective available potential energy
SB_LI	Lifted index at 500hPa
SB_CIN	Surface based parcel-derived convective inhibition
SB_MIXR	Mixing ratio at the height of the surface-based parcel
SB_E_LI	Entraining lifted index at 500hPa
SB_E_WMAXSHEAR	Square root of two times SB_CAPE multiplied by surface to 6 km above ground level (AGL) bulk wind shear
SB_E_WMAXSHEAR_3km	Square root of two times SB_CAPE multiplied by surface to 3 km AGL bulk wind shear
ML_CIN	Convective inhibition estimated from the mixed-layer parcel
ML_MIXR	Mixing ratio at the height of the mixed-layer parcel

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

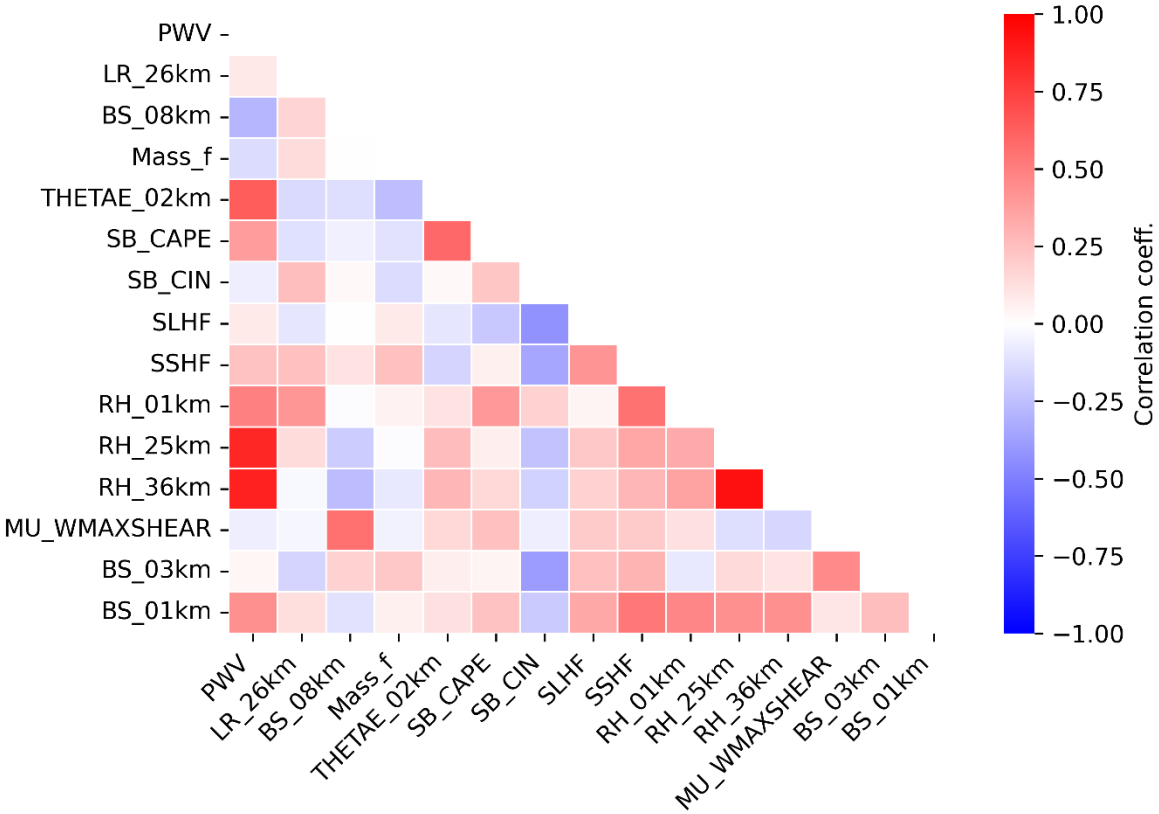
ML_WMAXSHEAR	Square root of two times ML_CAPE multiplied by surface to 6 km AGL bulk wind shear
ML_E_LI	Entraining lifted index at 500hPa
MU_CAPE_3km	Convective available potential energy between surface and 3 km estimated from the most-unstable parcel
MU_LI	Lifted index at 500hPa from the most unstable parcel
MU_CIN	Convective inhibition, estimated from the most-unstable parcel
MU_MIXR	Mixing ratio at the height of the most-unstable parcel
MU_WMAXSHEAR	Square root of two time MU_CAPE multiplied by surface to 6 km AGL bulk wind shear
MU_E_CAPE	Entraining convective available potential energy, estimated from the most-unstable parcel
MU_E_CAPE_3km	Entraining convective available potential energy between surface and 3 km AGL, estimated from the most-unstable parcel
MUML_CAPE	Convective available potential energy, estimated from the most unstable 500-m mean layer parcel
MUML_CAPE_3km	Convective available potential energy between surface and 3 km AGL, estimated from the most-unstable 500-m mean layer parcel
MUML_CIN	Convective inhibition, estimated from the most-unstable 500-m mean layer parcel
MUML_MIXR	Mixing ratio at the height of the most-unstable 500-m mean layer parcel

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

MUML_WMAXSHEAR	Square root of two time MUML_CAPE multiplied by surface to 6 km AGL bulk wind shear
MUML_E_CAPE	Entraining convective available potential energy, estimated from the most-unstable 500-m mean layer parcel
MUML_E_CAPE_3km	Entraining convective available potential energy between surface and 3 km, estimated from the most-unstable 500-m mean layer parcel
MUML_E_LI	Entraining lifted index at 500hPa from the most unstable 500-m mean layer parcel
LR_01km	Temperature lapse rate between surface and 1 km AGL (no negative sign)
LR_03km	Temperature lapse rate between surface and 3 km AGL (no negative sign)
LR_06km	Temperature lapse rate between surface and 6 km AGL (no negative sign)
LR_26km	Temperature lapse rate between 2 and 6 km AGL (no negative sign)
THETAE_01km	Mean equivalent potential temperature between surface and 1 km AGL
THETAE_02km	Mean equivalent potential temperature between surface and 2 km AGL
RH_01km	Mean relative humidity between surface and 1 km AGL
RH_02km	Mean relative humidity between surface and 2 km AGL
RH_25km	Mean relative humidity between 2- 5 km AGL

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

RH_36km	Mean relative humidity between 3- 6 km AGL
PWV	Precipitable water vapor (entire column)
BS_0500m	Bulk wind shear between surface and 500 m AGL
BS_01km	Bulk wind shear between surface and 1 km AGL
BS_03km	Bulk wind shear between surface and 3 km AGL
BS_06km	Bulk wind shear between surface and 6 km AGL
BS_08km	Bulk wind shear between surface and 8 km AGL
SLHF	Surface upward latent heat flux
SSHF	Surface upward sensible heat flux
Mass_f	Fine aerosol mass concentration



Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

Fig. A1 Heat map of the correlation matrix for select key variables, values represent pairwise Pearson correlation coefficients.

REFERENCES

Abdul-Razzak, H., and S. J. Ghan, 2000: A parameterization of aerosol activation: 2. Multiple aerosol types. *Journal of Geophysical Research: Atmospheres*, **105**, 6837–6844, https://doi.org/10.1029/1999JD901161.

Andreae, M. O., 2009: Correlation between cloud condensation nuclei concentration and aerosol optical thickness in remote and polluted regions. *Atmos Chem Phys*, **9**, 543–556, https://doi.org/10.5194/acp-9-543-2009.

Bergemann, M., and C. Jakob, 2016: How important is tropospheric humidity for coastal rainfall in the tropics? *Geophys Res Lett*, **43**, 5860–5868, https://doi.org/10.1002/2016GL069255.

Biau, G., and E. Scornet, 2016: A random forest guided tour. *TEST*, **25**, 197–227, https://doi.org/10.1007/s11749-016-0481-7.

Birch, C. E., S. Webster, S. C. Peatman, D. J. Parker, A. J. Matthews, Y. Li, and M. E. E. Hassim, 2016: Scale Interactions between the MJO and the Western Maritime Continent. *J Clim*, **29**, 2471–2492, https://doi.org/10.1175/JCLI-D-15-0557.1.

Block, K., M. Haghighatnasab, D. G. Partridge, P. Stier, and J. Quaas, 2024: Cloud condensation nuclei concentrations derived from the CAMS reanalysis. Earth Syst. Sci. Data, 16, 443–470, https://doi.org/10.5194/essd-16-443-2024

Bodas-Salcedo, A., J. P. Mulcahy, T. Andrews, K. D. Williams, M. A. Ringer, P. R. Field, and G. S. Elsaesser, 2019: Strong Dependence of Atmospheric Feedbacks on Mixed-Phase Microphysics and Aerosol-Cloud Interactions in HadGEM3. *J Adv Model Earth Syst*, **11**, 1735–1758, https://doi.org/10.1029/2019MS001688.

Borque, P., A. Varble, and J. Hardin, 2022: Peak rain rate sensitivity to observed cloud condensation nuclei and turbulence in continental warm shallow clouds during CACTI. *J. Geophys. Res. Atmos.*, **127**, e2022JD036864, https://doi.org/10.1029/2022JD036864

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

Breiman, L., 2001: Random Forests. *Mach Learn*, **45**, 5–32, https://doi.org/10.1023/A:1010933404324.

Brown, R. A., V. T. Wood, R. M. Steadham, R. R. Lee, B. A. Flickinger, and D. Sirmans, 2005: New WSR-88D Volume Coverage Pattern 12: Results of Field Tests. *Weather Forecast*, **20**, 385–393, https://doi.org/10.1175/WAF848.1.

Cai, R., E. Häkkinen, C. Yan, J. Jiang, M. Kulmala, and J. Kangasluoma, 2022: The effectiveness of the coagulation sink of 3–10 nm atmospheric particles. *Atmos Chem Phys*, **22**, 11529–11541, https://doi.org/10.5194/acp-22-11529-2022.

Chakraborty, S., R. Fu, S. T. Massie, and G. Stephens, 2016: Relative influence of meteorological conditions and aerosols on the lifetime of mesoscale convective systems. *Proceedings of the National Academy of Sciences*, **113**, 7426–7431, https://doi.org/10.1073/pnas.1601935113.

Cuesta-Valero, F. J., H. Beltrami, S. Gruber, A. García-García, and J. F. González-Rouco, 2022: A new bootstrap technique to quantify uncertainty in estimates of ground surface temperature and ground heat flux histories from geothermal data. *Geosci Model Dev*, **15**, 7913–7932, https://doi.org/10.5194/gmd-15-7913-2022.

Dalirian, M., A. Ylisirniö, A. Buchholz, D. Schlesinger, J. Ström, A. Virtanen, and I. Riipinen, 2018: Cloud droplet activation of black carbon particles coated with organic compounds of varying solubility. *Atmos Chem Phys*, **18**, 12477–12489, https://doi.org/10.5194/acp-18-12477-2018.

Davies, L., C. Jakob, P. May, V. V. Kumar, and S. Xie, 2013: Relationships between the large-scale atmosphere and the small-scale convective state for Darwin, Australia. *Journal of Geophysical Research: Atmospheres*, **118**, https://doi.org/10.1002/jgrd.50645.

Del Genio, A. D., 2012: Representing the Sensitivity of Convective Cloud Systems to Tropospheric Humidity in General Circulation Models. *Surv Geophys*, **33**, 637–656, https://doi.org/10.1007/s10712-011-9148-9.

Doswell, C. A., 1987: The Distinction between Large-Scale and Mesoscale Contribution to Severe Convection: A Case Study Example. *Weather Forecast*, **2**, 3–16, https://doi.org/10.1175/1520-0434(1987)002<0003:TDBLSA>2.0.CO;2.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

Efron, B., 1979: Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, **7**, https://doi.org/10.1214/aos/1176344552.

Efron, B., and R. J. Tibshirani, 1994: *An Introduction to the Bootstrap*. Chapman and Hall/CRC, https://doi.org/10.1201/9780429246593.

Emanuel, K. A., 1994: *Atmospheric Convection*. Oxford University PressNew York, NY, https://doi.org/10.1093/oso/9780195066302.001.0001.

Fan, J., R. Zhang, G. Li, and W. Tao, 2007: Effects of aerosols and relative humidity on cumulus clouds. *Journal of Geophysical Research: Atmospheres*, **112**, https://doi.org/10.1029/2006JD008136.

——, and Coauthors, 2009: Dominant role by vertical wind shear in regulating aerosol effects on deep convective clouds. *Journal of Geophysical Research: Atmospheres*, **114**, https://doi.org/10.1029/2009JD012352.

——, L. R. Leung, D. Rosenfeld, Q. Chen, Z. Li, J. Zhang, and H. Yan, 2013: Microphysical effects determine macrophysical response for aerosol impacts on deep convective clouds. *Proceedings of the National Academy of Sciences*, **110**, https://doi.org/10.1073/pnas.1316830110.

——, Y. Wang, D. Rosenfeld, and X. Liu, 2016: Review of Aerosol–Cloud Interactions: Mechanisms, Significance, and Challenges. *J Atmos Sci*, **73**, 4221–4252, https://doi.org/10.1175/JAS-D-16-0037.1.

——, and Coauthors, 2018: Substantial convection and precipitation enhancements by ultrafineaerosol particles. *Science (1979)*, **359**, 411–418, https://doi.org/10.1126/science.aan8461.

Feng, Z., A. Varble, J. Hardin, J. Marquis, A. Hunzinger, Z. Zhang, and M. Thieman, 2022: Deep Convection Initiation, Growth, and Environments in the Complex Terrain of Central Argentina during CACTI. *Mon Weather Rev*, **150**, 1135–1155, https://doi.org/10.1175/MWR-D-21-0237.1.

Fenwick, P. M., 1994: A new data structure for cumulative frequency tables. *Softw Pract Exp*, **24**, 327–336, https://doi.org/10.1002/spe.4380240306.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

Fridlind, A. M., and Coauthors, 2019: Use of polarimetric radar measurements to constrain simulated convective cell evolution: a pilot study with Lagrangian tracking. *Atmos Meas Tech*, **12**, 2979–3000, https://doi.org/10.5194/amt-12-2979-2019.

Gagin, A., D. Rosenfeld, and R. E. López, 1985: The Relationship between Height and Precipitation Characteristics of Summertime Convective Cells in South Florida. *J Atmos Sci*, 42, 84–94, https://doi.org/10.1175/1520-0469(1985)042<0084:TRBHAP>2.0.CO;2.

Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J Clim*, **30**, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1.

Genuer, R., J.-M. Poggi, C. Tuleau-Malot, and N. Villa-Vialaneix, 2017: Random Forests for Big Data. *Big Data Research*, **9**, 28–46, https://doi.org/10.1016/j.bdr.2017.07.003.

Giaiotti, D. B., R. Steinacker, and F. Stel, 2007: *Atmospheric Convection: Research and Operational Forecasting Aspects*. D.B. Giaiotti, R. Steinacker, and F. Stel, Eds. Springer Vienna, https://doi.org/10.1007/978-3-211-69291-2.

Gohil, K., R. Barrett, D. Rastogi, C.-N. Mao, Q. Yao, and A. Asa-Awuku, 2023: Solubility Considerations for Cloud Condensation Nuclei (CCN) Activity Analysis of Pure and Mixed Black Carbon Species. *J Phys Chem A*, **127**, 3873–3882, https://doi.org/10.1021/acs.jpca.2c08585.

Grabowski, W. W., 2018: Can the Impact of Aerosols on Deep Convection be Isolated from Meteorological Effects in Atmospheric Observations?. *J. Atmos. Sci.*, **75**, 3347–3363, https://doi.org/10.1175/JAS-D-18-0105.1.

Grabowski, W. W., and H. Morrison, 2020: Do Ultrafine Cloud Condensation Nuclei Invigorate Deep Convection?. *J. Atmos. Sci.*, **77**, 2567–2583, https://doi.org/10.1175/JAS-D-20-0012.1.

Hagos, S., Z. Feng, S. McFarlane, and L. R. Leung, 2013: Environment and the Lifetime of Tropical Deep Convection in a Cloud-Permitting Regional Model Simulation. *J Atmos Sci*, **70**, 2409–2425, https://doi.org/10.1175/JAS-D-12-0260.1.

Hartigan, J. A., and M. A. Wong, 1979: Algorithm AS 136: A K-Means Clustering Algorithm. *Appl Stat*, **28**, 100, https://doi.org/10.2307/2346830.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

Helmus, J. J., and S. M. Collis, 2016: The Python ARM Radar Toolkit (Py-ART), a Library for Working with Weather Radar Data in the Python Programming Language. *J Open Res Softw*, **4**, 25, https://doi.org/10.5334/jors.119.

Hersbach, H., and Coauthors, 2023: ERA5 hourly data on pressure levels from 1940 to present, https://doi.org/10.24381/cds.bd0915c6.

Holdridge, D., 2020: *Balloon-Borne Sounding System (SONDE) Instrument Handbook*, https://doi.org/10.2172/1020712.

Hu, J., Rosenfeld, D., Ryzhkov, A., Zrnic, D., Williams, E., Zhang, P., Snyder, J. C., Zhang, R., & Weitz, R. (2019). Polarimetric Radar Convective Cell Tracking Reveals Large Sensitivity of Cloud Precipitation and Electrification Properties to CCN. *Journal of Geophysical Research: Atmospheres*, *124*(22), 12194–12205. https://doi.org/10.1029/2019JD030857

Huang, Y., G. McFarquhar, S. Patil, L. Gao, M. Taszarek, M. Xue, A. Dzambo, M. Wolde, L. Nichman, C. Nguyen, K. Ranjbar, N. Bliankinshtein, K. Bala, P. Kollias, M. Jensen, Q. Mo, R. Bruintjes , C. Kuang , and T. Subba, 2025: Dependence of Convective Cloud Microphysical Properties on Environmental Conditions during the TRACER and ESCAPE Field Campaigns: A Synergistic Approach of Observations, Machine Learning and Numerical Models. *Journal of the Atmospheric Sciences*, conditionally accepted.

Igel, A. L., and S. C. van den Heever, 2021: Invigoration or Enervation of Convective Clouds by Aerosols? *Geophys Res Lett*, 48, https://doi.org/10.1029/2021GL093804.

Intergovernmental Panel on Climate Change (IPCC), 2023: *Climate Change 2021 – The Physical Science Basis*. Cambridge University Press, https://doi.org/10.1017/9781009157896.

Jackson, R., Sedlacek, A., Salwen, C., & Hayes, C. Single Particle Soot Photometer (AOSSP2BC60S), 2021-12-17 to 2022-10-02, ARM Mobile Facility (HOU), Houston, TX; AMF1 (main site for TRACER) (M1). Atmospheric Radiation Measurement (ARM) User Facility. https://doi.org/10.5439/1807910

Jensen, M. P. and 148 co-authors., "Studying Aerosol, Clouds and Air Quality in the Coastal Urban Environment of Southeastern Texas." *Bull. Amer. Meteor Soc.*, submitted.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

Kendall, M. G., 1938: A New Measure of Rank Correlation. *Biometrika*, **30**, 81, https://doi.org/10.2307/2332226.

——, 1945: THE TREATMENT OF TIES IN RANKING PROBLEMS. *Biometrika*, **33**, 239–251, https://doi.org/10.1093/biomet/33.3.239.

Kendall, M. G., and A. Stuart, 1973: *The Advanced Theory of Statistics, Volume 2: Inference and Relationship*. Griffin.

Kingfield, D. M., and M. M. French, 2022: The Influence of WSR-88D Intra-Volume Scanning Strategies on Thunderstorm Observations and Warnings in the Dual-Polarization Radar Era: 2011–20. *Weather Forecast*, **37**, 283–301, https://doi.org/10.1175/WAF-D-21-0127.1.

Kokoska, S., and D. Zwillinger, 2000: *CRC Standard Probability and Statistics Tables and Formulae, Student Edition*. CRC Press, https://doi.org/10.1201/b16923.

Kollias, P., E. Luke, M. Oue, and K. Lamer, 2020: Agile Adaptive Radar Sampling of Fast-Evolving Atmospheric Phenomena Guided by Satellite Imagery and Surface Cameras. *Geophys Res Lett*, **47**, https://doi.org/10.1029/2020GL088440.

Koontz, A., C. Flynn, C. Kuang, E. Andrews, C. Kuang, C. Hayes, A. Singh, and C. Salwen, 2021a: *Condensation Particle Counter (AOSCPCU), 2021-10-01 to 2022-10-01, ARM Mobile Facility (HOU), Houston, TX; AMF1 (main site for TRACER) (M1)*. Atmospheric Radiation Measurement (ARM) User Facility, accessed 9 Jun 2024, https://doi.org/10.5439/1228061.

Koontz, A., C. Kuang, E. Andrews, C. Hayes, A. Singh, and C. Salwen, 2021b: *Condensation Particle Counter (AOSCPCF), 2021-10-01 to 2022-10-01, ARM Mobile Facility (HOU), Houston, TX; AMF1 (main site for TRACER) (M1)*. Atmospheric Radiation Measurement (ARM) User Facility, accessed 9 Jun 2024, https://doi.org/10.5439/1352536.

Koren, I., G. Dagan, and O. Altaratz, 2014: From aerosol-limited to invigoration of warm convective clouds. *Science (1979)*, **344**, 1143–1146, https://doi.org/10.1126/science.1252595.

Kulmala, M., H. Vehkamäki, T. Petäjä, M. Dal Maso, A. Lauri, V.-M. Kerminen, W. Birmili, and P. H. McMurry, 2004: Formation and growth rates of ultrafine atmospheric particles:

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

a review of observations. *J Aerosol Sci*, **35**, 143–176, https://doi.org/10.1016/j.jaerosci.2003.10.003.

Ladino, L. A., A. Korolev, I. Heckman, M. Wolde, A. M. Fridlind, and A. S. Ackerman, 2017: On the role of ice-nucleating aerosol in the formation of ice particles in tropical mesoscale convective systems. *Geophys Res Lett*, **44**, 1574–1582, https://doi.org/10.1002/2016GL072455.

Laing, A. G., and J. M. Fritsch, 2000: The Large-Scale Environments of the Global Populations of Mesoscale Convective Complexes. *Mon Weather Rev*, **128**, 2756–2776, https://doi.org/10.1175/1520-0493(2000)128<2756:TLSEOT>2.0.CO;2.

Lakshmanan, V., K. Hondl, C. K. Potvin, and D. Preignitz, 2013: An Improved Method for Estimating Radar Echo-Top Height. *Weather Forecast*, **28**, 481–488, https://doi.org/10.1175/WAF-D-12-00084.1.

Lamer, K., P. Kollias, E. P. Luke, B. P. Treserras, M. Oue, and B. Dolan, 2023: Multisensor Agile Adaptive Sampling (MAAS): A Methodology to Collect Radar Observations of Convective Cell Life Cycle. *J Atmos Ocean Technol*, **40**, 1509–1522, https://doi.org/10.1175/JTECH-D-23-0043.1.

Lang, F., Y. Huang, S. T. Siems, and M. J. Manton, 2018: Characteristics of the Marine Atmospheric Boundary Layer Over the Southern Ocean in Response to the Synoptic Forcing. *Journal of Geophysical Research: Atmospheres*, **123**, 7799–7820, https://doi.org/10.1029/2018JD028700.

Lecoanet, D., and N. Jeevanjee, 2019: Entrainment in Resolved, Dry Thermals. *J Atmos Sci*, **76**, 3785–3801, https://doi.org/10.1175/JAS-D-18-0320.1.

Lei, T. M. T., S. W. I. Siu, J. Monjardino, L. Mendes, and F. Ferreira, 2022: Using Machine Learning Methods to Forecast Air Quality: A Case Study in Macao. *Atmosphere (Basel)*, **13**, 1412, https://doi.org/10.3390/atmos13091412.

Li, T., Y. Yu, X. Wang, X. Liu, Q. Mao, and Y. Yuan, 2025: A review of aerosol-cloud interactions: Mechanisms, climate effects, and observation methods. *Atmos Res*, 325, 108267, https://doi.org/10.1016/j.atmosres.2025.108267.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

Liu, B., B. Gan, F. Jia, and L. Wu, 2024: Impact of the North Pacific Meridional Mode on the Tropical Pacific Modulated by the Interdecadal Pacific Oscillation. *J Clim*, **37**, 2199–2216, https://doi.org/10.1175/JCLI-D-23-0448.1.

Mann, H. B., and D. R. Whitney, 1947: On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics*, **18**, 50–60, https://doi.org/10.1214/aoms/1177730491.

Marchand, R., N. Beagley, S. E. Thompson, T. P. Ackerman, and D. M. Schultz, 2006: A Bootstrap Technique for Testing the Relationship between Local-Scale Radar Observations of Cloud Occurrence and Large-Scale Atmospheric Fields. *J Atmos Sci*, **63**, 2813–2830, https://doi.org/10.1175/JAS3772.1.

Markowski, P., and Y. Richardson, 2010: *Mesoscale Meteorology in Midlatitudes*. Wiley, https://doi.org/10.1002/9780470682104.

May, A. A., and Coauthors, 2014: Aerosol emissions from prescribed fires in the United States: A synthesis of laboratory and aircraft measurements. *Journal of Geophysical Research: Atmospheres*, **119**, 11,826-11,849, https://doi.org/10.1002/2014JD021848.

McFarquhar, G. M., 2004: A New Representation of Collision-Induced Breakup of Raindrops and Its Implications for the Shapes of Raindrop Size Distributions. *J Atmos Sci*, **61**, 777–794, https://doi.org/10.1175/1520-0469(2004)061<0777:ANROCB>2.0.CO;2.

——, and A. J. Heymsfield, 1997: Parameterization of Tropical Cirrus Ice Crystal Size Distributions and Implications for Radiative Transfer: Results from CEPEX. *J Atmos Sci*, **54**, 2187–2200, https://doi.org/10.1175/1520-0469(1997)054<2187:POTCIC>2.0.CO;2.

McKnight, P. E., and J. Najab, 2010: Mann-Whitney <scp>U</scp> Test. *The Corsini Encyclopedia of Psychology*, Wiley, 1–1, https://doi.org/10.1002/9780470479216.corpsy0524.

Miller, M. A., K. Nitschke, T. P. Ackerman, W. R. Ferrell, N. Hickmon, and M. Ivey, 2016: The ARM Mobile Facilities. *Meteorological Monographs*, **57**, 9.1-9.15, https://doi.org/10.1175/AMSMONOGRAPHS-D-15-0051.1.

Miltenberger, A. K., P. R. Field, A. A. Hill, P. Rosenberg, B. J. Shipway, J. M. Wilkinson, R. Scovell, and A. M. Blyth, 2018: Aerosol–cloud interactions in mixed-phase convective

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

clouds – Part 1: Aerosol perturbations. *Atmos Chem Phys*, **18**, 3119–3145, https://doi.org/10.5194/acp-18-3119-2018.

Morrison, H., 2017: An Analytic Description of the Structure and Evolution of Growing Deep Cumulus Updrafts. *J Atmos Sci*, **74**, 809–834, https://doi.org/10.1175/JAS-D-16-0234.1.

——, and Coauthors, 2020: Confronting the Challenge of Modeling Cloud and Precipitation Microphysics. *J Adv Model Earth Syst*, **12**, https://doi.org/10.1029/2019MS001689.

Mudelsee, M., 2011: The Bootstrap in Climate Risk Analysis. *In Extremis*, Springer Berlin Heidelberg, 44–58, https://doi.org/10.1007/978-3-642-14863-7_2.

Ng, N. L., and Coauthors, 2011: An Aerosol Chemical Speciation Monitor (ACSM) for Routine Monitoring of the Composition and Mass Concentrations of Ambient Aerosol. *Aerosol Science and Technology*, **45**, 780–794, https://doi.org/10.1080/02786826.2011.560211.

Noether, G. E. (Gottfried E., 1967): Elements of nonparametric statistics.

Ortega, J., J. R. Snider, J. N. Smith, and J. M. Reeves, 2019: Comparison of aerosol measurement systems during the 2016 airborne ARISTO campaign. *Aerosol Science and Technology*, **53**, 871–885, https://doi.org/10.1080/02786826.2019.1610554.

Oue, M., B. Treserras, E. Luke, and P. Kollias, 2023: *CSAPR2 Optimized Convective Cell Tracking Data during TRACER*, https://doi.org/10.2172/2007152.

Pan, Z., D. Rosenfeld, Y. Zhu, F. Mao, W. Gong, L. Zang, and X. Lu, 2021: Observational quantification of aerosol invigoration for deep convective cloud lifecycle properties based on geostationary satellite. *J. Geophys. Res. Atmos.*, **126**, e2020JD034275, https://doi.org/10.1029/2020JD034275

Pedregosa, F., and Coauthors, 2011: Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, **12**, 2825–2830.

Pehl, J., G. McFarquhar, J. Snyder, and J. Furtado, 2025: Exploring Impacts of Differing Aerosol Environments on Radar Signatures of Deep Convection near Houston, Texas, using a Bulk Statistical Framework. *J Atmos Sci*, **82**, 749–773, https://doi.org/10.1175/JAS-D-24-0008.1.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

Peters, J. M., C. J. Nowotarski, and G. L. Mullendore, 2020: Are Supercells Resistant to Entrainment because of Their Rotation? *J Atmos Sci*, **77**, 1475–1495, https://doi.org/10.1175/JAS-D-19-0316.1.

——, J. P. Mulholland, and D. R. Chavas, 2022: Generalized lapse rate formulas for use in entraining CAPE calculations. *J. Atmos. Sci.*, **79**, 815–836, https://doi.org/10.1175/JAS-D-21-0118.1.

——, Z. J. Lebo, D. R. Chavas, and C.-Y. Su, 2023a: Entrainment makes pollution more likely to weaken deep convective updrafts than invigorate them. Geophys. Res. Lett., 50, e2023GL103314, https://doi.org/10.1029/2023GL103314

——, D. R. Chavas, C. Su, H. Morrison, and B. E. Coffer, 2023b: An analytic formula for entraining CAPE in midlatitude storm environments. *J. Atmos. Sci.*, **80**, 2165–2186, https://doi.org/10.1175/JAS-D-23-0003.1.

Pruppacher, H. R., and J. D. Klett, 2010: *Microphysics of Clouds and Precipitation*. Springer Netherlands, https://doi.org/10.1007/978-0-306-48100-0.

Rao, J. S., 2000: Bootstrapping to Assess and Improve Atmospheric Prediction Models. *Data Min Knowl Discov*, **4**, 29–41, https://doi.org/10.1023/A:1009876615946.

Rasmussen, K. L., and R. A. Houze, 2011: Orogenic Convection in Subtropical South America as Seen by the TRMM Satellite. *Mon Weather Rev*, **139**, 2399–2420, https://doi.org/10.1175/MWR-D-10-05006.1.

Rosenfeld, D., 2000: Suppression of Rain and Snow by Urban and Industrial Air Pollution. *Science (1979)*, **287**, 1793–1796, https://doi.org/10.1126/science.287.5459.1793.

——, U. Lohmann, G. B. Raga, C. D. O'Dowd, M. Kulmala, S. Fuzzi, A. Reissell, and M. O. Andreae, 2008: Flood or Drought: How Do Aerosols Affect Precipitation? *Science (1979)*, **321**, 1309–1313, https://doi.org/10.1126/science.1160606.

Sanderson, B. M., C. Piani, W. J. Ingram, D. A. Stone, and M. R. Allen, 2008: Towards constraining climate sensitivity by linear analysis of feedback patterns in thousands of perturbed-physics GCM simulations. *Clim Dyn*, **30**, 175–190, https://doi.org/10.1007/s00382-007-0280-7.

Saponaro, G., P. Kolmonen, L. Sogacheva, E. Rodriguez, T. Virtanen, and G. de Leeuw, 2017: Estimates of the aerosol indirect effect over the Baltic Sea region derived from 12 years

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

of MODIS observations. *Atmos Chem Phys*, **17**, 3133–3143, https://doi.org/10.5194/acp-17-3133-2017.

Sarmadi, F., Y. Huang, S. T. Siems, and M. J. Manton, 2017: Characteristics of Wintertime Daily Precipitation over the Australian Snowy Mountains. *J Hydrometeorol*, **18**, 2849–2867, https://doi.org/10.1175/JHM-D-17-0072.1.

Schmit, T. J., P. Griffith, M. M. Gunshor, J. M. Daniels, S. J. Goodman, and W. J. Lebair, 2017: A Closer Look at the ABI on the GOES-R Series. *Bull. Amer. Meteor. Soc.*, **98**, 681–698, https://doi.org/10.1175/BAMS-D-15-00230.1.

Schulte, R. M., M. D. Lebsock, J. M. Haynes, and Y. Hu, 2024: A random forest algorithm for the prediction of cloud liquid water content from combined CloudSat–CALIPSO observations. *Atmos Meas Tech*, **17**, 3583–3596, https://doi.org/10.5194/amt-17-3583-2024.

Scovell, R., and H. al-Sakka, 2016: A Point Cloud Method for Retrieval of High-Resolution 3D Gridded Reflectivity from Weather Radar Networks for Air Traffic Management. *J Atmos Ocean Technol*, **33**, 461–479, https://doi.org/10.1175/JTECH-D-15-0051.1.

Sedlacek AJ., 2017: Single-Particle Soot Photometer (SP2) Instrument Handbook. Ed. by Robert Stafford, . *ARM Climate Research Facility.*, DOE/SC-ARM-TR-169.

Seinfeld, J. H., and S. N. Pandis, 1998: *Atmospheric Chemistry and Physics*. 1326 pp., John Wiley, New York.

Shilling, J., and M. Levin, 2021: *Merged size distribution from SMPS and APS (MERGEDSMPSAPS), 2021-10-01 to 2022-09-30, ARM Mobile Facility (HOU), Houston, TX; AMF1 (main site for TRACER) (M1)*. Atmospheric Radiation Measurement (ARM) User Facility, accessed 9 Jun 2024, https://doi.org/10.5439/1871375.

Singh A, and C Kuang, 2024a: Condensation Particle Counter (CPC) Instrument Handbook. Ed. by Robert Stafford. . *ARM Climate Research Facility*, DOE/SC-ARM-TR-145.

——, and ——, 2024b: Scanning Mobility Particle Sizer (SMPS) Instrument Handbook. Ed. by Robert Stafford, . *ARM Climate Research Facility*, DOE/SC-ARM-TR-147.

Song, F., Z. Feng, L. R. Leung, R. A. Houze Jr., J. Wang, J. Hardin, and C. R. Homeyer, 2019: Contrasting Spring and Summer Large-Scale Environments Associated with Mesoscale

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

Convective Systems over the U.S. Great Plains. *J Clim*, **32**, 6749–6767, https://doi.org/10.1175/JCLI-D-18-0839.1.

Song, X., and G. J. Zhang, 2011: Microphysics parameterization for convective clouds in a global climate model: Description and single-column model tests. *J Geophys Res*, **116**, D02201, https://doi.org/10.1029/2010JD014833.

Stein, T. H. M., R. W. Scovell, K. E. Hanley, H. W. Lean, and N. H. Marsden, 2020: The potential use of operational radar network data to evaluate the representation of convective storms in NWP models. *Quarterly Journal of the Royal Meteorological Society*, **146**, 2315–2331, https://doi.org/10.1002/qj.3793.

Stevens, B., 2005: ATMOSPHERIC MOIST CONVECTION. *Annu Rev Earth Planet Sci*, **33**, 605–643, https://doi.org/10.1146/annurev.earth.33.092203.122658.

Székely, G. J., M. L. Rizzo, and N. K. Bakirov, 2007: Measuring and testing dependence by correlation of distances. *The Annals of Statistics*, **35**, https://doi.org/10.1214/009053607000000505.

Tao, W-K., S. Lang, X. Zeng, S. Shige, and Y. Takayabu, 2010: Relating Convective and Stratiform Rain to Latent Heating. *J Clim*, **23**, 1874–1893, https://doi.org/10.1175/2009JCLI3278.1.

Tao, W-K., J. Chen, Z. Li, C. Wang, and C. Zhang, 2012: Impact of aerosols on convective clouds and precipitation. *Reviews of Geophysics*, **50**, https://doi.org/10.1029/2011RG000369.Taszarek, M., N. Pilguj, J. T. Allen, V. Gensini, H. E. Brooks, and P. Szuster, 2021: Comparison of Convective Parameters Derived from ERA5 and MERRA-2 with Rawinsonde Data over Europe and North America. *J. Climate*, **34**, 3211–3237, https://doi.org/10.1175/JCLI-D-20-0484.1.

Taszarek, M., B. Czernecki, and P. Szuster, 2023: thundeR - a rawinsonde package for processing convective parameters and visualizing atmospheric profiles, https://doi.org/10.5194/ecss2023-28.

Tian, P., and Coauthors, 2024: Efficient droplet activation of ambient black carbon particles in a suburban environment. *Atmos Chem Phys*, **24**, 5149–5164, https://doi.org/10.5194/acp-24-5149-2024.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

Trapp, R. J., 2013: *Mesoscale-Convective Processes in the Atmosphere*. Cambridge University Press, https://doi.org/10.1017/CBO9781139047241.

——, N. S. Diffenbaugh, and A. Gluhovsky, 2009: Transient response of severe thunderstorm forcing to elevated greenhouse gas concentrations. *Geophys Res Lett*, **36**, https://doi.org/10.1029/2008GL036203.

Truong, S. C. H., Y. Huang, F. Lang, M. Messmer, I. Simmonds, S. T. Siems, and M. J. Manton, 2020: A Climatology of the Marine Atmospheric Boundary Layer Over the Southern Ocean From Four Field Campaigns During 2016–2018. *Journal of Geophysical Research: Atmospheres*, **125**, https://doi.org/10.1029/2020JD033214.

Tuftedal, K. S., B. P. Treserras, M. Oue, and P. Kollias, 2024: Shallow- and deep-convection characteristics in the greater Houston, Texas, area using cell tracking methodology. *Atmos Chem Phys*, **24**, 5637–5657, https://doi.org/10.5194/acp-24-5637-2024.

Uin, J., 2024: Ultra-High-Sensitivity Aerosol Spectrometer Instrument Handbook. *U.S. Department of Energy, Atmospheric Radiation Measurement user facility*.

Uin, J., J. Gasparik, and E. Lewis, 2024: Artifacts in the aerosol size distribution measured by the Ultra-High-Sensitivity Aerosol Spectrometer (UHSAS). *U.S. Department of Energy, Atmospheric Radiation Measurement User Facility*, Richland, WA, DOE/SC-ARM-TR-304.——, G. Senum, A. Koontz, C. Flynn, C. Salwen, and C. Hayes, 2021: Ultra-High Sensitivity Aerosol Spectrometer (AOSUHSAS), 2021-10-01 to 2022-10-01, ARM Mobile Facility (HOU), Houston, TX; AMF1 (main site for TRACER) (M1). *Atmospheric Radiation Measurement (ARM) User Facility*, accessed 1 Jun 2024, https://doi.org/10.5439/1409033.

Uin, J., and Coauthors, 2019: Atmospheric Radiation Measurement (ARM) Aerosol Observing Systems (AOS) for Surface-Based In Situ Atmospheric Aerosol and Trace Gas Measurements. *J Atmos Ocean Technol*, **36**, 2429–2447, https://doi.org/10.1175/JTECH-D-19-0077.1.

Varble, A., 2018: Erroneous Attribution of Deep Convective Invigoration to Aerosol Concentration. *J Atmos Sci*, **75**, 1351–1368, https://doi.org/10.1175/JAS-D-17-0217.1.

Varble, A. C., A. L. Igel, H. Morrison, W. W. Grabowski, and Z. J. Lebo, 2023: Opinion: A critical evaluation of the evidence for aerosol invigoration of deep convection. *Atmos Chem Phys*, **23**, 13791–13808, https://doi.org/10.5194/acp-23-13791-2023.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

Varble, A. C., Z. Feng, J. N. Marquis, Z. Zhang, A. Geiss, J. C. Hardin, and E. Jo, 2024: Updraft Width Modulates Ambient Atmospheric Controls on Convective Cloud Depth. *Journal of Geophysical Research: Atmospheres*, **129**, https://doi.org/10.1029/2024JD041769.

Vovk, T., M. Kryza, and M. Werner, 2024: Using random forest to improve EMEP4PL model estimates of daily PM2.5 in Poland. *Atmos Environ*, **332**, 120615, https://doi.org/10.1016/j.atmosenv.2024.120615.

Wang, D., R. Kobrosly, T. Zhang, T. Subba, S. van den Heever, S. Gupta, and M. Jensen, 2024: Causal Analysis of Aerosol Impacts on Isolated Deep Convection: Findings from TRACER, https://doi.org/10.5194/egusphere-2024-2436.

Watson TB, 2017: Aerosol Chemical Speciation Monitor (ACSM) Instrument Handbook. Ed. by Robert Stafford. *ARM Climate Research Facility*, **DOE/SC-ARM-TR-196.**

Weisman, M. L., and J. B. Klemp, 1982: The Dependence of Numerically Simulated Convective Storms on Vertical Wind Shear and Buoyancy. *Mon Weather Rev*, **110**, 504–520, https://doi.org/10.1175/1520-0493(1982)110<0504:TDONSC>2.0.CO;2.

Wilcoxon, F., 1945: Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, **1**, 80, https://doi.org/10.2307/3001968.

Wilhelm, J., K. Wapler, U. Blahak, R. Potthast, and M. Kunz, 2023: Statistical relevance of meteorological ambient conditions and cell attributes for nowcasting the life cycle of convective storms. *Quarterly Journal of the Royal Meteorological Society*, **149**, 2252–2280, https://doi.org/10.1002/qj.4505.

Xu, K.-M., 2006: Using the Bootstrap Method for a Statistical Significance Test of Differences between Summary Histograms. *Mon Weather Rev*, **134**, 1442–1453, https://doi.org/10.1175/MWR3133.1.

Yin, J., Z. Pan, F. Mao, D. Rosenfeld, L. Zang, J. Chen, and J. Gong, 2024: Large effects of fine and coarse aerosols on tropical deep convective systems throughout their lifecycle. *NPJ Clim Atmos Sci*, **7**, 195, https://doi.org/10.1038/s41612-024-00739-6.

Zang, L., D. Rosenfeld, Z. Pan, F. Mao, Y. Zhu, X. Lu, and W. Gong, 2023: Observing Aerosol Primary Convective Invigoration and Its Meteorological Feedback. *Geophys Res Lett*, **50**, https://doi.org/10.1029/2023GL104151.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.

Zawadowicz, M., J. Howie, C. Hayes, M. Allain, C. Salwen, and B. Behrens, 2021: Aerosol Chemical Speciation Monitor (AOSACSM), 2021-10-01 to 2022-10-01, ARM Mobile Facility (HOU), Houston, TX; AMF1 (main site for TRACER) (M1). Atmospheric Radiation Measurement (ARM) User Facility, accessed 9 Jun 2024, https://doi.org/10.5439/1762267.

Zöbisch, I., C. Forster, T. Zinner, L. Bugliaro, A. Tafferner, and K. Wapler, 2020: Characteristics of deep moist convection over Germany in multi-source data. *Meteorologische Zeitschrift*, **29**, 393–407, https://doi.org/10.1127/metz/2020/1011.

Accepted for publication in *Journal of the Atmospheric Sciences*. DOI 10.1175/JAS-D-25-0130.1.