Author/s:
Short, E;Lane, TP

Title:
Objectively Assessing Characteristics of Mesoscale Convective Organization in an Operational Convection-Permitting Model

Date:
2023-10-01

Citation:
Short, E. & Lane, T. P. (2023). Objectively Assessing Characteristics of Mesoscale Convective Organization in an Operational Convection-Permitting Model. Monthly Weather Review, 151 (10), pp.2819-2841. https://doi.org/10.1175/MWR-D-23-0033.1.

Persistent Link:
https://hdl.handle.net/11343/340244

Objectively Assessing Characteristics of Mesoscale Convective Organization in an Operational Convection Permitting Model

Ewan Short[a] , Todd P. Lane[a]

[a] *School of Geography, Earth & Atmospheric Sciences, and ARC Centre of Excellence for Climate Extremes, The University of Melbourne, Melbourne, Victoria, Australia.*

Corresponding author: Ewan Short, shorte1@student.unimelb.edu.au

ABSTRACT: The realism of convective organization in operational convection permitting model simulations is objectively assessed, with a particular focus on the mesoscale aspects, such as convective mode. A tracking and classification algorithm is applied to observed radar reflectivity and simulated radar reflectivity from the operational ACCESS-C convection permitting forecast domain over northern Australia between October 2020 and May 2022, and characteristics of real and simulated convective organization compared. Mesoscale convective systems from the operational forecast model are approximately twice as likely to be oriented parallel to the ambient wind and ambient wind shear than those observed by radar, indicating a bias toward the "training line" systems typically associated with more extreme rainfall. During highly humid active monsoon conditions, simulated convective systems have larger ground-relative speeds than systems observed in radar. Although there is less than 5% difference between the ratios of simulated and observed trailing, leading and parallel stratiform system observations, significant differences exist in other wind-shear-based classifications. For instance, in absolute terms, simulated systems are $10-35\%$ less likely to be up-shear tilted, and 15-30% less likely to be down-shear propagating than observed systems, suggesting errors in simulated cold pool characteristics.

SIGNIFICANCE STATEMENT: Remarkable progress has been made simulating thunderstorms in operational weather forecasting computer models. While some details of individual storm-clouds may be unrealistic, how these storm-clouds self-organise, i.e. cluster and regenerate, can be explicitly simulated, with this organisation often appearing realistic. However, assessing the realism of this organisation in an objective, systematic way has proven challenging. Here we assess organised convection in Australia's current high resolution weather prediction model. In some respects, simulated storm clouds organize realistically. High-altitude icy cloud mostly trails behind groups of storm clouds in both simulations and reality. In other respects organization is unrealistic. Simulated storm clouds are twice as likely to orient along the mean wind direction than in reality, likely contributing to extreme rainfall biases.

1. Introduction

Towering cumulonimbus convection produces hazardous weather, including heavy rainfall, lightning, hail, and wind extremes. Excluding mid-latitude supercells, this convection occurs via intense, narrow updrafts with median diameters of $\approx 2\,\mathrm{km}$ (e.g. Byers and Braham 1948; LeMone and Zipser 1980; Giangrande et al. 2013; Wang et al. 2020). Within the kilometre scale updraft of a developing cumulonimbus cloud, aircraft observations show there is smaller scale variability over horizontal distances of $\approx 100\,\mathrm{m}$ produced by individual thermals and their associated toroidal circulations (Damiani et al. 2006, 2008). These 100 m scale flows not only produce the "parapets" that give cumuliform clouds their striking appearance, but also mix cloud air with non-cloud air, modifying the intensity of the convection (Bryan et al. 2003).

Cumulonimbi often group together into larger mesoscale convective systems (MCSs), which can span thousands of square kilometres and last for many hours, much longer than the lifetime of the individual cumulonimbi that comprise them. Grouping of cumulonimbi may occur through processes like the land-sea breeze, or by orography, but can also occur through mesoscale "self-organization", i.e. through processes inherent to the convection itself. Theories of self-organisation typically invoke cold pools (e.g. Seitter and Kuo 1983; Emanuel 1986; Rotunno et al. 1988; Weisman et al. 1988; Weisman 1993; Moncrieff and So 1989; Moncrieff and Liu 1999; Weisman and Rotunno 2004; Houze 2004; Coniglio et al. 2006, 2012), internal gravity waves (e.g. Mapes et al. 2003; Kawashima 2003; Lane and Zhang 2011), and hydrostatic steady-flow ideas (e.g.

Moncrieff and Green 1972; Moncrieff and Miller 1976; Moncrieff 1981, 1992), with diverse mechanisms often thought to act in concert (e.g. Lafore and Moncrieff 1989; Moncrieff and Lane 2015). These processes can group cumulonimbi into lines, bows, clusters, and more complex combined structures (e.g. Keene and Schumacher 2013).

Early operational weather-forecasting models used grid-spacings too large to explicitly simulate convection, which was instead parametrized. Over the last decade it has become computationally feasible to run operational weather-forecasting models with horizontal grid spacings $\Delta x \approx 1\,\text{km}$, in which cumulonimbus-like features are explicitly produced. However, if a numerical simulation can only explicitly resolve spatial scales at $\approx 7\Delta x$ (Skamarock 2004), then $\Delta x = 1\,\text{km}$ is too large to explicitly simulate realistic individual cumulonimbus as, for instance, cumulus updraft widths will be too large. Indeed, Wang et al. (2020) showed that even at $\Delta x = 250\,\text{m}$, simulated updraft widths larger than 3 km were more than twice as frequent in numerical simulations than observations. Furthermore, 100 m scale thermals, and their associated entrainment, certainly can not be simulated with a $\Delta x = 1\,\text{km}$ model. Thus, $\Delta x \approx 1\,\text{km}$ simulations are convection "permitting" rather than convection-resolving, as cumulonimbus processes are only explicitly simulated in an emblematic way (Bryan et al. 2003; Emanuel 2020).

Nevertheless, convection-permitting simulations may still be of significant operational value (Lean et al. 2008; Roberts and Lean 2008). Because cumulonimbi often form MCSs, and MCSs are typically associated with 10-100 km mesoscale mean flows which a $\Delta x = 1\,\text{km}$ model can simulate explicitly, it is plausible that a model at this gridlength can explicitly resolve core MCS flows in a realistic way (Weisman et al. 1997), despite the component cumulonimbi themselves being necessarily unrealistic. Because MCSs represent the localization and regeneration of cumulonimbi, and therefore the concentration and accumulation of their hazards (e.g. Schumacher and Johnson 2006), assessing how realistically convective permitting models simulate MCSs is an operationally pressing goal.

Object based tracking algorithms provide one strategy for assessing the realism of MCSs in operational models. Tracking algorithms can be used to identify MCSs in both observed and simulated radar reflectivity data, with descriptive statistics from each then compared. This approach circumvents the small-scale predictability limits which hamper the direct comparison of gridded datasets.

Previous studies have used an object-based approach to assess the realism of smaller scale convective features in convective permitting models. Caine et al. (2013) applied the Thunderstorm Identification Tracking Analysis and Nowcasting (TITAN) algorithm to 5 days of $\Delta x = 1.25\,\text{km}$ simulation and radar data encompassing the Darwin region of Northern Australia, identifying different model biases in the horizontal and vertical extents of the convective objects associated with different reflectivity thresholds. In general, simulated convection reached higher altitudes than observed.

Hanley et al. (2015) used rain rates and area thresholds to identify convective objects in a 15 hour, $\Delta x = 1.5\,\text{km}$ United Kingdom Met Office Unified Model (UM) simulation of the southern United Kingdom, and in radar composites of the same time period and region. They found that simulated cumulonimbi were too intense, and too far apart, with insufficient light rain, and too few small cumulonimbi. Reducing Δx improved some convective characteristics, but worsened others. Similarly, Stein et al. (2020) analysed convective objects in four days of UM and composite radar data encompassing the southern United Kingdom, finding that when $\Delta x = 1.5\,\text{km}$ cumulonimbi were too wide, and when $\Delta x = 200\,\text{m}$ they were too narrow.

Bergemann et al. (2022) studied an ensemble of nested UM simulations spanning 8 days, encompassing the Tiwi Islands north of Darwin Airport in Australia's Northern Territory (see Fig. 1). They found that when $\Delta x = 1.33\,\text{km}$, simulated thunderstorms mostly comprised individual, disorganised cumulonimbi. Furthermore, observed thunderstorms typically formed over the north-east of Melville island, the larger of the Tiwi islands, propagating south-west to the southern flank of the smaller Bathurst island, whereas simulated storms typically stalled over Melville island, resulting in significant precipitation biases across these islands. The rainfall from the simulated storms also peaked a few hours too early. These biases did not markedly improve with $\Delta x = 0.44\,\text{km}$, as reducing the resolution introduced significant changes to the behaviour of cold-pools and the microphysics, making the simulated storms even more intense.

Collectively, the above studies indicate that choosing an operationally feasible $\Delta x \gg 10\,\text{m}$ is even more difficult than argued by Bryan et al. (2003). Unrealistic convective motion introduces a myriad of complex feedbacks, particularly with microphysics schemes, that degrade core features of small simulated storms. As argued above, it is nevertheless possible that convection permitting models at operationally feasible grid-spacings can still explicitly and realistically simulate larger

mesoscale features, and in the present study we also employ object based methods to assess this. We use an MCS tracking and analysis algorithm (Short et al. 2023) to compare the statistics of organizational characteristics of larger MCS objects, such as whether the MCS is trailing or leading stratiform (Parker and Johnson 2000), between radar observations and operational model data. Such characteristics collectively determine the organizational "mode". While the studies cited above evaluated spatial and temporal characteristics of organized systems (e.g. size, shape and depth) in this study we focus on evaluating the organizational mode. This represents a more stringent convection-permitting model test, as the organisational mode is determined by multiple physical processes and interactions, and assesses the ability of the model to form and maintain mesoscale systems in a realistic way.

Recent developments have made a systematic comparison of Australian operational convection permitting model simulations with radar observations possible. In 2020 the Australian Bureau of Meteorology began archiving their operational convection-permitting model (Bureau of Meteorology 2022) and radar datasets (Soderholm and Protat 2022) on public databases, from which characteristics of simulated and radar observed convective organization can be compared. In this study we focus on the Darwin region of Northern Australia, as a climatology of mesoscale organized convection exists there (Short et al. 2023).

To systematically compare these radar and model datasets, we use the "Mesoscale TINT is not TITAN" (MINT) algorithm. MINT is described in detail by Short et al. (2023): in essence the algorithm identifies convective cloud objects at low levels using the Steiner et al. (1995) scheme, stratiform cloud objects at upper levels using a 15 dBZ reflectivity threshold, and pairs the two into larger MCS objects if the stratiform cloud lies over the convective cloud. MCS objects are matched and tracked between subsequent scans by minimising a cost function of distance and size, as in TINT and TITAN (Dixon and Wiener 1993; Raut et al. 2021). MINT uses cross-correlation vectors derived from successive radar scans to estimate system ground-relative velocities, comparing these with the ambient-wind vectors, and the stratiform offset vector, i.e. the displacement vector between the convective and stratiform cloud regions, to perform various organizational classifications, as depicted in Fig. 2. To ensure stratiform offset, velocity vectors, and resultant classifications are accurate, MINT quality controls MCS observations in multiple ways. For instance, to ensure velocities and stratiform offsets are accurate, MCS observations must lie almost entirely within the

given radar domain. In the study of (Short et al. 2023), the radar domain comprised a circle of radius 150 km, necessarily restricting their analysis to MCS objects usually smaller than typical MCS size threshold definitions (Houze 2004; Hitchcock et al. 2021).

By applying MINT to a 15-year research radar record from northern Australia, Short et al. (2023) found that $65-80\%$ of MCS observations were consistent with the classical front-fed, trailing stratiform, up-shear tilted, down-shear propagating model, at least when all the relevant classfications could be performed unambiguously. The dominance of observations consistent with the classical model, despite the identified systems often being smaller than conventional MCS size thresholds, suggested the classical model mechanisms of convective organisation may also apply to systems much smaller than previously thought. During highly humid active monsoon periods, in which cold pools are potentially weaker, classical model MCS observations dropped to $< 10\%$ of the total, with non-classical rear-fed trailing-stratiform observations now accounting for $10-40\%$ of the total.

The rest of this paper is organized as follows. Section 2 describes the datasets and methodology, and section 3 presents results. Sections 4 and 5 provide a discussion and conclusion respectively.

2. Approach

In this study we compare characteristics of convective organization in the Australian Bureau of Meteorology's operational mesoscale convection permitting configuration of the Australian Community Climate and Earth System Simulator (ACCESS), with those observed by operational weather radar. We focus on the Darwin region of Australia's Northern Territory, as depicted in Fig. 1 a), over the two northern Australian wet seasons encompassing 01/10/2020 to 01/05/2021, and 01/10/2021 to 01/05/2022. Convective systems are tracked and organizational characteristics diagnosed using a new configuration of the "Mesoscale TINT is not TITAN" (MINT) algorithm Short et al. (2023). Whereas the original MINT configuration used fully 3D radar reflectivity data, only 1 km altitude and column maximum simulated reflectivity fields are available for ACCESS, requiring a new MINT configuration to accommodate this: further details are provided in section 2 b.

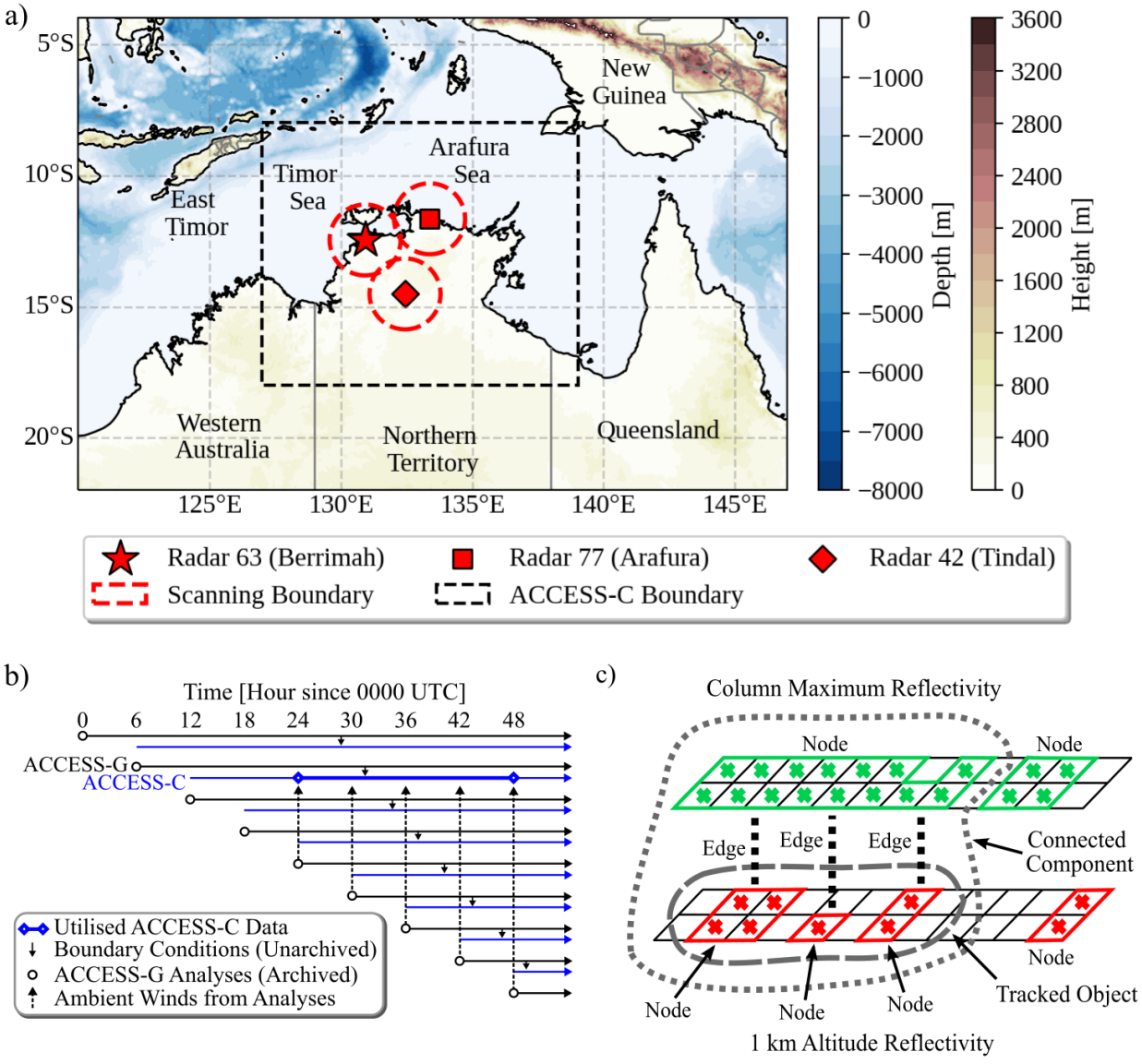


FIG. 1. Methodology schematics depicting, a), the Darwin ACCESS-C domain and locations of select operational radar and airport weather stations, b) how different ACCESS datasets are used, and c), the methodology for identifying convective systems in reflectivity datasets. See text for further details.

a. Datasets

In 2020 the Australian Bureau of Meterology updated its nested ACCESS Australian Parallel Suite (APS) weather forecasting system from version 2 to version 3 (APS3). The APS3 system comprises a global atmospheric model (ACCESS-G) with a horizontal grid spacing $\Delta x \approx 12\,\text{km}$ (Bureau of Meteorology 2019). Within ACCESS-G are nested $\Delta x \approx 1.5\,\text{km}$ convection permitting models

over major Australian cities (ACCESS-C). Figure 1 a) shows the $\approx 1305 \times 1110$ km ACCESS-C domain encompassing Darwin, the capital of Australia's Northern Territory.

ACCESS-G analyses are conducted, and forecasts initialized, at 0000, 0600, 1200 and 1800 UTC each day, with a forecast range of at least 84 hours in each case, i.e. forecast lead-times from +0, to at least +84 hours. The ACCESS-C models are coupled to ACCESS-G via hourly cycling data-assimilation. ACCESS-C analyses are conducted, and forecasts initialized, at each hour UTC, with the 0000, 0600, 1200 and 1800 UTC initialized forecast ranges at least 36 hours in each case: these forecasts receive initial and boundary conditions from the ACCESS-G forecast initialized 6 hours earlier (Rennie et al. 2022; Bush et al. 2020) as depicted in Fig. 1 b).

ACCESS-G and ACCESS-C use version 10.6 of the United Kingdom Met Office Unified Model (UM) to simulate atmospheric processes (Walters et al. 2019; Bush et al. 2020; Rennie et al. 2022). The ACCESS-C nested models are configured differently for tropical and mid-latitude regions: the Darwin model uses the tropical configuration (Rennie et al. 2022; Bush et al. 2020; Dipankar et al. 2020). Simulated radar reflectivities are calculated from ACCESS-C microphysical variables assuming the Rayleigh scattering limit (McBeath et al. 2014; Stein et al. 2014; Rennie et al. 2022), and the 1 km altitude and column maximum reflectivities are archived every 10 minutes.

We consider the two North Australian wet seasons encompassing 01/10/2020 to 01/05/2021, and 01/10/2021 to 01/05/2022. We use the ACCESS-C simulated reflectivities, and the horizontal winds from the ACCESS-G analyses and forecasts, and from the ACCESS-C forecasts (Bureau of Meteorology 2022). For a given day we consider the ACCESS-C forecast initialized at 1200 UTC on the previous day and use the simulated reflectivities at lead times of 12 to 36 hours, as depicted in Fig. 1 b).

We also consider operational radar data over the same time period from the C-Band Berrimah, Arafura and Tindal radars, shown in Fig. 1 a). These radars have scanning patterns encompassing elevations from 0.5° to 32°, repeating in 5 minute intervals from each hour UTC. Because ACCESS-C simulated reflectivities are archived in ten minute intervals from each hour UTC, and for consistency with previous work (Short et al. 2023), we only use every alternate operational radar scan. For consistency with Short et al. (2023), reflectivities are gridded onto $2.5 \times 2.5 \times 0.5$ km Cartesian grids using the Py-ART library (Helmus and Collis 2016), masking data at hori-

zontal distances of more than 150 km from each radar. The ACCESS-G winds and ACCESS-C reflectivities are linearly interpolated onto these same grids.

We also use horizontal winds from the 5th European Center for Medium Range Forecasting Reanalysis (ERA5), publicly available in Australia from the 1st of January 1979 to the present (European Centre for Medium-Range Weather Forecasts 2020). These winds are provided hourly on a 31 by 31 km horizontal grid, with 137 vertical levels (Hersbach et al. 2020), which we again interpolate onto the $2.5 \times 2.5 \times 0.5$ km radar grids.

As noted above, only the 1 km and column maximum ACCESS-C simulated reflectivities are archived, and we cannot reconstruct the complete 3 dimensional reflectivities ourselves as not all ACCESS-C microphysical fields are archived. This limited ACCESS-C data availability necessitated modification of the MINT algorithm configuration used by Short et al. (2023), which used the full 3D radar reflectivity volume. To assess the new MINT configuration we consider reflectivities from the C-band polarimetric radar (CPOL) located near Darwin, operational between 1998 to 2016 (Louf and Protat 2020; Louf et al. 2019), from which convective organization characteristics have previously been studied using MINT (Short et al. 2023). Also, when interpreting our ACCESS-C results we consider radiosonde data from the Darwin Airport and Gove Airport stations (University of Wyoming 2022) depicted in Fig. 1 a).

To demarcate synoptic conditions Short et al. (2023) used the Pope et al. (2009) regimes, which between 1998 and 2017 were provided by the Australian Bureau of Meteorology. Short et al. (2023) defined the "weak monsoon" as the combination of the dry-east, east, shallow-west, and moist-east Pope regimes (regimes 1, 3, 4 and 5), with the active monsoon category given by the deep-west Pope regime (regime 2). In the present study we distinguish synoptic environments using a simple binary criteria applied to ERA5 data at the grid-cell closest to Darwin Airport. We consider the daily mean wind between 800 and 600 hPa, demarcating "active monsoon" and "weak monsoon" days when this mean wind is $> 5\,\mathrm{m\,s^{-1}}$ and $\leq 5\,\mathrm{m\,s^{-1}}$ respectively. This definition is analogous to that of Drosdowsky (1996), and matches that of Short et al. (2023) for 91% of the days between 1998 and 2017.

b. Methodology

In this study we apply the MINT algorithm (Short et al. 2023) to the ACCESS and radar datasets described above. Whereas previous storm tracking algorithms like TITAN and TINT (nominally "TINT is not TITAN") were designed to identify and track disjoint convective cloud regions, MINT (Mesoscale TINT) uses upper level stratiform cloud to group multiple disjoint convective cloud regions into larger MCSs, with the larger system tracked as a single object.

MINT is described in detail by Short et al. (2023). The algorithm acts on gridded radar reflectivity datasets in Cartesian coordinates, and estimates of the ambient winds provided by reanalysis or model datasets. Core outputs are the convective and stratiform cloud regions, which must be at least 80 and 800 km^2 respectively, and the stratiform offset vector, which measures the displacement between these regions, and from which various MCS classifications can be performed (see Fig. 2). Short et al. (2023) found that MINT correctly identifies organised convective systems, and produces classifications that are typically very robust across diverse sensitivity tests. Because there is minimal overlap between the three radar domains considered in this study we apply MINT independently to each domain, pooling results when reporting statistics.

Unfiltered MINT output comprises all MCS observations, where an "observation" consists of a single system in a single scan. Most classification results consist of counting different types of observations thus defined. Imposing the additional requirement that systems last for at least 30 min, we obtain what we term the "raw sample". For example, two systems tracked across four radar scans would contribute eight "observations" to the raw sample.

In order to diagnose and classify observations the raw sample must be restricted as follows. The magnitude of a given system's ground-relative velocity must be $\geq 5\,\mathrm{m\,s^{-1}}$, the magnitude of its stratiform offset must be $\geq 10\,\mathrm{km}$, the ratio of its reflectivity pixels on the radar scanning boundary to its total pixels must be $\leq 10^{-3}$ for both its convective and stratiform cloud regions, and these conditions must all be met for at least 30 min contiguously: we call the observations satisfying these criteria the "restricted sample". Considering again the two hypothetical systems tracked across four radar scans: if the first system has too many pixels on the radar scanning boundary during the first scan, and the second systems's velocity drops below $5\,\mathrm{m\,s^{-1}}$ on its final two scans, only five of the eight raw sample observations would be included in the restricted sample.

The boundary intersection requirement has the largest impact on the restricted samples, removing $\approx 70\%$ of the observations from both the radar and ACCESS-C raw samples. As discussed by Short et al. (2023), these restrictions are imposed to quality control the resulting measurements and classifications: for instance, velocities less than $5\,\mathrm{m\,s^{-1}}$ were found to have inaccurate directions, and systems with small stratiform offsets cannot be classified due to errors associated with radar scanning patterns. Similarly, in the classifications depicted in Fig. 2 a), c) and d), observations are classified as ambiguous if the stratiform-offset vector forms an angle of $< \theta_e$ with a quadrant boundary, as illustrated in Fig. 2, with these ambiguous cases ignored when reporting statistics. In the classical MCS model (Houze 2004), sufficiently large stratiform offsets are necessary to prevent the cold pools associated with stratiform cloud from blocking low-level inflow to new cumulonimbi: in this sense the restricted sample observations are on average more organized than those of the raw sample.

Because only the 1 km altitude and column maximum ACCESS-C simulated reflectivities are archived, we use the former to identify convective objects, and the latter to identify stratiform objects, grouping multiple convective objects into a single system if they all lie under the same region of stratiform cloud, as depicted in Fig. 1 c). MINT then works on the two dimensional objects identified in each of these datasets. A potential limitation of this new configuration is that the detected stratiform cloud does not necessarily occur at upper levels, and is not necessarily connected to the convective cells through mid-tropospheric cloud, as in the original configuration of Short et al. (2023). Despite this, a comparison between the original and new MINT configuration reveals the various classification results are essentially identical, varying by at most 5%, with further details provided in the online supplement.

MINT performs the Parker and Johnson (2000) classification depicted in Fig. 2 a) solely from reflectivity data. When paired with estimates of the undisturbed ambient winds, MINT also performs four other classifications relevant to convective organization. For systems identified in the operational radar datasets, we use ERA5 for the ambient winds. For systems identified in ACCESS-C, ideally we would use the winds from the same ACCESS-G forecast providing the ACCESS-C boundary conditions. MINT requires wind estimates throughout the lower troposphere, but unfortunately only near-surface ACCESS-G forecast winds are archived. Instead, we therefore use the ACCESS-G analyses wind data, provided at all model levels at 6 hourly intervals each day, as de-

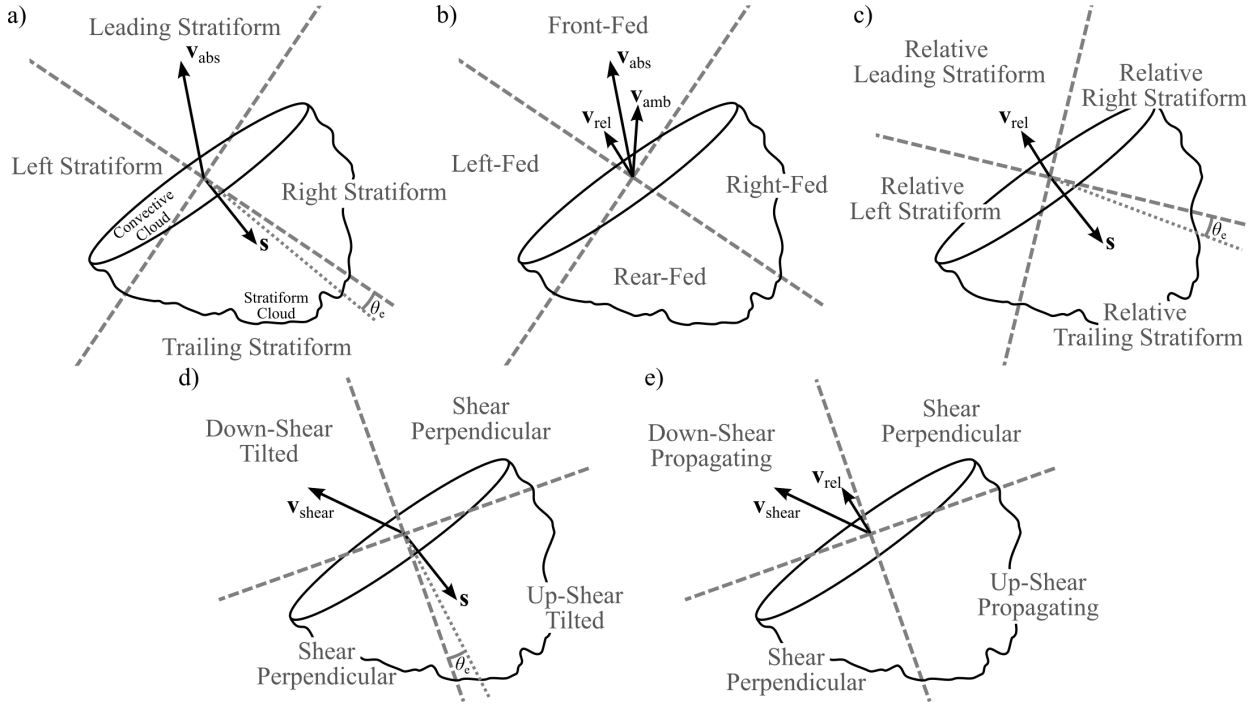


FIG. 2. Diagrams adapted from Short et al. (2023) depicting MINT's different methods for classifying convective systems based on, a), the stratiform offset $\mathbf{s}$ and the system's ground-relative velocity $\mathbf{v}_{abs}$, b), the system's ground-relative velocity and flow-relative velocity $\mathbf{v}_{rel} = \mathbf{v}_{abs} - \mathbf{v}_{amb}$, where $\mathbf{v}_{amb}$ is the ambient low-level wind, c), the system's flow-relative velocity, stratiform offset, d) the stratiform offset and the low-level wind shear $\mathbf{v}_{shear}$ and e) the system's flow-relative velocity and the low-level wind shear. The dashed grey lines demarcate the quadrants used in each classification. Note the first quadrant is centred on $\mathbf{v}_{abs}$ in a) and b), on $\mathbf{v}_{rel}$ in c), and on $\mathbf{v}_{shear}$ in d) and e). Also, in a), c) and d), the classification is deemed ambiguous if the stratiform offset vector forms an angle $< \theta_e$ with any quadrant boundary. We refer to the classifications depicted by a) to e) as the stratiform offset, inflow, relative stratiform offset, tilt and propagation classifications respectively. See section 2 and Short et al. (2023) for further details.

picted in Fig. 1 b). This approach assumes the ACCESS-G forecasts are consistent with subsequent ACCESS-G analyses at lead times of 18-42 hours throughout the lower troposphere. Because only near surface ACCESS-G forecast winds are archived, we cannot check this assumption. Instead, we develop intuition for possible differences by comparing the 10 m altitude ACCESS-G analyses and forecast winds. In section 3 we also consider the sensitivity of our results to switching the ambient winds used with each reflectivity dataset, i.e. using ERA5 winds with ACCESS-C reflectivities and ACCESS-G winds with radar reflectivities.

While Short et al. (2023) used hourly ERA5 wind data temporally interpolated to 10 minute radar scanning intervals, in the present study the ACCESS-G analyses are only performed every six hours, so for consistency we restrict the ERA5 winds to the same 6 hourly intervals, then temporally interpolate both onto the 10 minute time intervals of the radar and ACCESS-C reflectivities. We consider winds between 500 m and 3 km inclusive, taking the mean as an estimate of the ambient wind, and the difference between the 3 km and 500 m winds as a low-level shear vector. In section 3 we consider how our results change when winds over different altitudes are used.

From the mean ambient wind and shear vectors, the additional classifications depicted in Figs. 2 b) to e), and described in Short et al. (2023) can be performed. We refer to the Fig. 2 a) to e) classifications as the stratiform offset, inflow, relative stratiform offset, tilt, and propagation classifications respectively. To perform these classifications we must further restrict the restricted sample: the magnitude of the flow-relative velocity must be $\geq 2\,\mathrm{m\,s^{-1}}$ for the inflow, relative stratiform offset and propagation classifications, and the magnitude of the low-level shear vector must be $\geq 2\,\mathrm{m\,s^{-1}}$ for the tilt and propagation classifications. Expressions like "trailing stratiform" and "up-shear tilted" appearing in Fig. 2 are defined by Short et al. (2023), but their meaning should be clear from Fig. 2 and its caption. As discussed by Short et al. (2023), relative trailing stratiform, up-shear tilted, down-shear propagating observations, of which front-fed, trailing stratiform are the canonical example, are consistent with the classical MCS model (Houze 2004).

Because the ACCESS-C data considered each day represents a different model run, system identification and tracking must be reset at 0000 UTC ($\approx$ 0900 LST) each day to avoid spurious matching of systems between days, as convective-scale predictability limits imply distinct runs produce distinct systems in different locations, at different times. For consistency we also do this for systems identified in the radar data. The thresholds and parameter choices given above, with $\theta_e = 10^\circ$ (see Fig. 2), form the "baseline" MINT configuration of this study, which mirrors the baseline configuration used by Short et al. (2023). The modified configurations for the various sensitivity tests are summarised in Table 1, noting test "C" represents a combined test, as even if results are robust when parameters are varied individually, they may not be when parameters are varied jointly.

Test Name	Description
Base*	Baseline MINT run. See text for details.
SA	Swapped ambient wind datasets, i.e. pairing ACCESS-G ambient winds with the operational radar reflectivities, and ERA5 ambient winds with ACCESS-C reflectivities.
W2*	A shallower low-level ambient wind layer of $[500\,\mathrm{m}, 2500\,\mathrm{m})$.
W4*	A deeper low-level ambient wind layer of $[500\,\mathrm{m}, 4500\,\mathrm{m})$.
S4*	A higher shear magnitude threshold of $4\,\mathrm{m\,s^{-1}}$.
RV4*	A higher flow-relative velocity magnitude threshold of $4\,\mathrm{m\,s^{-1}}$.
CA24	A higher area threshold of $240\,\mathrm{km^2}$ for convective regions.
LT*	Lower stratiform offset threshold of 5 km and a smaller quadrant buffer angle θ_e of 5 degrees (see Fig. 2).
D	Restriction to just daytime observations between 0000 and 1200 UTC (roughly 0900 and 2100 LST) inclusive each day.
C*	A low-level ambient wind layer of $[500\,\mathrm{m}, 2500\,\mathrm{m})$, a higher stratiform offset threshold of 12.5 km, and higher shear and flow-relative velocity magnitude thresholds of $3\,\mathrm{m\,s^{-1}}$.
NS*	A 40 dBZ reflectivity threshold is used in place of the Steiner scheme to identify convective objects in the 1 km reflectivity data.
LR*	Lower reflectivity thresholds of 35 dBZ and 10 dBZ are used to detect convective and stratiform objects in the 1 km and column-maximum reflectively data respectively.
T15*	A larger quadrant buffer angle θ_e of 15 degrees (see Fig. 2).
S15*	A higher stratiform offset threshold of 15 km.
A2*	A higher area threshold of $2000\,\mathrm{km^2}$, which must be satisfied by either or both of the convective or stratiform regions.
B5*	A stricter radar boundary intersection ratio threshold of 5×10^{-4} pixels/pixels.
L50*	The ellipses fitted to convective cloud regions must have a major axis length greater than 50 km, and a major to minor axis ratio of at least 2:1.
L25*	The ellipses fitted to convective regions must have a major axis length greater than 25 km, and a major to minor axis ratio of at least 2:1.

TABLE 1. Names and descriptions of the sensitivity tests. Asterisks indicate tests analogous to those performed by Short et al. (2023).

3. Results

This section summarizes the new MINT configuration assessment results, and provides comparisons between the convective organization characteristics of the radar and ACCESS-C datasets.

a. Algorithm Assessment

As a preliminary assessment, we repeated the classification of MCSs in the CPOL radar record performed by Short et al. (2023) with the new MINT configuration, which as discussed in section 2 b, uses only column maximum and 1 km reflectivities. Results were essentially the same: the main difference was a change in sample sizes. All classification results were within 5% of those obtained by Short et al. (2023), for example, ≈ 70 and 75% of restricted sample observations were trailing stratiform with the original and new configurations respectively. Further details are provided in the online supplement.

We also compared the 10 m winds from the ACCESS-G analyses and forecasts, to develop intuition for when and where the low-level winds may differ. Differences were greatest over ocean near coastlines at 1800 UTC, corresponding to ACCESS-G forecasts at 36-hour lead times, consistent with a biased ACCESS-G land-breeze (Jucker et al. 2020). This difference motivated sensitivity test D, where we consider only those observations occurring between 0000 and 1200 UTC. Further details on the ambient wind comparison are provided in the online supplement.

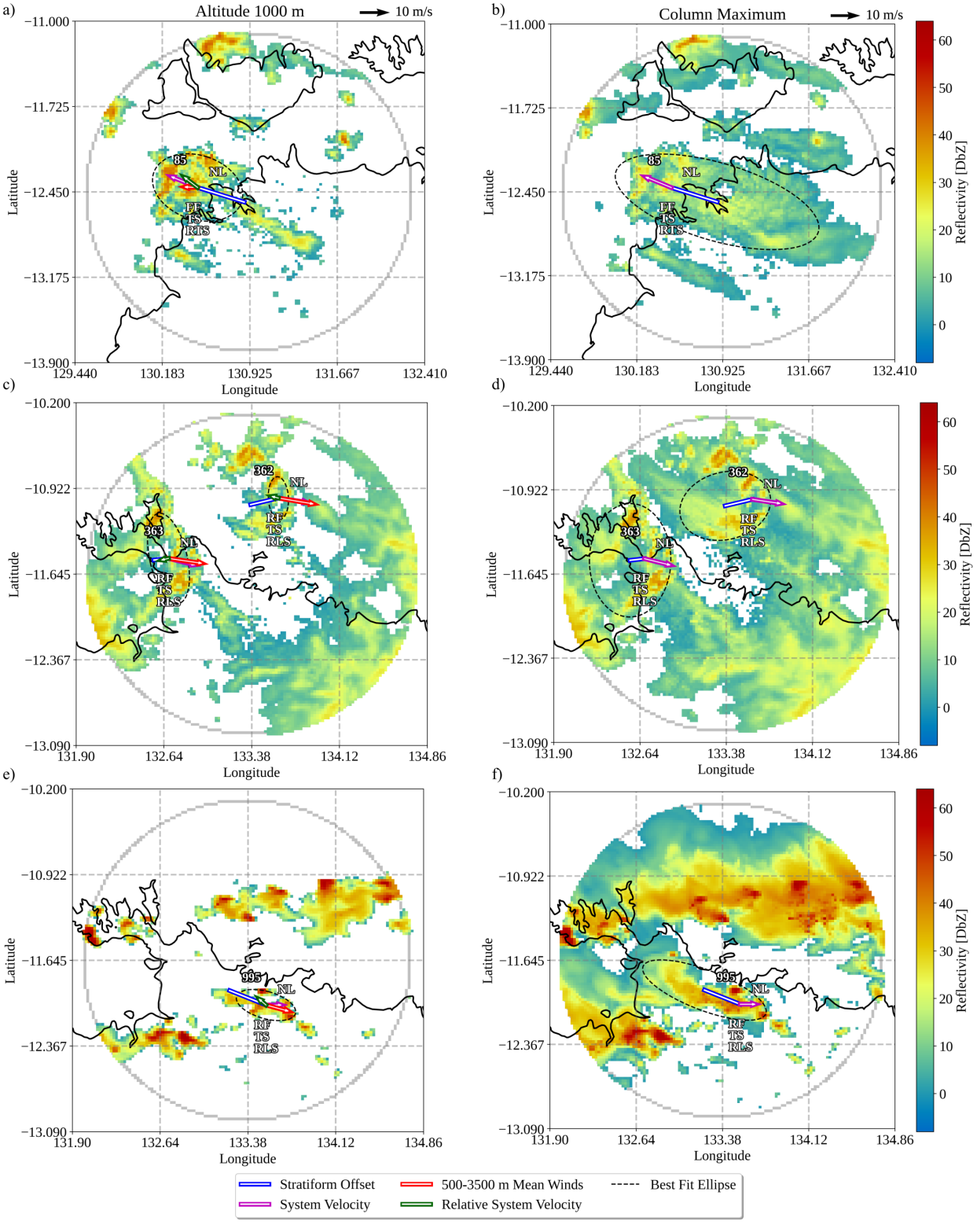

a)
Altitude 1000 m
10 m/s
b)
Column Maximum
10 m/s
c)
d)
e)
f)
Latitude
Longitude
Reflectivity [DbZ]
85
NL
RF
TS
RLS
362
363
995
Stratiform Offset
System Velocity
500-3500 m Mean Winds
Relative System Velocity
Best Fit Ellipse

FIG. 3. Example systems (identity numbers 85, 362, 363 and 995) identified in the radar dataset, a) to d), and ACCESS-C dataset, e) and f), at, a), b), 20/11/2020 20:30 UTC, c), d), 25/01/2021 19:50 UTC, e), f), 26/01/2021 04:00 UTC. Example systems are either front fed (FF) trailing stratiform (TS), relative trailing stratiform (RTS), or rear fed (RF) trailing stratiform (TS), relative leading stratiform (RLS). Reflectivity is at, a), c), e), 1 km altitudes, or the column maximum b), d) and f). The stratiform offset, ground-relative and flow-relative velocities are defined in section 2 b. A $10\,\mathrm{m\,s^{-1}}$ vector key is in the top-right of a) and b).

Finally, we assessed the new MINT configuration by manually checking its output on a subset of the ACCESS-C and operational radar datasets. In particular we generated assessment diagrams analogous to those of Short et al. (2023), showing ellipses fitted to convective and stratiform cloud regions, stratiform offsets, ground-relative and flow-relative MCS velocities, and the resulting MCS observation classifications, and checked that these were correct for each radar domain for the scans associated with the restricted sample over January 2021. Figure 3 provides example diagrams. Analogous diagrams depicting the shear vectors were generated to assess the shear based classifications.

The only obvious error occurred in the Arafura radar domain on 25/01/2021 20:10:00 (not shown), where the ellipse fitted to one system's convective region was much too eccentric. However, the centroid of this ellipse was still in the right position, and hence the stratiform offset vector, and dependent classifications, were still correct. This single example represents less than 0.1% percent of the observations, and less than 1% percent of the systems identified by MINT during this period. The January 2021 assessment diagrams are available online (Short 2022).

b. Preliminary Comparison of ACCESS-C and Radar

In this section we report some basic differences between radar and ACCESS-C MCS observations. Figures 4 a) to d) provide the radar and ACCESS-C raw and restricted sample sizes subdivided by monsoon regime and radar domain. Consistent with Bergemann et al. (2022), the ACCESS-C raw sample size is 27% percent larger than the radar raw sample size, suggesting ACCESS-C is producing too much convection overall. The differences between the restricted sample sizes is even more pronounced: there are almost twice as many observations in the ACCESS-C restricted sample than the radar restricted sample, and almost three times as many during the active monsoon, and in the Arafura radar domain.

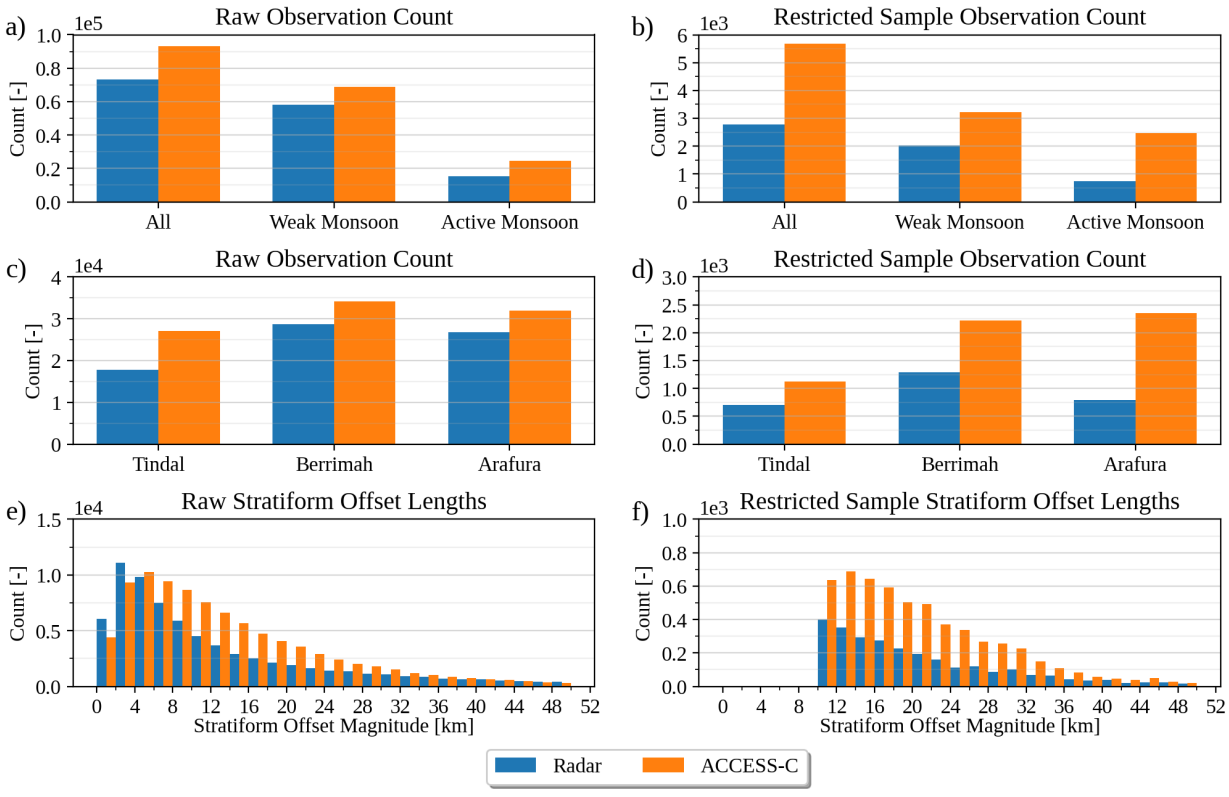
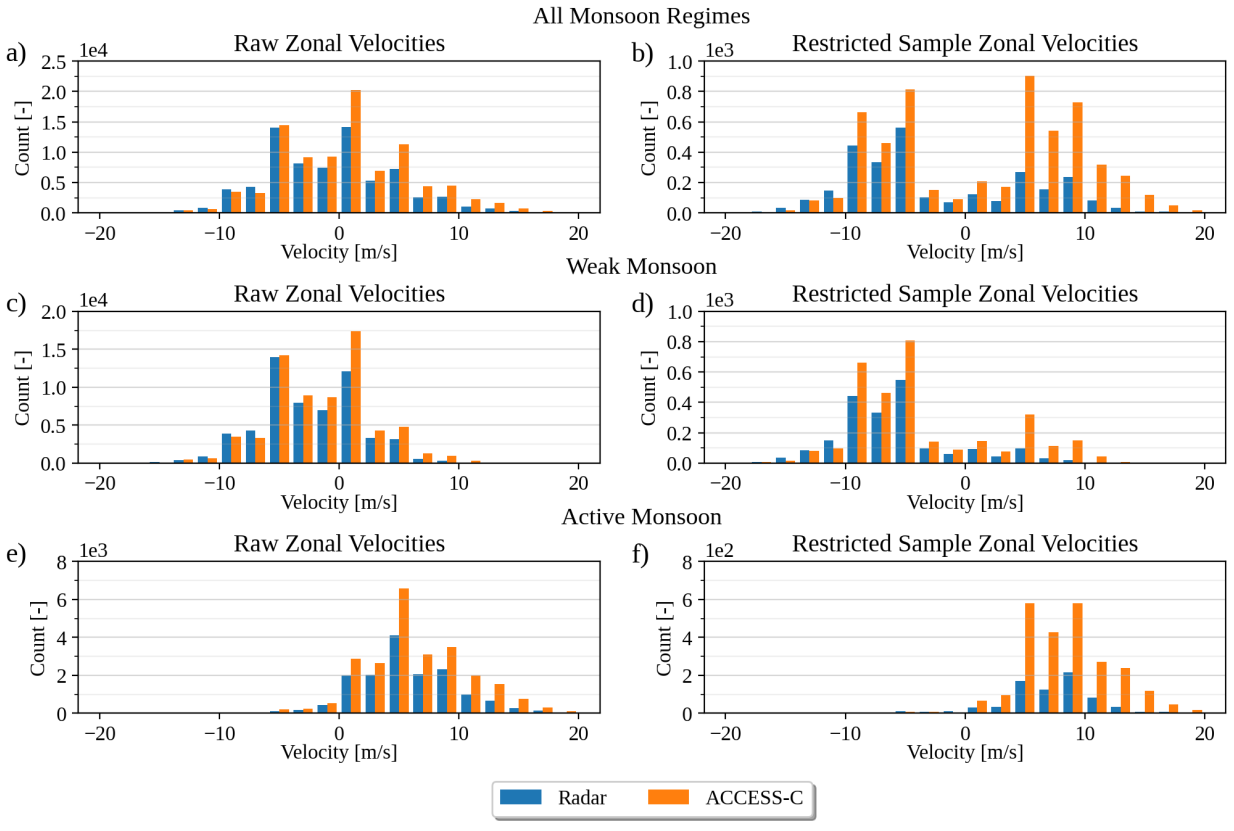


FIG. 4. The raw, a), c), e), and restricted b), d), f), sample sizes, a) to d), and stratiform offset magnitudes, e), f), for the radar and ACCESS-C system observations. The 2 km wide bins for the paired histograms in e) and f) are indicated by the x-axis ticks. Sample sizes are reported by monsoon regime, a), b), and radar domain, c) and d). System observations, stratiform offsets, and the raw and restricted samples are defined in section 2 b.

The large difference in restricted sample sizes occurs primarily because stratiform offsets are larger in ACCESS-C than in radar, resulting in relatively more ACCESS-C observations included in the restricted sample, as depicted in Figs. 4 e) and f). Note that the magnitude of the measured stratiform offset contains error associated with the radar scanning patterns. However, because the radar considered in this study have short 5 minute scanning patterns, geometric arguments show this error is usually ≤ 1 km, less than the 2.5 km gridlength of the reflectivity datasets. It is therefore very unlikely the differences in both the mode and shape of the histograms in Figs. 4 e) and f) solely represent scanning pattern errors. Appendix A1 provides further details.

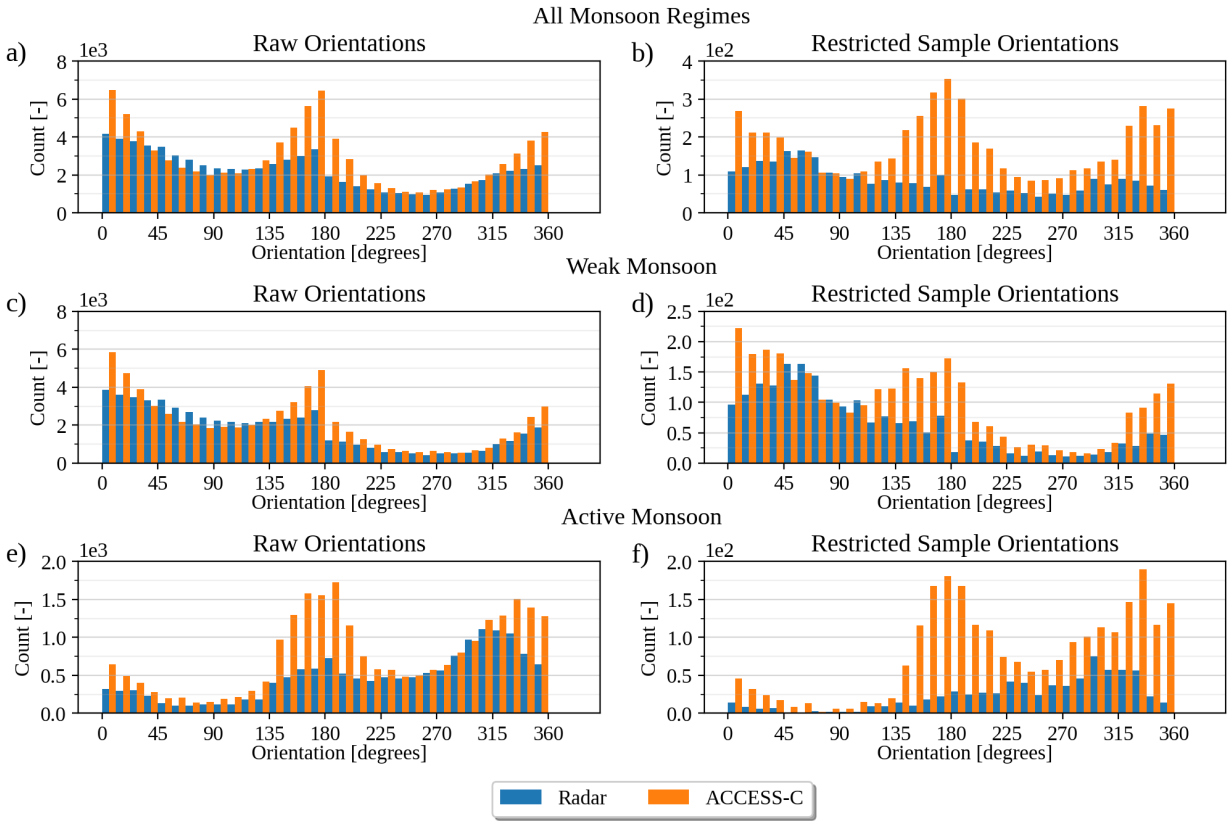
One reason the difference between the radar and ACCESS-C restricted sample sizes further increases during the active monsoon is that ground-relative system velocities, which are primarily zonal, are larger for ACCESS-C than radar observed systems, as depicted in Figs. 5 e) and f).



396 FIG. 5. Histograms of the raw sample, a), c), e), and restricted sample b), d), f), system ground-relative zonal
397 velocities, for all monsoon regimes, a), b), the weak monsoon, c) and d), and the active monsoon, e) and f). The
398 raw and restricted samples, system velocities $\mathbf{v} = (u, v)$, and monsoon regimes are defined in section 2 b. Note
399 that while observations for which $|\mathbf{v}| < 5\,\mathrm{m\,s^{-1}}$ are excluded from the restricted sample, this does not imply all
400 observations for which the system zonal velocity $|u| < 5\,\mathrm{m\,s^{-1}}$ are excluded.

404 In particular, 7% and 17% of the restricted sample radar and ACCESS-C velocities respectively
405 are greater than $12\,\mathrm{m\,s^{-1}}$. The overall larger velocities result in relatively more ACCESS-C
406 observations included in the restricted sample. During the weak monsoon there is good agreement
407 between ACCESS-C and radar ground-relative zonal velocities, as shown in Figs. 5 c) and d),
408 with the differences in the overall distributions shown in a) and b) driven almost entirely by the
409 active monsoon differences. Normalised versions of the histograms in Fig. 5, and the histograms
410 presented subsequently, are available in the online supplement.

415 Another striking difference between radar and ACCESS-C systems is the orientation of their
416 convective regions, where orientations are measured in degrees from lines of latitude using the



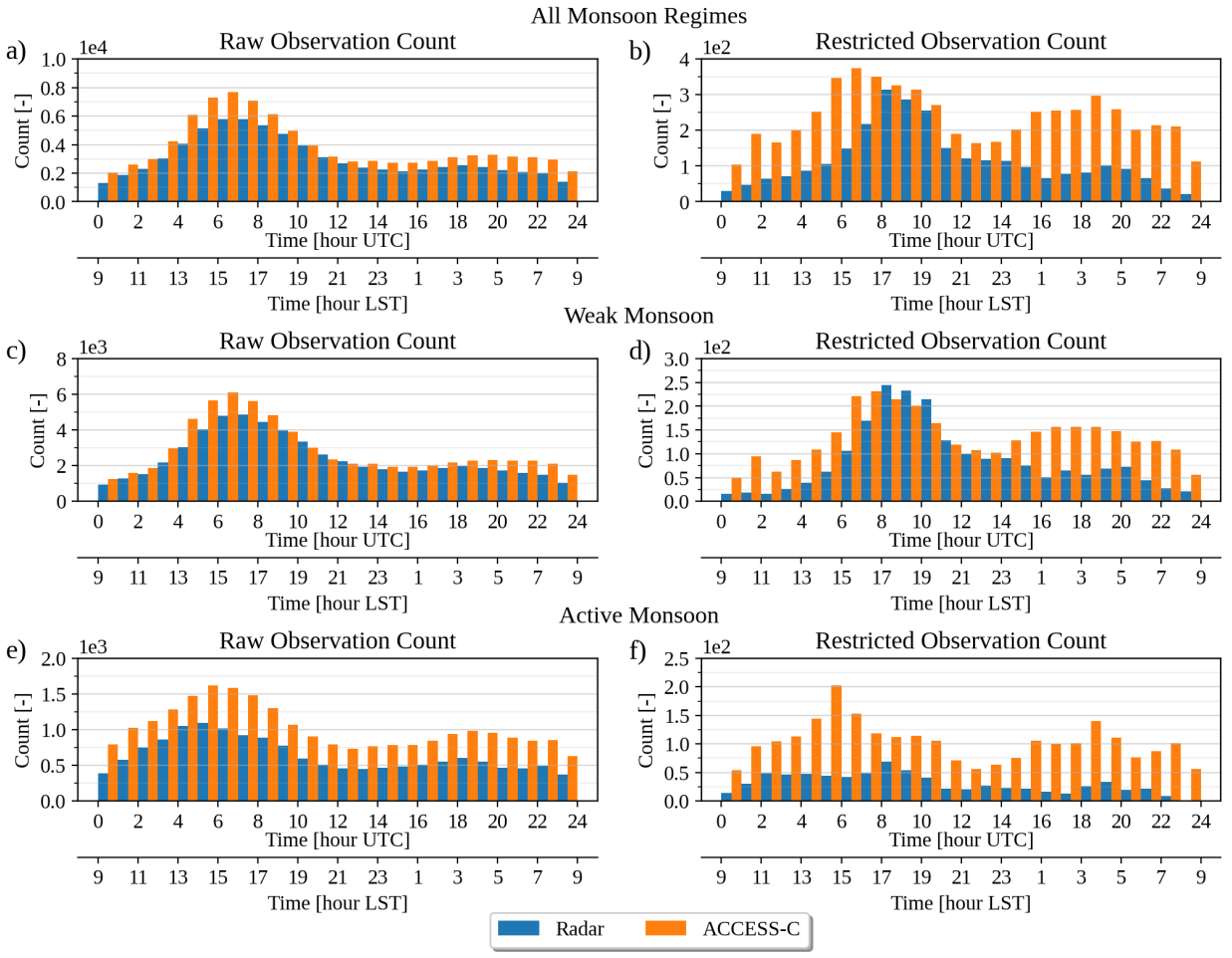
411 FIG. 6. As in Fig. 5, but for the orientation of system convective regions. Orientations are given by the
412 major axis of the ellipse fit to a given system's convective region, measured in degrees from lines of latitude,
413 with the orientation chosen to that the normal vector 90° anticlockwise lies in the same half plane as system
414 ground-relative velocity.

417 major-axis of the ellipse fitted to the convective region, choosing the orientation such that the
418 normal vector 90° counter-clockwise lies in the same half-plane as the system's ground-relative
419 velocity. Figure 6 provides histograms of the radar and ACCESS-C orientations. The ACCESS-C
420 and radar histograms are most similar for the raw sample during the weak monsoon as depicted
421 in Fig. 6 c), with both histograms exhibiting peaks and secondary peaks at 0 and 180 degrees
422 respectively, although the ACCESS-C histogram is more skewed towards these orientations. The
423 ambient wind, and ambient wind shear, is primarily zonal in this region, implying systems in the
424 raw sample tend to be oriented with the zonal wind, and the zonal wind shear, recalling that in
425 the raw sample orientation estimates are potentially erroneous as system's can intersect the radar
426 scanning boundary.

Figure 6 d) shows the restricted sample histograms during the weak monsoon. Now the radar system orientations peak at ≈ 50 degrees, indicating a more meridional orientation. This is consistent with classical, cold-pool driven systems dominating during the weak monsoon, with convective lines orienting in a roughly shear perpendicular direction. By contrast, the ACCESS-C histogram still has a peak and secondary peak at 0 and 180 degrees respectively, suggesting convective lines do not develop appropriately in ACCESS-C, and are more closely aligned with the ambient flow. Orientation differences are similar, but even greater during the active monsoon, depicted in Figs. 6 e) and f), with Figs. 3 c) and e) providing representative examples. The overall orientation biases shown in Figs. 6 a) and b) reflect biases from both the weak and active monsoon, with 11% and 24% of the radar and ACCESS restricted sample orientations respectively within 20° of 0° or 180°.

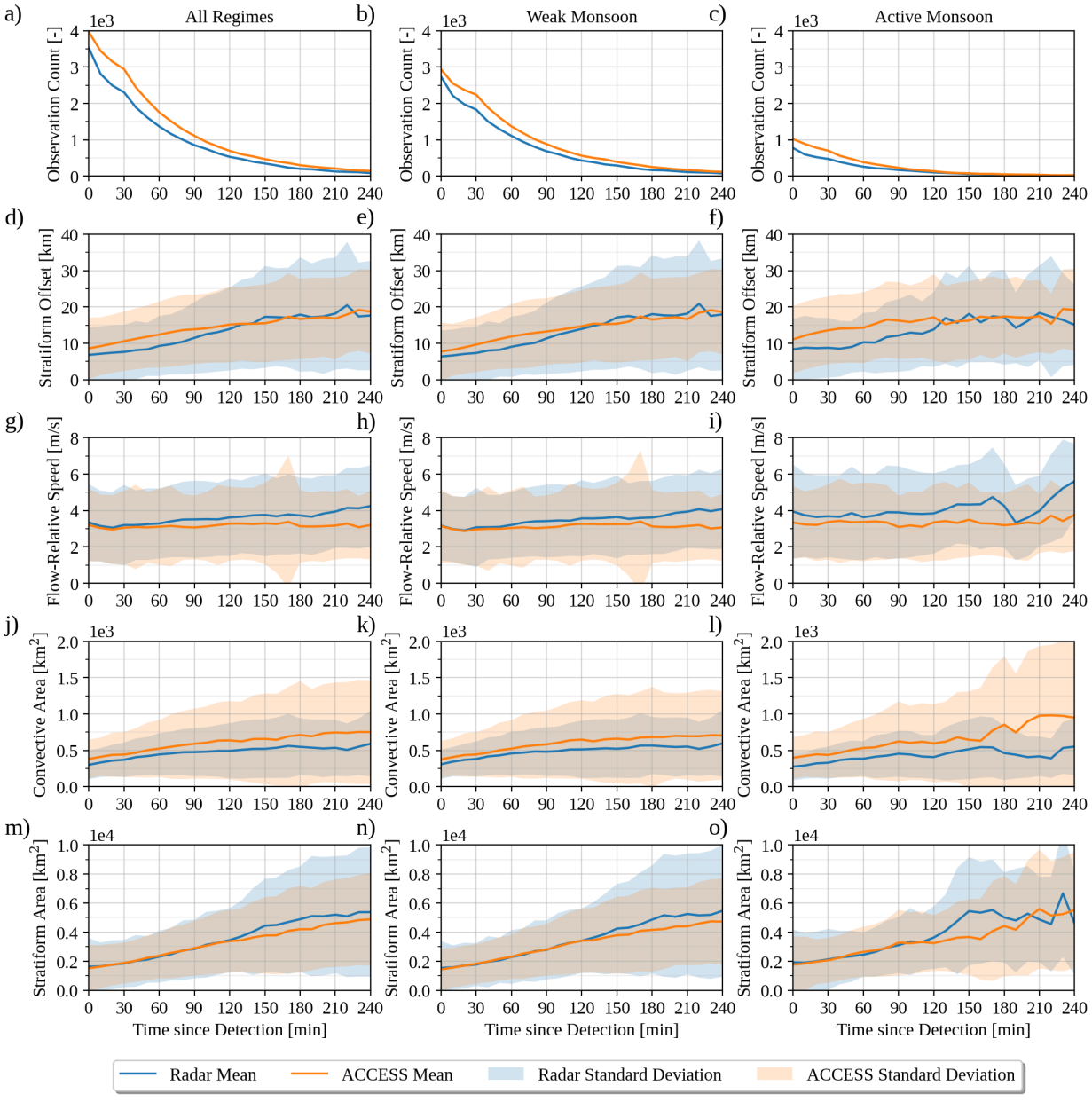
Another difference between the radar and ACCESS-C systems is the time of day they occur. Figure 7 shows histograms of the raw and restricted sample sizes as a function of hour of the day. For the raw sample during the weak monsoon depicted in Fig. 7 c), there is good agreement between the radar and ACCESS-C histograms: both show the most observations occur during the afternoon and early evening between 13:00 and 19:00 local solar time (LST), consistent with thermal and sea-breeze forcing. However, in the restricted sample shown in Fig. 7 d), the ACCESS-C histogram peaks $1-2$ hours before the radar histogram, and the ACCESS-C histogram shows a wide, secondary peak centred at ≈ 0300 LST.

Jucker et al. (2020) observed that in the Darwin region during weak monsoon conditions, at $\Delta x \approx 1\,\text{km}$ horizontal grid-lengths the UM's diurnal precipitation peaked ≈ 2 hours earlier than radar observations and also produced a large secondary peak in the early morning local solar time entirely absent from observations. Jucker et al. (2020) attributed this secondary peak to a biased UM land-breeze. Figures 7 c) and d) indicate analogous biases occur with the more organized system observations of the restricted sample, but not in the raw sample, suggesting that diurnal precipitation biases relate not only to an exaggerated land-breeze, but to issues with how the simulated convection organizes in response to the land-breeze. Figures 7 e) and f) show the diurnal cycles during the active monsoon. Now there are simply $30-50$% more observations in the ACCESS-C raw and restricted samples than those of radar at all times of day, although time differences in the small peaks in the restricted sample again suggest a timing bias.



438 FIG. 7. As in Fig. 5, but for the sample sizes at each hour of the day coordinated universal time (UTC) and
439 approximate local solar time (LST).

463 Histograms of other basic properties of systems are similar for ACCESS and radar systems,
464 although there are some interesting time dependent differences. Figure 8 shows the sample sizes,
465 and means and standard deviations of the magnitudes of the stratiform offset and flow-relative
466 system velocity vectors, and the convective and stratiform areas, as a function of the time since
467 detection of each system. To better reveal the temporal evolution of these properties we consider
468 a subset of the raw sample comprising those observations that meet the boundary intersection
469 threshold described in section 2 b, as this condition is sufficient to ensure stratiform offset and
470 velocity magnitudes, and system areas are accurate. Note that although time since system detection
471 does not necessarily correspond to time since initiation for an individual system, as systems can



459 FIG. 8. The sample sizes, a), b), c), average stratiform offset magnitude, d), e), f), average flow-relative system
460 speed, g), h) i), average convective area, j), k), l), and average stratiform area, m), n), o), as a function of time
461 since system detection. Shaded regions depict standard deviations. The three columns show results for all
462 monsoon regimes, the weak monsoon, and active monsoon respectively.

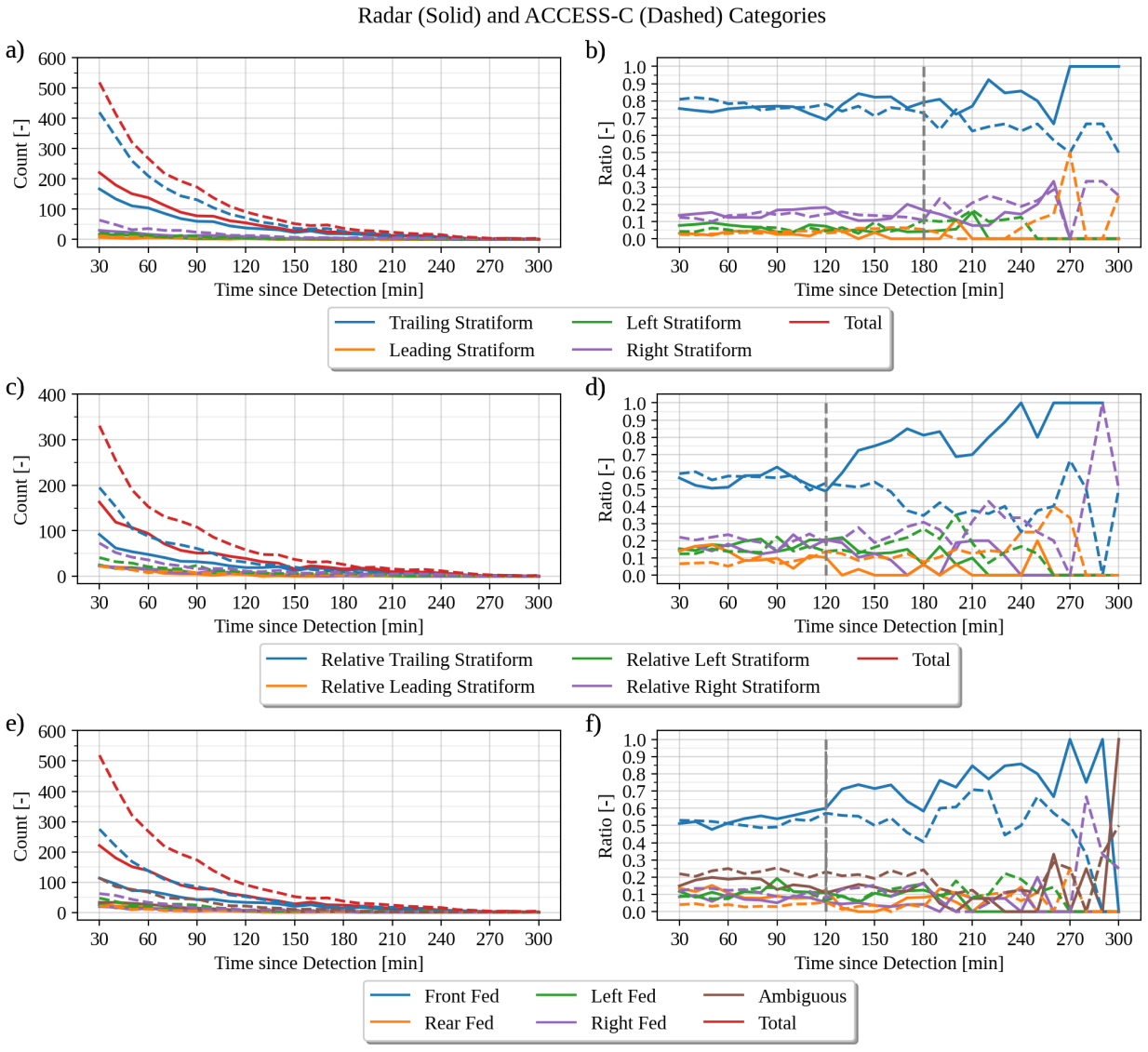
472 enter the radar domains fully formed, the average time dependent behaviour may still reflect how
473 the given variables evolve in a crude, averaged sense, as, all else being equal, systems that are
474 tracked for longer within the radar domain are more likely to have existed for longer.

Figures 8 d), e) and f) show the mean stratiform offset magnitudes and their standard deviations. For the first 120 minutes since detection, the average ACCESS-C stratiform offset magnitude is 2-3 km larger than that of radar, increasing approximately linearly, whereas the radar time series is flatter between 0-60 minutes. This difference is greatest during the active monsoon, suggesting the ground-relative velocity biases shown in Figs. 5 e) and f) also reflect simulated system convective regions advancing further ahead of their stratiform regions. After 120 minutes the average radar and ACCESS-C stratiform offset magnitudes are approximately equal.

Figures 8 g), h) and i) show the average radar and ACCESS-C flow-relative system speeds. Initially both are 3 to $4\,\mathrm{m\,s^{-1}}$, with the radar average speeds increasing by $\approx 1\,\mathrm{m\,s^{-1}}$ between 0 and 240 minutes, whereas the ACCESS-C speeds stay approximately constant. During the active monsoon, the radar flow-relative speeds are $\approx 0.5\,\mathrm{m\,s^{-1}}$ faster than the ACCESS-C speeds, implying that the radar flow-relative velocities must be more often directed against the mean winds, and the ACCESS-C flow-relative velocities directed with the mean winds, to produce the ground-relative zonal velocity biases shown in Figs. 5 e) and f).

Figures 8 j), k) and l) show the average area of system convective regions, indicating ACCESS-C convective regions are on average 100 to $200\,\mathrm{km^2}$ larger than those of radar, with larger standard deviations, particularly during the active monsoon, consistent with the cumulonimbus width biases discussed in section 1. The average stratiform areas shown in Figs. 8 m), n) and o) are remarkably consistent between radar and ACCESS-C for the first 120 minutes, after which observed stratiform areas become up to 20% larger those of ACCESS-C during the weak monsoon, and up to 50% larger during the active monsoon, although sample sizes in the latter are small. A physical interpretation of these stratiform area differences, consistent with previous studies (Hanley et al. 2015; Bergemann et al. 2022), is that when MCSs are first detected by MINT, they typically comprise only a few cumulonimbi, and the stratiform cloud associated with these are consistent between ACCESS-C and radar. However, as MCSs become more developed, ACCESS-C MCSs comprise fewer, larger cumulonimbi than radar, which fail to organise the same way as observed MCSs, with resulting errors in their stratiform cloud regions.

The comparisons considered in this section were replicated for each of the sensitivity tests summarized in Table 1. Results were very similar across all tests, with the exception of the NS test, likely due to small sample sizes in that test: details are provided in the online supplement.



506 FIG. 9. Counts a), c) and e), and ratios b), d) and f), of the stratiform offset categories, a) and b), relative
507 stratiform offset categories, c) and d), and low level inflow categories, e) and f), as a function of the time since
508 each system's first detection, derived from the modified two-layer version of MINT described in section 2 applied
509 to the Berrimah, Arafura and Tindal operational radar datasets. The vertical grey dashed lines in b), d) and f)
510 indicate the times at which the ratios are subjectively deemed less reliable due to low sample sizes. See Fig. 2
511 and section 2 b for further details.

505 *c. ACCESS-C and Radar Organizational Characteristics*

512 In this section we compare the results of applying the classifications depicted in Fig. 2 to the
513 radar and ACCESS-C observations. Figures 9 a) and b) provide time series of the stratiform offset

classification counts and ratios, ignoring ambiguous classifications where the stratiform offset vector forms an angle $\leq \theta_e = 10^\circ$ with a quadrant boundary. There is excellent agreement among the ratios of these various categories. While on its own this classification does not necessarily say much about the underlying organizational mechanisms (Short et al. 2023), it is nevertheless encouraging that ACCESS-C reproduces it so well.

Figures 9 c) and d) provide the relative stratiform offset classification counts and ratios. Up to 120 minutes from detection there is good agreement between the ratios, the most notable difference being in the relative leading stratiform category, for which the radar ratio is $\approx 10\%$ larger during the first 60 min. After 120 minutes the radar and ACCESS-C ratios diverge, noting sample sizes become small after this time, making ratios less reliable. The relative trailing stratiform category becomes more common for radar systems, consistent with previous work (Short et al. 2023), suggesting this represents a real difference between radar and ACCESS-C, at least up to 180 minutes.

Figures 9 e) and f), which provide the inflow classification counts and ratios, are consistent with Figs. 9 c) and d), showing that rear-fed systems are more common in radar than ACCESS-C observations for the first 120 minutes since detection, noting a rear-fed trailing stratiform observation is, roughly speaking, an example of a relative leading stratiform observation (Short et al. 2023). After 120 minutes front-fed systems become more common in radar, but roughly as common in ACCESS-C as before. Inflow directions are $\approx 10\%$ more likely to be classified as ambiguous in ACCESS-C than radar, meaning flow relative velocity magnitudes are $< 2\,\mathrm{m\,s^{-1}}$, indicating ACCESS-C systems are more likely to move with the mean low-level ambient winds than radar systems.

Figure 10 shows time series of the tilt and propagation classification counts and ratios. For the first 150 minutes, radar observations are $\approx 10\%$ and $\approx 20\%$ more likely to be up-shear tilted and down-shear propagating than ACCESS-C, with ACCESS-C observations $\approx$ 5-10% more likely to be shear-perpendicular for both classifications. This is consistent with the orientation differences shown in Fig. 6, indicating ACCESS-C systems are less likely to exhibit the relationships between shear and propagation directions, and convective region orientation, implied by the classical MCS model (Houze 2004; Short et al. 2023): additional figures provided in the online supplement corroborate this. In the online supplement we also consider how the ratios presented in Figs. 9 and 10 vary across the sensitivity tests detailed in Table 1. The core differences between radar

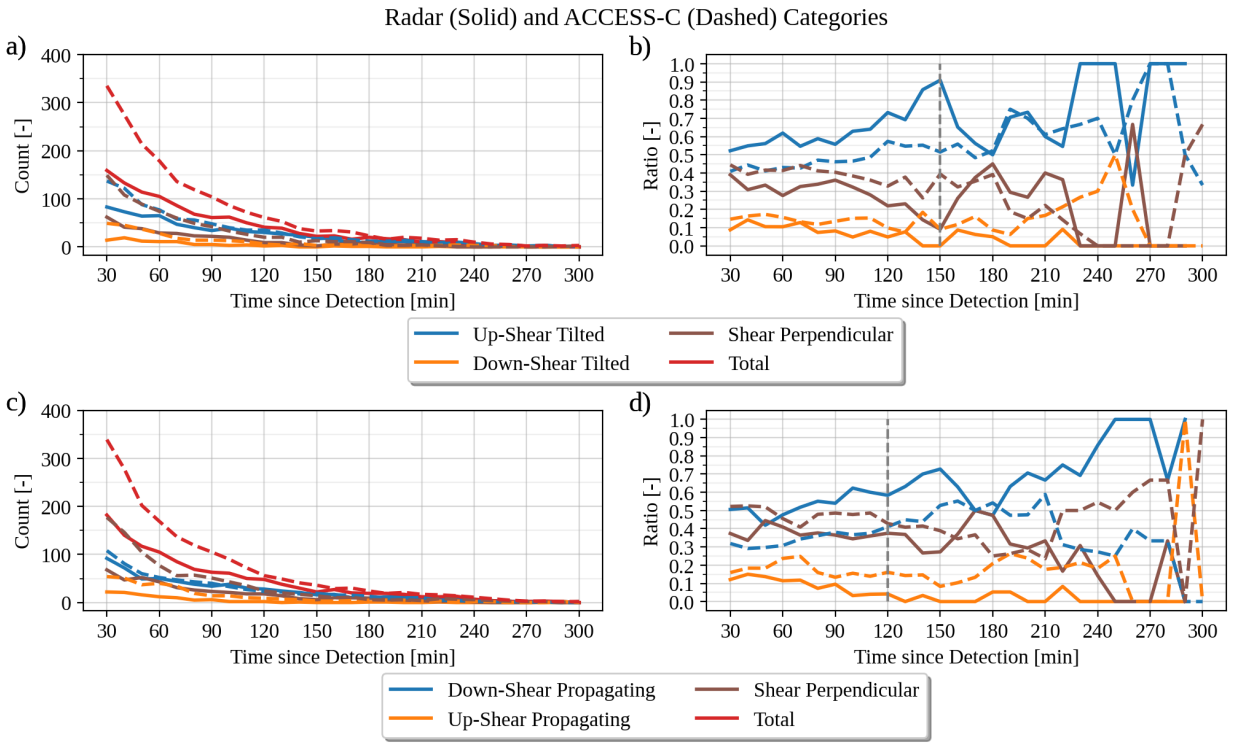
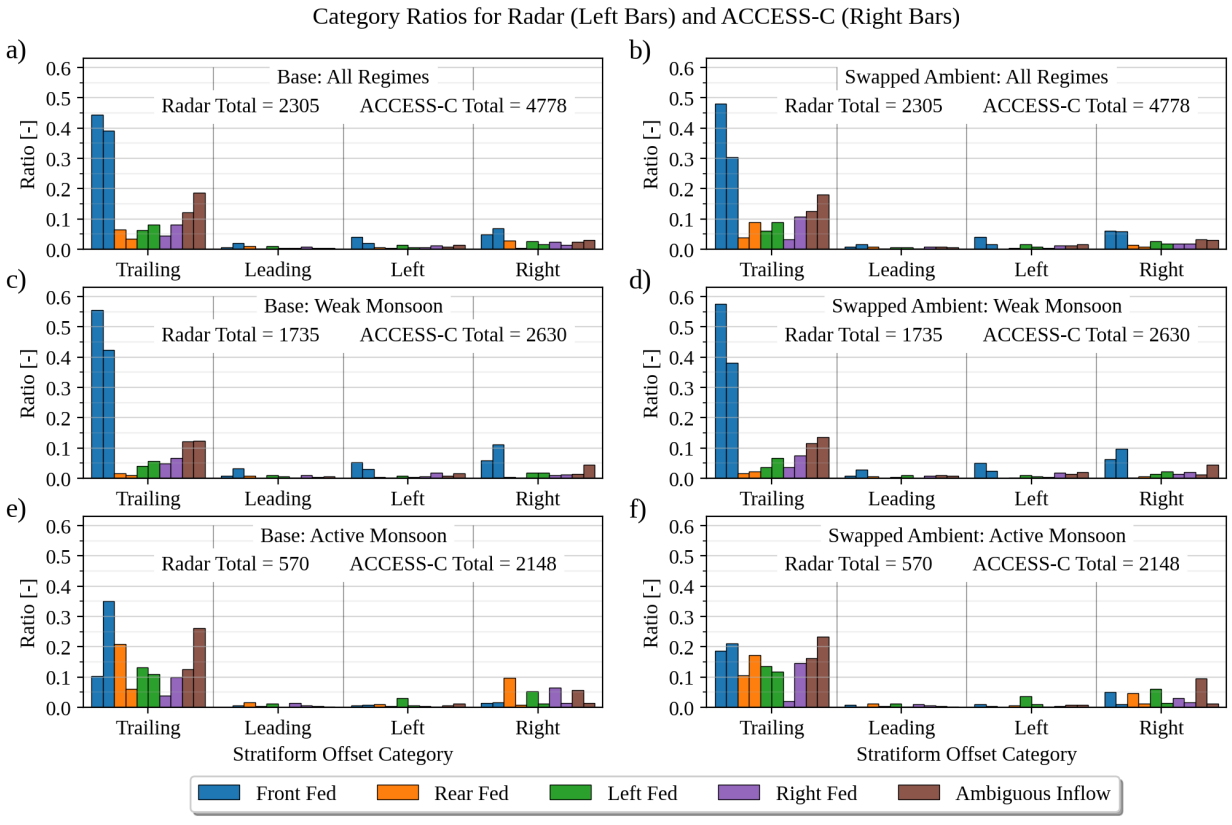


FIG. 10. As in Fig. 9, but for the tilt and propagation classifications. See Fig. 2 and section 2 b for further details.

and ACCESS-C systems described above are replicated across all tests, with the precise ratios exhibiting comparable sensitivities as in previous work (Short et al. 2023).

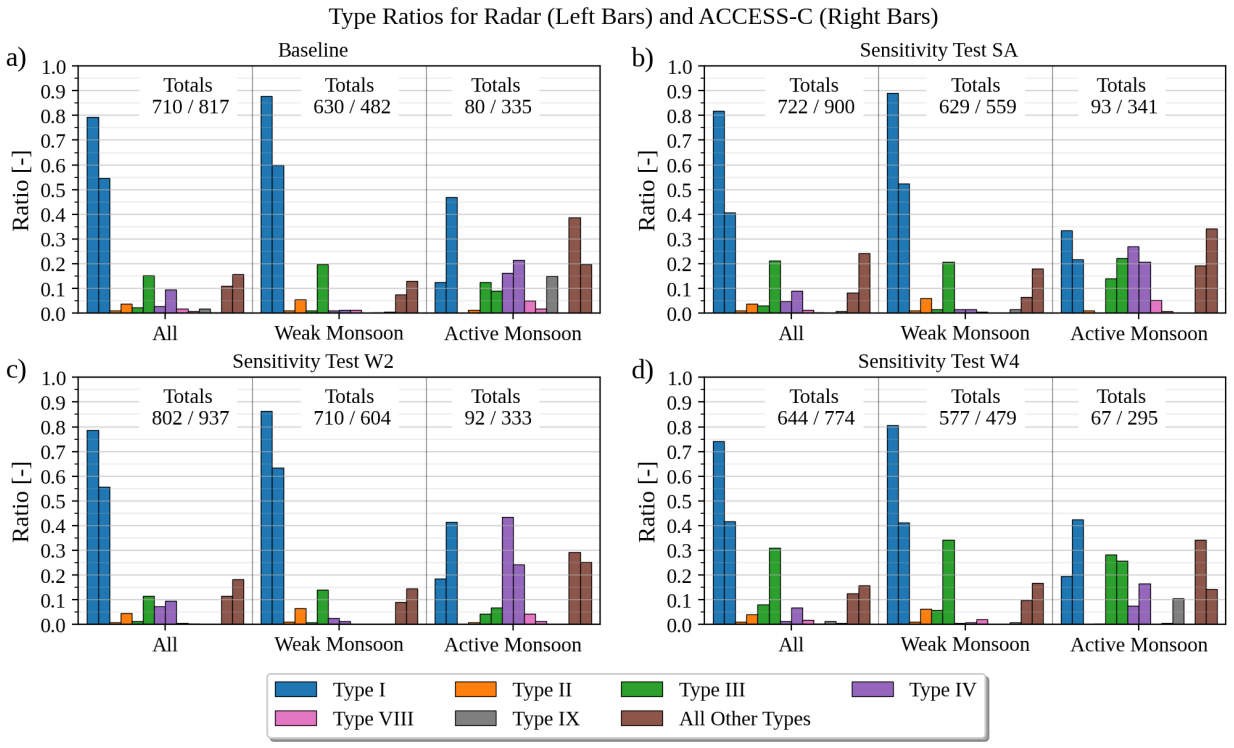
To investigate the co-occurrence of different categories, and how this changes across monsoon regimes, Fig. 11 shows ratios for the combined stratiform offset and inflow direction classifications, contrasting the baseline with the SA test, which exhibits the largest sensitivities. The SA test is arguably the most important new test not considered by Short et al. (2023), as it involves changing the dataset used to estimate ambient winds, a key methodological decision affecting all results dependant on these winds, i.e. the classifications depicted in Fig. 2 b) to e). During the weak monsoon, radar observations are $\approx 15\%$ more likely to be front-fed trailing stratiform than ACCESS-C, with ACCESS-C observations more likely to fall into the other trailing stratiform categories, or be front-fed right stratiform. The ratios change little between the baseline and SA test. By contrast, substantial differences between the baseline and SA test occur during the active monsoon, implying this regime accounts for most of the overall sensitivity.



546 FIG. 11. Ratios of observations in each combined stratiform offset and inflow category for radar, left bars, and
547 ACCESS-C, right bars. Results shown for, a), b), all monsoon regimes, c), d), the weak monsoon, and e), f), the
548 active monsoon, with the baseline results, a), c), e), contrasted with the SA sensitivity test results, b), d), f).

	Inflow	Stratiform	Tilt	Propagation
Type I	Front	Trailing	Up-Shear	Down-Shear
Type II	Front	Leading	Down-Shear	Down-Shear
Type III	Front	Trailing	Down-Shear	Up-Shear
Type IV	Rear	Trailing	Up-Shear	Up-Shear
Type VIII	Rear	Leading	Down-Shear	Up-Shear
Type IX	Right	Right	Down-Shear	Down-Shear

560 TABLE 2. Properties of the six most common category combinations, ignoring ambiguous inflow and shear
561 perpendicular categories, when classifications a), b), d) and e) in Fig. 2 can all be performed for the same system
562 observation.



563 FIG. 12. Histograms depicting, a), ratios of the six most common observation types (see Table 2) for which
564 the stratiform offset, inflow, and shear-relative tilt and propagation direction classifications can be performed
565 unambiguously, and how these ratios change between weak and active monsoon regimes. Histograms b), c) and
566 d) are analogous to a), but for the SA, W2 and W4 sensitivity tests described in Tab. 1 respectively.

567 To further investigate the co-occurrence of different organizational characteristics, we calculate
568 the ratio of observations for each combination of the stratiform offset, inflow, tilt and propagation
569 direction classifications, for simplicity neglecting observations with ambiguous inflow directions
570 or shear-perpendicular tilt or propagation directions. Table 2 details the six most common category
571 combinations, with numbering chosen for consistency with previous work (Short et al. 2023). Type
572 I observations are consistent with the classical MCS model (e.g. Houze 2004), being front-fed,
573 trailing stratiform, up-shear tilted, down-shear propagating, whereas type IV observations exhibit
574 low level inflows which pass underneath the stratiform cloud, and are thus inconsistent with the
575 classical model (Short et al. 2023).

576 Figure 12 presents the ratios of these types, and how they vary between the weak and active
577 monsoon, and across the four sensitivity tests showing the greatest differences in these ratios from

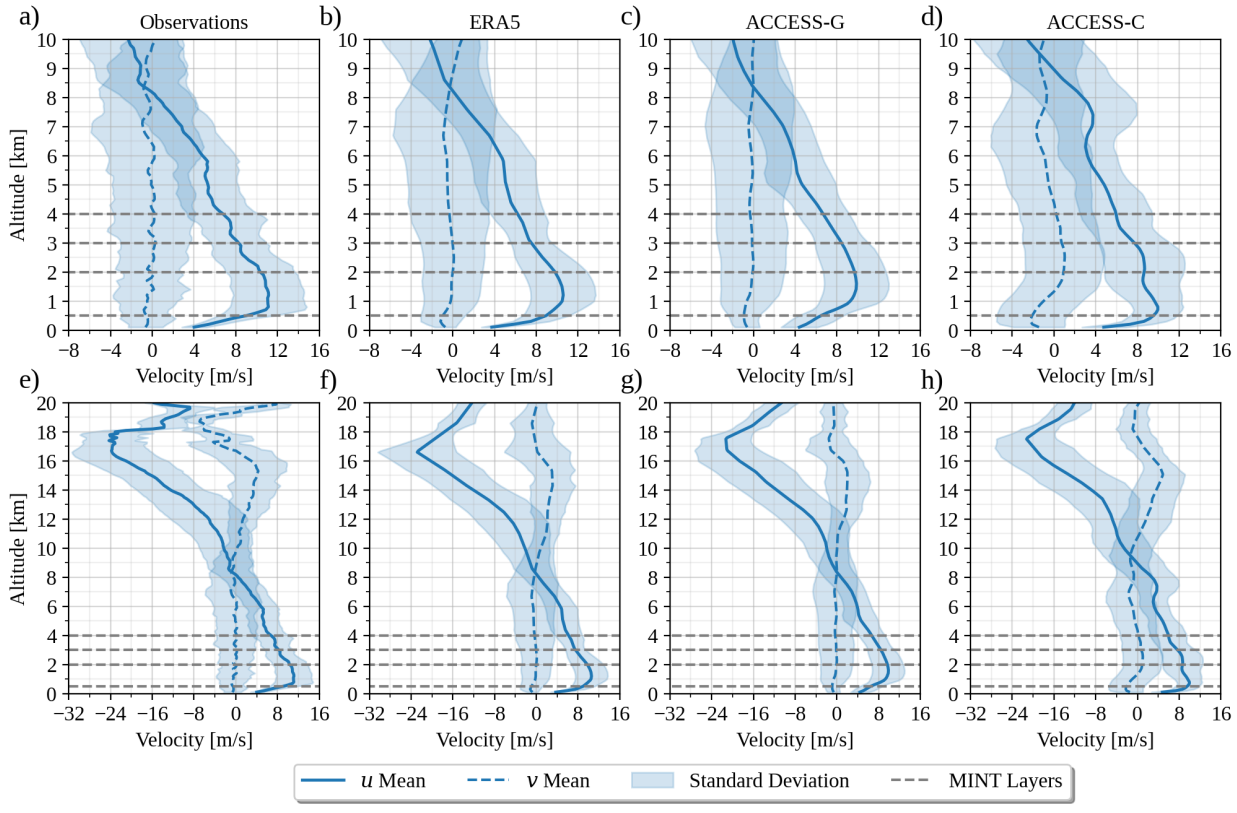
the baseline. The weak monsoon ratios for the baseline radar observations are nearly identical to those reported by Short et al. (2023), with comparable consistency across sensitivity tests: 80-90% of observations are type I. By contrast, only 40 to 65% of ACCESS-C observations are type I, with commensurate increases in type II and III, and miscellaneous observations, relative to the radar ratios.

During the active monsoon there is significant variability across sensitivity tests, but across all tests only 10 to 35% of radar observations are type I, consistent with Short et al. (2023). For ACCESS-C, the type I ratios also reduce from their weak-monsoon values, but less dramatically, now accounting for 20 to 50% of observations. During the active monsoon, observations are redistributed across types I, III and IV over the four sensitivity tests shown in Fig. 12, behaviour again consistent with previous work (Short et al. 2023).

The active monsoon sensitivities occur because of slight differences in how ERA5 and ACCESS-G resolve the low level winds, and because during the active monsoon the zonal shear changes direction over the vertical levels considered by MINT. Figures 13 a) to d) show the mean winds and their standard deviations at Darwin Airport (see Fig. 1) at 1200 UTC through the first 10 km of the atmosphere. These winds are derived from radiosonde observations, the ERA5 and ACCESS-G analyses, and the ACCESS-C forecast at a 24-hour lead time. The average is taken over the 18 active monsoon days from January and Febuary 2021: not all ACCESS-C fields were archived for 2022, so for consistency with later figures we consider just 2021. An extended discussion of how these wind profiles account for the active monsoon sensitivities in Fig. 12 is provided in the online supplement.

4. Discussion

To realistically simulate and predict organized convection at $\Delta x \approx 1$ km horizontal grid lengths, many diverse processes must be simulated realistically. The synoptic environment must be realistic. External mesoscale forcings like the land-sea breeze must be realistic, and the convection must respond to such forcing in the right way. Cumulus convection, when considered in aggregate, must self-organize for the right reasons, despite individual cumulonimbi being necessarily unrealistic in particular ways.



589 FIG. 13. Average wind profiles and their standard deviations at Darwin Airport at 1200 UTC for the 18 active
590 monsoon days from January and Febuary 2021. Profiles are derived from, a), e), radiosonde observations, b),
591 f), ERA5, c), g) ACCESS-G analyses, and d), h), ACCESS-C. Depicted are the winds between 0 and 10 km
592 altitudes, a) to d), and 0 to 20 km altitudes, e) to h).

610 In some respects ACCESS-C simulates organized convection remarkably well. Figure 9 b) shows
611 that ACCESS-C and radar observations are split between trailing, leading and parallel stratiform
612 categories in similar ratios, with no more than 5% difference between ACCESS-C and radar for each
613 category and test. However, major differences between radar and ACCESS-C organized systems
614 also exist. During the weak monsoon the diurnal cycles of the ACCESS-C and radar restricted
615 sample sizes are different, noting the "restricted sample" is defined in section 2 b. Consistent
616 with previous studies (Jucker et al. 2020), the ACCESS-C sample sizes exhibit a primary peak
617 1-2 hours earlier than radar, and a secondary peak in the early morning local solar time, that is
618 absent from radar observations. This difference suggests ACCESS-C systems organize differently
619 in response to land-breeze forcing than do radar systems. Curiously, Bergemann et al. (2022) found

the opposite behaviour when only the Tiwi Islands were considered, with a small, early morning peak present in CPOL observations, but not the UM simulations, suggesting the UM is producing this early morning convection in the wrong location.

Also, during the active monsoon ACCESS-C systems are more likely to have larger westerly ground-relative velocities. Finally, there are dramatic differences in the orientations of the convective cloud regions of ACCESS-C and radar systems, particularly during the active monsoon. These orientation differences are reflected in the tilt and propagation classification results, suggesting ACCESS-C systems organize differently to radar systems. Diagnosing the reasons for the above differences is beyond the scope of this study: instead we explore some hypotheses to be tested in future work.

At least three distinct mechanisms have been proposed to explain long-lived, self-organized convective systems. The first are mesoscale flows, which are "steady" in the sense that they persist longer than the individual cumulonimbi: these steady flows may both force new cumulonimbi, and be maintained by cumulonimbi in aggregate (e.g. Moncrieff and Green 1972; Moncrieff and Miller 1976; Moncrieff and So 1989). The second mechanism is the cold-pool density current, which is assumed to form under stratiform and convective cloud, and which forces new convection where the magnitude of the cold-pool vorticity matches or exceeds that of the ambient shear (e.g. Rotunno et al. 1988; Weisman et al. 1988; Weisman and Rotunno 2004). The final mechanism is the organizational effect of internal gravity waves (e.g. Bretherton and Smolarkiewicz 1989; Mapes et al. 2003; Kawashima 2003; Lane and Zhang 2011). Numerical experiments typically invoke a combination of these mechanisms to explain simulated self-organization (e.g. Lane and Moncrieff 2015; Moncrieff and Lane 2015).

The simplest hypothesis for the active monsoon ground-relative zonal velocity differences between radar and ACCESS-C systems, is that the ACCESS-C systems are advected faster, by stronger ACCESS-C ambient winds. However, Fig. 13 shows ACCESS-C ambient winds are not stronger at 1200 UTC, and the 0000 UTC comparison (not shown) is similar. An alternative hypothesis is that ACCESS-C systems are more likely to develop cold pools during the active monsoon, resulting in systems which propagate faster than the low-level ambient winds. Figure 14 provides skew-T diagrams from the 12 UTC Darwin airport radiosonde observations, ERA5 and ACCESS-G analyses, and ACCESS-C forecasts, averaged over the 18 active monsoon days from January and

Febuary 2021, as in Fig. 13. Between 0 and 1 km altitudes, ACCESS-C is on average $\approx 1\,^{\circ}$C warmer than observations, which is either inconsistent with exaggerated ACCESS-C cold pools, or reflects a compensating ACCESS-C mean state bias. Furthermore, Grant et al. (2020) demonstrated in a high-resolution convection permitting model that weaker cold-pools produce more flow-aligned MCSs, behaviour consistent with the orientation differences between ACCESS-C and radar observations given by Fig. 6.

Steady-flow considerations provide another hypothesis for the active monsoon speed differences. As argued by Short et al. (2023), type I observations are consistent with the classical MCS model (Rotunno et al. 1988; Houze 2004), but type IV rear-fed trailing stratiform systems are not. Figure 12 c) shows that when classified using winds between 500 m and 2 km, type IV systems are more common during the active monsoon for both radar and ACCESS-C.

While type IV observations are up-shear tilted and up-shear propagating with respect to the boundary layer westerly shear, during the active monsoon they are down-shear tilted, down-shear propagating with respect to the easterly shear between 2 km and the tropopause, as depicted in Figs. 13 e) to h). If we ignore boundary layer shear, type IV observations are therefore analogous to the overturning, or "steering-level" steady-flow model (Moncrieff and Green 1972). In this model a mesoscale inflow ascends, completely overturns, and flows back the way it came. The flow structure propagates relative to the ground at a speed c equal to the undisturbed ambient wind at a particular height - the steering level.

The non-dimensional Richardson number which determines the steering level, and hence c, is

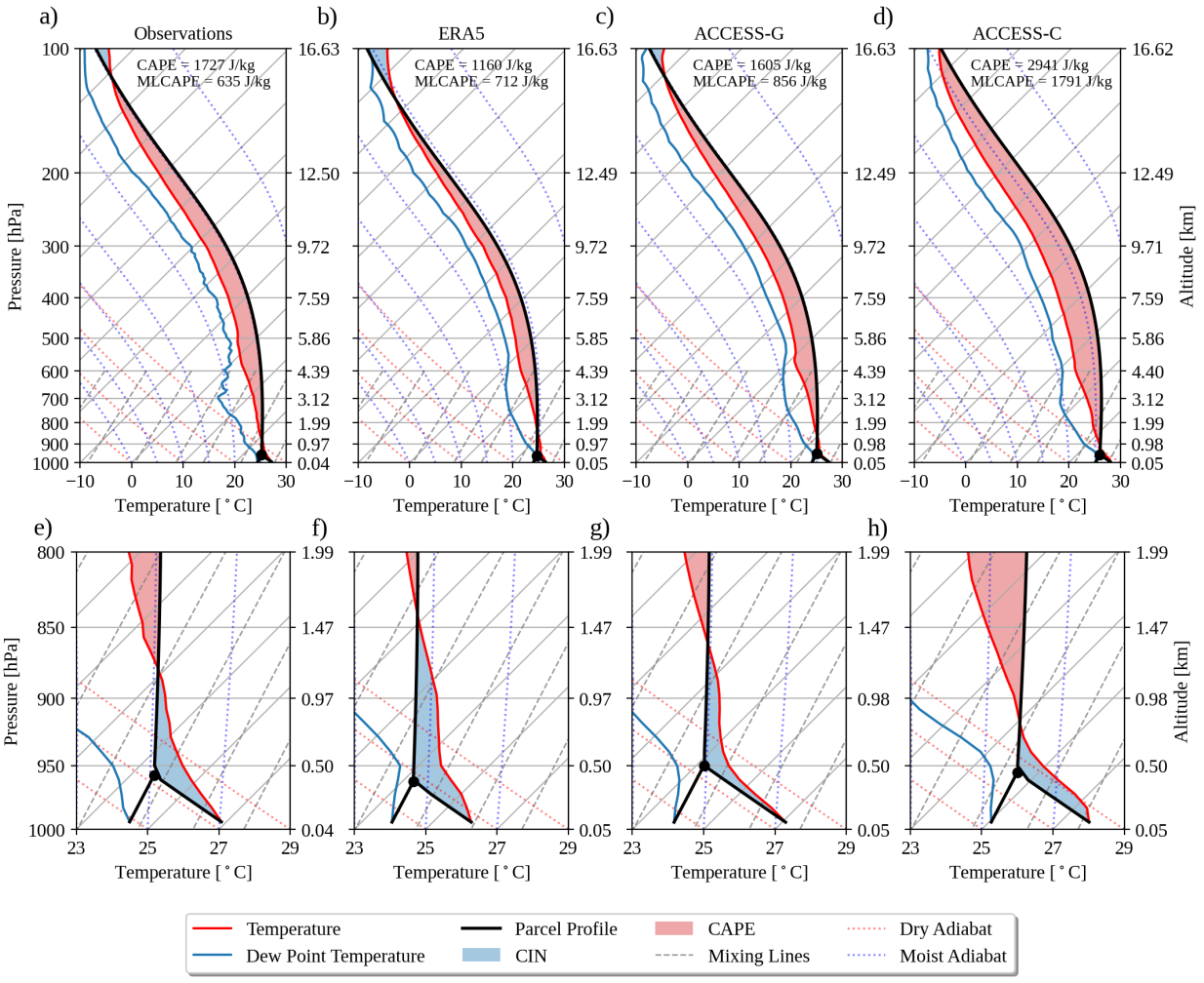
$$R = \frac{\text{CAPE}}{\frac{1}{2}(\Delta U)^2}, \tag{1}$$

where ΔU is the difference between the undisturbed horizontal winds at the inflow level and the tropopause. If conservation of mass is maintained by a symmetric down-draft, steady overturning requires

$$R \lessapprox 1 \tag{2}$$

for compressible flow (Moncrieff 1978).

Figure 14 shows the convective available potential energy (CAPE) and 100 hPa mixed-layer CAPE (MLCAPE) calculated from the composite profiles for each dataset: these are similar



674 FIG. 14. Skew-*T* diagrams from Darwin Airport vertical profiles of temperature and dew-point temperature
675 at 1200 UTC, for the 18 active monsoon days from January and Febuary 2021. Profiles are derived from, a),
676 e), radiosonde observations, b), f), ERA5, c), g) ACCESS-G analyses, and d), h), ACCESS-C. Depicted are the
677 convective available potential energy (CAPE) and convective inhibition (CIN) between 1000 and 100 hPa, a) to
678 d), and 1000 to 800 hPa, e) to h).

681 to the average of the CAPE and MLCAPE calculated for each day individually, and from the
682 CAPE values provided with the soundings (not shown). For ACCESS-C, CAPE=2941 $\mathrm{J\,kg^{-1}}$,
683 approximately twice that of the other datasets, whereas the MLCAPE=1791 $\mathrm{J\,kg^{-1}}$, nearly triple
684 that of the other datasets. This difference occurs because ACCESS-C is on average $\approx 1\,^{\circ}\mathrm{C}$ warmer
685 at low levels, as noted above.

From Fig. 13 we estimate $\Delta U = 30\,\mathrm{m\,s^{-1}}$ and $\Delta U = 28\,\mathrm{m\,s^{-1}}$ for observations and ACCESS-C respectively. Using MLCAPE as a representative average CAPE gives $R = 1.4$ and $R = 4.6$ for observations and ACCESS-C respectively. Equation 2 thus suggests it is less likely ACCESS-C systems comprise steering-level steady flows than observed systems. Indeed, when $R > 2.8$ steady-flow arguments show "purely propagating" systems, which move relative to the winds throughout the troposphere become possible (Moncrieff and Miller 1976), suggesting ACCESS-C systems are more likely to be of this type during the active monsoon. For active monsoon ambient zonal winds, purely-propagating systems must have ground-relative velocities either $\mathbf{v} \lessapprox -20\,\mathrm{m\,s^{-1}}$ or $\mathbf{v} \gtrapprox 10\,\mathrm{m\,s^{-1}}$, the choice determined by how such steady-flows develop (Moncrieff and Miller 1976), possibly through the influence of gravity waves (Lane and Moncrieff 2015). Furthermore, the front to rear flow in the purely propagating model (Moncrieff and Miller 1976) may still produce a front-fed trailing-stratiform structure, even without cold-pools, consistent with the results discussed above.

While there is ambiguity in the above argument, as different choices of average CAPE could be made, the central idea is that for the same shear, larger CAPE makes it harder for the flow to overturn. These ideas are also consistent with the greater proportion of type I observations in ACCESS-C in Fig. 12 c), although there is sensitivity in these classifications across the other tests. A more detailed study focusing on the active monsoon is required to properly test these ideas.

5. Conclusions

This study described a new configuration of the convective system tracking algorithm MINT (Short et al. 2023), and applied this to radar and operational ACCESS-C convection permitting model data encompassing Australia's Northern Territory for the two wet seasons spanning October to May between 2020 and 2022. We compared radar and ACCESS-C MINT output to systematically assess differences in the mesoscale organizational characteristics of observed and simulated convective systems. The radar results were essentially identical to previously established climatologies (Short et al. 2023). The most significant limitation of the present study, and that of Short et al. (2023), is arguably the reliance on individual radar domains, which necessarily restricts the analysis to convective systems typically smaller than common MCS size threshold definitions (Houze 2004; Hitchcock et al. 2021): future work could address this using radar composites.

The ACCESS-C results exhibited similarities and differences with observations. There were 24 – 28% more system observations in ACCESS-C raw MINT output than in radar, suggesting overactive ACCESS-C convection. There were approximately twice as many observations in the ACCESS-C restricted samples on which classifications were performed, primarily because ACCESS-C observations had larger stratiform offsets, likely because these offsets grew too quickly.

During the active monsoon, ACCESS-C restricted sample ground-relative velocities were in absolute terms 3-11% more likely to exceed $12\,\mathrm{m\,s^{-1}}$ than radar velocities across all sensitivity tests except one, with the exception likely due to small sample sizes. Mesoscale dynamic considerations (Moncrieff and Green 1972; Moncrieff and Miller 1976; Moncrieff 1978), coupled with ACCESS-C thermodynamic biases, provide one hypothesis for these ground-relative velocity differences.

During weak monsoon conditions, ACCESS-C convective systems form at the right times of day. However, counts of the more organized system observations comprising the restricted samples showed ACCESS-C counts peaked 1-2 hours earlier than in radar, and featured a secondary peak in the early morning local solar time, absent from observations, consistent with Unified Model (UM) diurnal precipitation cycle biases in northern Australia (Jucker et al. 2020), suggesting issues with how land-breeze convection subsequently organizes.

There were major differences in the orientations of radar and ACCESS-C convective regions, with 11-14%, and 21 – 27% of the radar and ACCESS-C restricted sample orientations respectively within 20° of 0° or 180°, i.e. oriented zonally. Radar systems were oriented more meridionally, perpendicular to the primarily zonal ambient wind, and ambient wind shear, consistent with the classical MCS model (Houze 2004). These orientation biases imply ACCESS-C produces too many "training-lines" (Schumacher and Johnson 2006), likely contributing to ACCESS-C extreme rainfall biases in the Darwin region (Rennie et al. 2022).

Radar and ACCESS-C observations were distributed remarkably consistently among the Parker and Johnson (2000) classifications, with trailing, leading, left and right stratiform observations accounting for 65 – 85%, 0 – 7%, 5 – 10%, and 10 – 20% of observations respectively, with no more than 5% difference between ACCESS-C and radar across the sensitivity tests considered. However, in absolute terms ACCESS-C observations were 5 – 10% more likely to be relative right stratiform, and 5 – 15% more likely to have ambiguous system relative inflows, corresponding to system ground-relative velocties roughly the same speed as the mean low level winds. ACCESS-C

systems were also 10 – 35% less likely to be up-shear tilted, and 15 – 30% less likely to be down-shear propagating, with commensurate increases in the down-shear tilted, up-shear propagating, and shear perpendicular categories.

When the above classifications were considered in conjunction, during the weak monsoon 80 – 90% of radar observations were front-fed, trailing stratiform, up-shear tilted and down-shear propagating, consistent with the classical MCS model (Houze 2004; Short et al. 2023). However, during the active monsoon the frequency of such observations dropped dramatically, accounting for only 10 – 35% of the overall observations. Instead, rear-fed trailing stratiform systems, inconsistent with the classical model, become more common, accounting for 15 – 45% of observations. These radar results were consistent with previously established climatologies (Short et al. 2023).

ACCESS-C partially captured this difference between weak and active monsoon convective organization, although the difference was less stark. During the weak monsoon, observations consistent with cold-pool theory accounted for between 40 – 70% of total observations, with front-fed leading-stratiform, and front-fed trailing stratiform, down-shear tilted, up-shear propagating observations now accounting for 5-10%, and 5-35% of observations respectively. During the active monsoon, observations consistent with the classical model still accounted for 20 – 55% of observations, although other evidence suggesting weaker ACCESS-C cold pools indicate alternative mechanisms may be contributing to this structure.

These differences show ACCESS-C convection organizes differently to radar convection, particularly during the active monsoon. These differences may stem from different mesoscale forcings, in particular a biased land-breeze (Jucker et al. 2020), ACCESS-C thermodynamic biases, or fundamental differences in self-organizational mechanisms such as cold pool strength, suggested by the orientation, tilt and propagation classification differences. Numerical experiments are necessary to diagnose the reasons for these differences.

Even if the ACCESS-C environmental thermodynamic and diurnal cycle biases could be corrected, the fundamental question of whether $\Delta x \approx 1\,\text{km}$ grid-length convection permitting models can realistically simulate convective organization and explicitly resolve the associated mesoscale flows, despite the component cumulonimbi being necessarily unrealistic, remains. While this study documents some remarkable progress in the operational simulation of convective organization, many operationally significant biases, such as orientation biases, persist. More work is

necessary if convection permitting models are to realistically simulate convective organization and its seasonal variability in current climates, let alone in simulated future climates.

Acknowledgments. Funding for this study was provided for Ewan Short by the Australian Research Council's Centre of Excellence for Climate Extremes (CE170100023). We would also like to acknowledge the dedicated work of Josh Soderholm, Alain Protat, Valentin Louf and Jordon Brook in producing the radar datasets, Valentin Louf for contributing the Steiner classification code, and Stacey Hitchcock, Claire Vincent and Michael Reeder for some helpful discussions.

Data availability statement. The radar, ERA5, and ACCESS datasets are publicly available online through the National Computational Infrastructure data catalogue (Soderholm and Protat 2022; European Centre for Medium-Range Weather Forecasts 2020; Bureau of Meteorology 2022). The radiosonde data, and code for the TINT and MINT algorithms, is also freely available online (Argonne National Laboratory 2022; Short 2022; University of Wyoming 2022).

APPENDIX

A1. Stratiform Offset Error Model

In this appendix we elaborate on the error model introduced by Short et al. (2023). First note the radars considered in this study have scanning patterns encompassing elevations ϵ from 0.5° to 32°, with the pattern beginning at the reference time $t = 0$ and ending at $t = 252$ seconds. We approximate

$$\epsilon = a \sinh\left(\frac{t}{c}\right) + \frac{1}{2}, \tag{A1}$$

where $c \approx 62.57$ and $a \approx 1.12$. For those elevations in the scanning pattern, and the times at which each elevation is first scanned, equation A1 is accurate to within 3%. For a given point in the radar domain $\mathbf{x}$ with altitude z and horizontal distance from the radar r, we approximate

$$\epsilon = \arctan\left(\frac{z}{r}\right). \tag{A2}$$

Using equations A1 and A2 we can estimate the time since the beginning of the pattern at which a given point $\mathbf{x}$ is scanned.

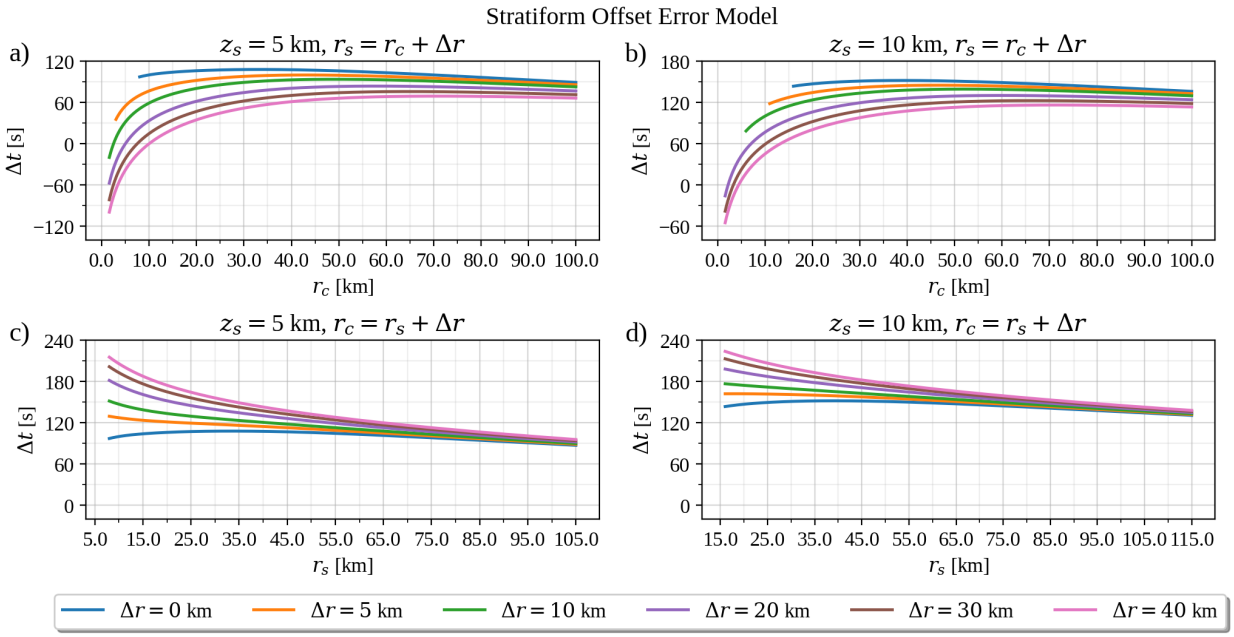


FIG. A1. The time delay Δt between radar scans of a "stratiform" point at, a), c), 5 km altitude, and b), d), 10 km altitude, and a "convective" point at 1 km altitude. Depicted are examples where the stratiform point is further from the radar than the convective point, a), b), and vice versa, c), d).

Consider "convective" and "stratiform" points $\mathbf{x}_c$ and $\mathbf{x}_s$, with altitudes $z_c = 1$ km and $z_s > 1$ km, and horizontal distances from the radar r_c and r_s, sampled by the radar at two different times t_c and t_s, calculated as above. If we treat these two points as a convective system, the stratiform offset as measured by the radar $\mathbf{s}$ would be the horizontal component of $\mathbf{x}_s - \mathbf{x}_c$.

Suppose that during the scanning pattern $\mathbf{x}_s$ and $\mathbf{x}_c$ move at a known constant horizontal velocity $\mathbf{u}$. Let $\mathbf{x'}_s$ and $\mathbf{x'}_c$ denote the positions of the convective and stratiform points $\mathbf{x}_s$ and $\mathbf{x}_c$ at $t = 0$, and $\mathbf{s'} = \mathbf{x'}_s - \mathbf{x'}_c$ denote the "true" stratiform offset, i.e. the stratiform offset corresponding to the simultaneous observation of $\mathbf{x'}_s$ and $\mathbf{x'}_c$ at $t = 0$. Define $\mathbf{e} = \mathbf{u}\Delta t$, with $\Delta t = t_s - t_c$. Then

$$\mathbf{s} = \mathbf{x'}_s + \mathbf{u} \cdot t_s - (\mathbf{x'}_c + \mathbf{u} \cdot t_c) = \mathbf{s'} + \mathbf{e}, \tag{A3}$$

Note the triangle inequalities imply,

$$|\|\mathbf{s'}\| - \|\mathbf{e}\|| \leq \|\mathbf{s'} + \mathbf{e}\| \leq \|\mathbf{s'}\| + \|\mathbf{e}\|, \tag{A4}$$

with equalities if and only if $\mathbf{s}'$ and $\mathbf{u}$ are collinear.

Figures A1 a) and b) show Δt as a function of r_c, assuming $\mathbf{x}_s$ is further from the radar than $\mathbf{x}_c$, i.e. that $r_s = r_c + \Delta r$, where Δr is a non-negative number. When r_c is small, $\mathbf{x}_s$ may be sampled before $\mathbf{x}_c$, resulting in a negative Δt. As r_c increases so does Δt, reaching 0 when the convective and stratiform points are sampled at the same elevation. As r_c increases further, Δt peaks, then slowly decreases. Figures A1 c) and d) are analogous to a) and b), but for situations where $\mathbf{x}_s$ is closer to the radar than $\mathbf{x}_c$. In such cases $\mathbf{x}_c$ is sampled before $\mathbf{x}_s$, and for $\Delta r > 0$ errors are largest nearest the radar.

For the representative range of Δr and z_s shown in Fig. A1, $|\Delta t| \leq 240$. If the maximum system ground-relative speed $\|\mathbf{u}\| \leq 20\,\mathrm{m\,s^{-1}}$, then the worst possible error is $|\|\mathbf{s}\| - \|\mathbf{s}'\|| = 4.8\,\mathrm{km}$.

Errors this large are unlikely to occur. A typical system ground-relative speed $\|\mathbf{u}\|$ is $10\,\mathrm{m\,s^{-1}}$ (see Fig. 5) and a typical choice for Δr is 10 km (see Fig. 7), giving $\Delta t \approx 150\,\mathrm{s}$ and $|\|\mathbf{s}\| - \|\mathbf{s}'\|| \leq 1.5\,\mathrm{km}$, recalling the equality is attained if and only if $\mathbf{u}$ and s′ are collinear: if they are at right angles $|\|\mathbf{s}\| - \|\mathbf{s}'\|| \approx 120\,\mathrm{m}$. We therefore estimate a typical stratiform offset magnitude error associated with the scanning pattern is $\approx 1\,\mathrm{km}$, noting this is less than the 2.5 km horizontal grid length of the datasets considered in this study.

The MINT algorithm calculates stratiform offsets from the centroids of ellipses fitted to the convex hulls of convective and stratiform cloud regions (Short et al. 2023), i.e. it uses many points, not just two. Adapting the above argument we can assess how the scanning pattern deforms two-dimensional cloud regions, and how this alters the measured stratiform offset: examples are provided in the online supplement. The basic conclusion, errors of $\approx 1\,\mathrm{km}$, remains the same.

Another potential source of error is how the un-scanned upper levels near the radar, i.e. those corresponding to elevations $> 32°$, affect the observed stratiform cloud regions. However, for the assessment diagrams (see section 2) considered in this study, these un-scanned regions had no effect on stratiform offsets, as MINT fits ellipses to the convex hulls of the detected regions (Short et al. 2023), and stratiform cloud extends low enough that missing cloud regions are always small relative to the overall size of the region: Fig. 3 b) provides an example.

The above ideas could be further investigated with an observing system simulation experiment (e.g. Oue et al. 2019).

References

Argonne National Laboratory, 2022: openradar/TINT. Accessed 01/10/2022, https://github.com/openradar/TINT.

Bergemann, M., T. P. Lane, S. Wales, S. Narsey, and V. Louf, 2022: High-resolution simulations of tropical island thunderstorms: Does an increase in resolution improve the representation of extreme rainfall? *Quarterly Journal of the Royal Meteorological Society*, **148 (748)**, 3303–3318, https://doi.org/10.1002/qj.4360.

Bretherton, C. S., and P. K. Smolarkiewicz, 1989: Gravity waves, compensating subsidence and detrainment around cumulus clouds. *Journal of the Atmospheric Sciences*, **46 (6)**, 740–759, https://doi.org/10.1175/1520-0469(1989)046⟨0740:GWCSAD⟩2.0.CO;2.

Bryan, G. H., J. C. Wyngaard, and J. M. Fritsch, 2003: Resolution requirements for the simulation of deep moist convection. *Monthly Weather Review*, **131 (10)**, 2394–2416, https://doi.org/10.1175/1520-0493(2003)131⟨2394:RRFTSO⟩2.0.CO;2.

Bureau of Meteorology, 2019: NOC operations bulletin number 125. Tech. rep., http://www.bom.gov.au/australia/charts/bulletins/opsbull_G3GE3_external_v3.pdf, Bureau of Meteorology, Melbourne, Victoria. Accessed 11/04/2023.

Bureau of Meteorology, 2022: APS3 ACCESS numerical weather prediction (NWP) models - operational reference data collection. National Computational Infrastructure Australia, accessed 11/10/2022, https://doi.org/10.25914/608a993391647.

Bush, M., and Coauthors, 2020: The first met office unified model–jules regional atmosphere and land configuration, ral1. *Geoscientific Model Development*, **13 (4)**, 1999–2029, https://doi.org/10.5194/gmd-13-1999-2020.

Byers, H. R., and R. R. Braham, 1948: Thunderstorm structure and circulation. *Journal of Atmospheric Sciences*, **5 (3)**, 71–86, https://doi.org/10.1175/1520-0469(1948)005⟨0071:TSAC⟩2.0.CO;2.

Caine, S., T. P. Lane, P. T. May, C. Jakob, S. T. Siems, M. J. Manton, and J. Pinto, 2013: Statistical assessment of tropical convection-permitting model simulations using a cell-tracking algorithm. *Monthly Weather Review*, **141 (2)**, 557–581, https://doi.org/10.1175/MWR-D-11-00274.1.

867 Coniglio, M. C., S. F. Corfidi, and J. S. Kain, 2012: Views on applying rkw theory: An il-
868 lustration using the 8 may 2009 derecho-producing convective system. *Monthly Weather Re-*
869 *view*, **140 (3)**, 1023–1043, https://doi.org/10.1175/MWR-D-11-00026.1, URL https://journals.
870 ametsoc.org/view/journals/mwre/140/3/mwr-d-11-00026.1.xml.

871 Coniglio, M. C., D. J. Stensrud, and L. J. Wicker, 2006: Effects of upper-level shear on the
872 structure and maintenance of strong quasi-linear mesoscale convective systems. *Journal of the*
873 *Atmospheric Sciences*, **63 (4)**, 1231–1252, https://doi.org/10.1175/JAS3681.1, https://doi.org/
874 10.1175/JAS3681.1.

875 Damiani, R., G. Vali, and S. Haimov, 2006: The structure of thermals in cumulus from airborne
876 dual-Doppler radar observations. *Journal of the Atmospheric Sciences*, **63 (5)**, 1432–1450,
877 https://doi.org/10.1175/JAS3701.1.

878 Damiani, R., and Coauthors, 2008: The cumulus, photogrammetric, in situ, and Doppler obser-
879 vations experiment of 2006. *Bulletin of the American Meteorological Society*, **89 (1)**, 57–74,
880 https://doi.org/10.1175/BAMS-89-1-57.

881 Dipankar, A., and Coauthors, 2020: SINGV: A convective-scale weather forecast model for
882 Singapore. *Quarterly Journal of the Royal Meteorological Society*, **146 (733)**, 4131–4146,
883 https://doi.org/10.1002/qj.3895.

884 Dixon, M., and G. Wiener, 1993: Titan: Thunderstorm identification, tracking, analysis, and
885 nowcasting—a radar-based methodology. *Journal of Atmospheric and Oceanic Technology*,
886 **10 (6)**, 785–797, https://doi.org/10.1175/1520-0426(1993)010⟨0785:TTITAA⟩2.0.CO;2.

887 Drosdowsky, W., 1996: Variability of the Australian summer monsoon at Darwin: 1957–
888 1992. *Journal of Climate*, **9 (1)**, 85–96, https://doi.org/10.1175/1520-0442(1996)009⟨0085:
889 VOTASM⟩2.0.CO;2.

890 Emanuel, K., 2020: The relevance of theory for contemporary research in atmospheres, oceans,
891 and climate. *AGU Advances*, **1 (2)**, https://doi.org/10.1029/2019AV000129.

892 Emanuel, K. A., 1986: Some dynamical aspects of precipitating convection. *Journal of At-*
893 *mospheric Sciences*, **43 (20)**, 2183–2198, https://doi.org/10.1175/1520-0469(1986)043⟨2183:
894 SDAOPC⟩2.0.CO;2.

European Centre for Medium-Range Weather Forecasts, 2020: ERA5 hourly data on pressure levels. Copernicus and National Computational Infrastructure, accessed 01/10/2022, https://doi.org/10.25914/5fb115b82e2ba.

Giangrande, S. E., S. Collis, J. Straka, A. Protat, C. Williams, and S. Krueger, 2013: A summary of convective-core vertical velocity properties using ARM UHF wind profilers in Oklahoma. *Journal of Applied Meteorology and Climatology*, **52 (10)**, 2278–2295, https://doi.org/10.1175/JAMC-D-12-0185.1.

Grant, L. D., M. W. Moncrieff, T. P. Lane, and S. C. van den Heever, 2020: Shear-parallel tropical convective systems: Importance of cold pools and wind shear. *Geophysical Research Letters*, **47 (12)**, https://doi.org/10.1029/2020GL087720.

Hanley, K. E., R. S. Plant, T. H. M. Stein, R. J. Hogan, J. C. Nicol, H. W. Lean, C. Halliwell, and P. A. Clark, 2015: Mixing-length controls on high-resolution simulations of convective storms. *Quarterly Journal of the Royal Meteorological Society*, **141 (686)**, 272–284, https://doi.org/10.1002/qj.2356.

Helmus, J., and S. Collis, 2016: The Python ARM radar toolkit (Py-ART), a library for working with weather radar data in the Python programming language. *Journal of Open Research Software*, **4**, https://doi.org/10.5334/jors.119.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146 (730)**, 1999–2049, https://doi.org/10.1002/qj.3803.

Hitchcock, S. M., T. P. Lane, R. A. Warren, and J. S. Soderholm, 2021: Linear rainfall features and their association with rainfall extremes near Melbourne, Australia. *Monthly Weather Review*, https://doi.org/10.1175/MWR-D-21-0007.1.

Houze, J., Robert A., 2004: Mesoscale convective systems. *Reviews of Geophysics*, **42 (4)**, https://doi.org/10.1029/2004RG000150.

Jucker, M., T. P. Lane, C. L. Vincent, S. Webster, S. A. Wales, and V. Louf, 2020: Locally forced convection in subkilometre-scale simulations with the Unified Model and WRF. *Quarterly Journal of the Royal Meteorological Society*, **146 (732)**, 3450–3465, https://doi.org/10.1002/qj.3855.

Kawashima, M., 2003: The role of gravity waves in the meso-β-scale cycle of squall-line type convective systems. *Journal of the Meteorological Society of Japan. Ser. II*, **81 (4)**, 713–746, https://doi.org/10.2151/jmsj.81.713.

Keene, K. M., and R. S. Schumacher, 2013: The bow and arrow mesoscale convective structure. *Monthly Weather Review*, **141 (5)**, 1648–1672, https://doi.org/10.1175/MWR-D-12-00172.1.

Lafore, J.-P., and M. W. Moncrieff, 1989: A numerical investigation of the organization and interaction of the convective and stratiform regions of tropical squall lines. *Journal of the Atmospheric Sciences*, **46 (4)**, 521–544, https://doi.org/10.1175/1520-0469(1989)046⟨0521:ANIOTO⟩2.0.CO;2.

Lane, T. P., and M. W. Moncrieff, 2015: Long-lived mesoscale systems in a low–convective inhibition environment. Part I: upshear propagation. *Journal of the Atmospheric Sciences*, **72 (11)**, 4297–4318, https://doi.org/10.1175/JAS-D-15-0073.1.

Lane, T. P., and F. Zhang, 2011: Coupling between gravity waves and tropical convection at mesoscales. *Journal of the Atmospheric Sciences*, **68 (11)**, 2582–2598, https://doi.org/10.1175/2011JAS3577.1.

Lean, H. W., P. A. Clark, M. Dixon, N. M. Roberts, A. Fitch, R. Forbes, and C. Halliwell, 2008: Characteristics of high-resolution versions of the Met Office Unified Model for forecasting convection over the United Kingdom. *Monthly Weather Review*, **136 (9)**, 3408–3424, https://doi.org/10.1175/2008MWR2332.1.

LeMone, M. A., and E. J. Zipser, 1980: Cumulonimbus vertical velocity events in GATE. Part I: diameter, intensity and mass flux. *Journal of Atmospheric Sciences*, **37 (11)**, 2444–2457, https://doi.org/10.1175/1520-0469(1980)037⟨2444:CVVEIG⟩2.0.CO;2.

Louf, V., and A. Protat, 2020: CPOL dual-polarisation C-band research radar - level-1b v2020. National Computational Infrastructure and Bureau of Meteorology, https://doi.org/10.25914/5f4c857de1bb7.

Louf, V., A. Protat, R. A. Warren, S. M. Collis, D. B. Wolff, S. Raunyiar, C. Jakob, and W. A. Petersen, 2019: An integrated approach to weather radar calibration and monitoring using

ground clutter and satellite comparisons. *Journal of Atmospheric and Oceanic Technology*, **36 (1)**, 17–39, https://doi.org/10.1175/JTECH-D-18-0007.1.

Mapes, B. E., T. T. Warner, and M. Xu, 2003: Diurnal patterns of rainfall in northwestern South America. Part III: Diurnal gravity waves and nocturnal convection offshore. *Monthly Weather Review*, **131 (5)**, 830, https://doi.org/10.1175/1520-0493(2003)131⟨0830:DPORIN⟩2.0.CO;2.

McBeath, K., P. R. Field, and R. J. Cotton, 2014: Using operational weather radar to assess high-resolution numerical weather prediction over the british isles for a cold air outbreak case-study. *Quarterly Journal of the Royal Meteorological Society*, **140 (678)**, 225–239, https://doi.org/10.1002/qj.2123.

Moncrieff, M. W., 1978: The dynamical structure of two-dimensional steady convection in constant vertical shear. *Quarterly Journal of the Royal Meteorological Society*, **104 (441)**, 543–567, https://doi.org/10.1002/qj.49710444102.

Moncrieff, M. W., 1981: A theory of organized steady convection and its transport properties. *Quarterly Journal of the Royal Meteorological Society*, **107 (451)**, 29–50, https://doi.org/10.1002/qj.49710745103.

Moncrieff, M. W., 1992: Organized convective systems: Archetypal dynamical models, mass and momentum flux theory, and parametrization. *Quarterly Journal of the Royal Meteorological Society*, **118 (507)**, 819–850, https://doi.org/10.1002/qj.49711850703.

Moncrieff, M. W., and J. S. A. Green, 1972: The propagation and transfer properties of steady convective overturning in shear. *Quarterly Journal of the Royal Meteorological Society*, **98 (416)**, 336–352, https://doi.org/10.1002/qj.49709841607.

Moncrieff, M. W., and T. P. Lane, 2015: Long-lived mesoscale systems in a low–convective inhibition environment. Part II: Downshear propagation. *Journal of the Atmospheric Sciences*, **72 (11)**, 4319–4336, https://doi.org/10.1175/JAS-D-15-0074.1.

Moncrieff, M. W., and C. Liu, 1999: Convection initiation by density currents: Role of convergence, shear, and dynamical organization. *Monthly Weather Review*, **127 (10)**, 2455–2464, https://doi.org/10.1175/1520-0493(1999)127⟨2455:CIBDCR⟩2.0.CO;2.

977 Moncrieff, M. W., and M. J. Miller, 1976: The dynamics and simulation of tropical cumulonimbus
978 and squall lines. *Quarterly Journal of the Royal Meteorological Society*, **102 (432)**, 373–394,
979 https://doi.org/10.1002/qj.49710243208.

980 Moncrieff, M. W., and D. W. K. So, 1989: A hydrodynamical theory of conservative
981 bounded density currents. *Journal of Fluid Mechanics*, **198**, 177–197, https://doi.org/10.1017/
982 S0022112089000091.

983 Oue, M., P. Kollias, A. Shapiro, A. Tatarevic, and T. Matsui, 2019: Investigation of obser-
984 vational error sources in multi-Doppler-radar three-dimensional variational vertical air mo-
985 tion retrievals. *Atmospheric Measurement Techniques*, **12 (3)**, 1999–2018, https://doi.org/
986 10.5194/amt-12-1999-2019.

987 Parker, M. D., and R. H. Johnson, 2000: Organizational modes of midlatitude mesoscale
988 convective systems. *Monthly Weather Review*, **128 (10)**, 3413–3436, https://doi.org/10.1175/
989 1520-0493(2001)129⟨3413:OMOMMC⟩2.0.CO;2.

990 Pope, M., C. Jakob, and M. J. Reeder, 2009: Regimes of the north Australian wet season. *Journal*
991 *of Climate*, **22 (24)**, 6699–6715, https://doi.org/10.1175/2009JCLI3057.1.

992 Raut, B. A., R. Jackson, M. Picel, S. M. Collis, M. Bergemann, and C. Jakob, 2021: An adaptive
993 tracking algorithm for convection in simulated and remote sensing data. *Journal of Applied*
994 *Meteorology and Climatology*, **60 (4)**, 513–526, https://doi.org/10.1175/JAMC-D-20-0119.1.

995 Rennie, S., and Coauthors, 2022: ACCESS-C: Australian convective-scale NWP with hourly 4D-
996 Var data assimilation. *Weather and Forecasting*, **37 (7)**, 1287–1303, https://doi.org/10.1175/
997 WAF-D-21-0183.1.

998 Roberts, N. M., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations
999 from high-resolution forecasts of convective events. *Monthly Weather Review*, **136 (1)**, 78–97,
1000 https://doi.org/10.1175/2007MWR2123.1.

1001 Rotunno, R., J. B. Klemp, and M. L. Weisman, 1988: A theory for strong, long-lived squall lines.
1002 *Journal of the Atmospheric Sciences*, **45 (3)**, 463–485, https://doi.org/10.1175/1520-0469(1988)
1003 045⟨0463:ATFSLL⟩2.0.CO;2.

Schumacher, R. S., and R. H. Johnson, 2006: Characteristics of U.S. extreme rain events during 1999–2003. *Weather and Forecasting*, **21 (1)**, 69–85, https://doi.org/10.1175/WAF900.1.

Seitter, K. L., and H.-L. Kuo, 1983: The dynamical structure of squall-line type thunderstorms. *Journal of Atmospheric Sciences*, **40 (12)**, 2831–2854, https://doi.org/10.1175/1520-0469(1983)040⟨2831:TDSOSL⟩2.0.CO;2.

Short, E., 2022: eshort0401/TINT. Accessed 01/10/2022, https://github.com/eshort0401/TINT.

Short, E., T. P. Lane, and C. L. Vincent, 2023: Objectively diagnosing characteristics of mesoscale organization from radar reflectivity and ambient winds. *Monthly Weather Review*, **151 (3)**, 643–662, https://doi.org/10.1175/MWR-D-22-0146.1.

Skamarock, W. C., 2004: Evaluating mesoscale NWP models using kinetic energy spectra. *Monthly Weather Review*, **132 (12)**, 3019–3032, https://doi.org/10.1175/MWR2830.1.

Soderholm, J., and A. Protat, 2022: AURA - operational radar network level 1 archive. National Computational Infrastructure and Bureau of Meteorology, accessed 01/10/2022, https://doi.org/10.25914/5cb686a8d9450.

Stein, T., R. Scovell, K. Hanley, H. Lean, and N. 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**, https://doi.org/10.1002/qj.3793.

Stein, T. H. M., R. J. Hogan, K. E. Hanley, J. C. Nicol, H. W. Lean, R. S. Plant, P. A. Clark, and C. E. Halliwell, 2014: The three-dimensional morphology of simulated and observed convective storms over Southern England. *Monthly Weather Review*, **142 (9)**, 3264–3283, https://doi.org/10.1175/MWR-D-13-00372.1.

Steiner, M., R. A. Houze, and S. E. Yuter, 1995: Climatological characterization of three-dimensional storm structure from operational radar and rain gauge data. *Journal of Applied Meteorology*, **34 (9)**, 1978–2007, https://doi.org/10.1175/1520-0450(1995)034⟨1978:CCOTDS⟩2.0.CO;2.

University of Wyoming, 2022: Atmospheric soundings. Accessed 04/10/2022, https://weather.uwyo.edu/upperair/sounding.html.

Walters, D., and Coauthors, 2019: The met office unified model global atmosphere 7.0/7.1 and jules global land 7.0 configurations. *Geoscientific Model Development*, **12 (5)**, 1909–1963, https://doi.org/10.5194/gmd-12-1909-2019.

Wang, D., S. E. Giangrande, Z. Feng, J. C. Hardin, and A. F. Prein, 2020: Updraft and downdraft core size and intensity as revealed by radar wind profilers: Mcs observations and idealized model comparisons. *Journal of Geophysical Research: Atmospheres*, **125 (11)**, https://doi.org/10.1029/2019JD031774.

Weisman, M. L., 1993: The genesis of severe, long-lived bow echoes. *Journal of the Atmospheric Sciences*, **50 (4)**, 645–670, https://doi.org/10.1175/1520-0469(1993)050⟨0645:TGOSLL⟩2.0.CO;2.

Weisman, M. L., J. B. Klemp, and R. Rotunno, 1988: Structure and evolution of numerically simulated squall lines. *Journal of the Atmospheric Sciences*, **45 (14)**, 1990–2013, https://doi.org/10.1175/1520-0469(1988)045⟨1990:SAEONS⟩2.0.CO;2.

Weisman, M. L., and R. Rotunno, 2004: "A theory for strong long-lived squall lines" revisited. *Journal of the Atmospheric Sciences*, **61 (4)**, 361–382, https://doi.org/10.1175/1520-0469(2004)061⟨0361:ATFSLS⟩2.0.CO;2.

Weisman, M. L., W. C. Skamarock, and J. B. Klemp, 1997: The resolution dependence of explicitly modeled convective systems. *Monthly Weather Review*, **125 (4)**, 527–548, https://doi.org/10.1175/1520-0493(1997)125⟨0527:TRDOEM⟩2.0.CO;2.