Characterizing the Mesoscale Cellular Convection in Marine Cold Air Outbreaks with a Machine Learning Approach

Christian. P. Lackner[1], Bart Geerts[1], Timothy W. Juliano[2], Branko Kosovic[2] and Lulin Xue[2]

[1]Department of Atmospheric Science, University of Wyoming, Laramie, WY, USA.

[2]Research Applications Laboratory, National Center for Atmospheric Research, Boulder, CO, USA.

Corresponding author: Christian P. Lackner (clackner@uwyo.edu)

Key Points:

- Cloud cells in marine cold-air outbreaks are objectively identified and classified using a profiling mm-wave radar
- The classification reveals that open-cellular clouds undergo a lifecycle with distinct characteristics at each lifecycle stage
- Closed-cellular clouds contain occasional convection but generally have more stratiform characteristics

Abstract

During marine cold-air outbreaks (MCAOs), when cold polar air moves over warmer ocean, a well-recognized cloud pattern develops, with open or closed mesoscale cellular convection (MCC) at larger fetch over open water. The Cold-Air Outbreaks in the Marine Boundary Layer Experiment (COMBLE) provided a comprehensive set of ground-based in-situ and remote sensing observations of MCAOs at a coastal location in northern Norway. We determine MCAO periods that unambiguously exhibit open or closed MCC. Individual cells observed with a profiling Ka-band radar are identified using a water segmentation method. Using self-organizing maps (SOMs), these cells are then objectively classified based on the variability in their vertical structure. The SOM-based classification shows that comparatively intense convection occurs only in open MCC. This convection undergoes an apparent lifecycle. Developing cells are associated with stronger updrafts, large spectral width, larger amounts of liquid water, lower precipitation rates, and lower cloud tops than mature and weakening cells. The weakening of these cells is associated with the development of precipitation-induced cold pools. The SOM classification also reveals less intense convection, with a similar lifecycle. Such convection, when weakening, becomes virtually indistinguishable from the more intense stratiform precipitation cores in closed MCC. Non-precipitating stratiform cores have weak vertical drafts and are almost exclusively found during closed MCC periods. Convection is observed only occasionally in the closed MCC environment.

Plain Language Summary

During marine cold-air outbreaks (MCAOs) a characteristic cloud field develops over the open ocean. At large fetch, this cloud field is characterized by open and closed mesoscale cellular convection (MCC). The Cold-Air Outbreaks in the Marine Boundary Layer Experiment (COMBLE) in northern Norway provided comprehensive observations of MCC using ground-based and remote sensing instruments. Distinct open or closed MCC periods are identified using a vertically scanning cloud radar. Within these periods individual cells are identified. Using a machine learning algorithm, these cells are objectively classified based on their vertical structure as observed by the radar. The classification reveals that intense convection primarily occurs in open MCC, displaying a lifecycle with developing cells characterized by strong updrafts, substantial liquid water, lower precipitation rates, and lower cloud tops compared to mature and weakening cells. Weakening cells are associated with precipitation-induced cold pools. Less intense convection with a similar lifecycle is observed during open and closed MCC. Most frequently during closed MCC, stratiform cells with weak vertical motions are observed. These are classified as non-precipitating and precipitating cells. The vertical cloud structure of the precipitating stratiform cells is very similar to weakening convective cells.

1 Introduction

The Arctic is an integral part of the global climate system. It has been observed that the Arctic is warming at a rate 2-4 times faster than the rest of the globe (e.g., Serreze & Francis, 2006; Serreze et al., 2009; Rantanen et al., 2022), a trend that models predict will continue in the future (Davy & Outten, 2020). This well-documented trend is referred to as Arctic amplification. This amplification is attributed to a variety of feedback processes. Among them are Arctic cloud feedbacks (Taylor et al., 2013), for which the magnitude and sign of the feedback remains uncertain (IPCC, 2021, p. 974).

Here, we examine one specific high-latitude cloud regime, one that occurs in marine cold-air outbreaks (MCAOs). Such synoptic events produce unique cloud patterns that through vertical mixing and radiative effects may play an important role in high-latitude cloud feedbacks (Fletcher et al., 2016). When Arctic air masses move from the sea ice (or boreal continents) over the open ocean, large surface heat fluxes develop that destabilize the boundary layer (Pithan et al., 2018). This, combined with intense surface moisture fluxes, leads to shallow moist convection that deepens with fetch. Initially, as the air moves over the ocean, linear cloud patterns form, referred to as cloud streets, and at a larger fetch downstream, the cloud streets transition to open or closed mesoscale cellular convection (MCC) (e.g., Atkinson & Zhang, 1996; Brümmer, 1999; Brümmer & Pohlmann, 2000). At high latitudes, these clouds are generally mixed-phase (e.g., Geerts et al. 2022).

Understanding the differences in the characteristics of open and closed MCC is important since their differing cloud morphology and microphysical properties impact the radiative properties of clouds as well as surface fluxes (e.g., Agee, 1987). Open and closed MCC have been studied extensively in the subtropics and the mid-latitudes (e.g., Eastman et al., 2021, 2022; Jensen et al., 2021; McCoy et al., 2017, 2023; Mohrmann et al., 2021; Muhlbauer et al., 2014; Wood & Hartmann, 2006). These studies show that a variety of environmental factors, such as wind, precipitation, surface forcing, and aerosol can influence the development of open or closed MCC. McCoy et al. (2017) find that the occurrence of open and closed MCC during MCAOs is correlated with the MCAO-index M, a measure of low-level static stability driven by the sea surface ($M = \theta_{SST} - \theta_{850hPa}$). Mesoscale cloud morphology is strongly influenced by surface heat fluxes and low-level stability: environments with large positive M values favor open MCC regimes, while under low M values, a closed MCC regime prevails, with more stratiform clouds (McCoy et al., 2017). This has potential implications for the cloud feedbacks in a globally warming climate since the intensity of MCAOs is predicted to weaken in the future (Kolstad & Bracegirdle, 2008; Landgren et al., 2019).

A comprehensive dataset of ground based observations of MCAO clouds was collected as part of the Cold-Air Outbreaks in the Marine Boundary Layer Experiment in the 2019/20 cold season (COMBLE, Geerts et al., 2022). Geerts et al. (2022) and Lackner et al. (2023a) illustrate MCAOs with episodes of rather deep, intense open-cellular convection observed during COMBLE. These examples demonstrate a relationship between vertical velocity, supercooled liquid water path (LWP), and surface temperature anomalies: cells with strong updrafts tend to have large LWP values, while others have weak vertical motions, low LWP values, and a surface cold pool. These differences probably can be attributed to different cloud lifecycle stages. Geerts et al. (2022) and Lackner et al. (2023a) also illustrate closed MCC observed during COMBLE. The closed MCC periods are characterized by a lower M value and weaker winds, and clouds exhibit continuous cloud cover, weaker vertical motions, and lower average cloud top heights when compared to the open MCC. The objective of this study is to comprehensively characterize the vertical structure of MCC using the spectrum of MCAOs observed during the 6-month COMBLE campaign, to better understand the linked MCAO cloud dynamics and microphysics, i.e. the linkages between cell vertical structure, its lifecycle, and its mesoscale organization.

This study is structured as follows. Section 2 describes the COMBLE campaign, the periods with open and closed MCC, and the methods used to identify and classify individual cells. The results of the classification of cells and an analysis of additional measurements is presented in Section 3. A discussion of the results is provided in Section 4 and the study is summarized in Section 5.

2 Data Sources and Analysis Methods

2.1 The COMBLE Campaign

The data used in this study were collected as part of COMBLE, which deployed the ARM mobile facility 1 (AMF1, Miller et al., 2016) to Nordmela harbor (69.141 ºN; 15.684 ºE) on the Norwegian island Andoya from December 2019 to May 2020. A detailed summary of the COMBLE campaign can be found in Geerts et al. (2022), and a description of the datasets can be found in Lackner et al. (2023a). The key instrument used for analysis is the Ka-band ARM zenith radar (KAZR), providing radar observations of vertical cloud structures. Additionally, the *MICROBASEKAPLUS* (Zhao et al., 2014) and the *THERMOCLDPHASE* (Van Weverberg et al., 2023) products are utilized. All datasets are summarized in Table 1.

The *MICROBASEKAPLUS* product provides radar retrieved ice water content, which is used to calculate ice water path (IWP). We found the IWP values to be exceptionally high for the more intense, 4-5 km deep convective cells. Following a recommendation from the ARM instrument mentors (S. Giangrande, personal communication, March 28, 2023), we reduced the KAZR reflectivity uniformly by 5 dBZ, to achieve more realistic IWP values. The *THERMOCLDPHASE* product provides a vertically resolved classification of cloud phase. This product uses KAZR and micropulse lidar data, and is based on the algorithm developed in Shupe (2007). The micropulse lidar at the AMF1 site was only deployed starting 11 February 2020. Thus, *THERMOCLDPHASE* is only available for some of the periods described in Section 2.2.

Table 1. COMBLE AMF1 Data Sets Used in This Study

ARM Data Products	Description	Variable(s)	Units
MET (Kyrouac et al., 2019; Ritsche & Prell, 2011)	Surface meteorological instrumentation	Atmospheric temperature Horizontal wind direction Horizontal wind speed Atmospheric pressure Precipitation rate	°C ° m s^{-1} hPa mm hr^{-1}
MAWS (Keeler et al., 2019)	Vaisala automatic weather station	Atmospheric temperature Horizontal wind direction Horizontal wind speed Atmospheric pressure	°C ° m s^{-1} hPa
ARSCLKAZR1KOLLIAS (Clothiaux et al., 2001; Johnson & Jensen, 2019)	KAZR ARSCL.c1: multiple outputs from first Kollias algorithm	Equivalent reflectivity factor Mean Doppler velocity Spectral width	dBZ m s^{-1} m s^{-1}
ARSCLKAZRBND1KOLLIAS (Johnson et al., 2019)	Cloud boundaries retrieved from KAZRARSCL	Cloud top height Cloud base height	m m
MWRRET1LILJCLOU (Turner et al., 2007; Zhang, 2019)	Microwave radiometer retrievals	Liquid water path	kg m^{-2}

INTERPOLATEDSONDE (Fairless et al., 2021; Jensen et al., 2019)	Sounding data interpolated to 1 min	Atmospheric temperature Atmospheric pressure Horizontal wind speed	°C hPa m s^{-1}
MICROBASEKAPLUS (Dunn et al., 2011; Wang et al., 2019; Zhao et al., 2014)	Improved MICROBASE product with uncertainties	Ice water content	kg m^{-3}
THERMOCLDPHASE (Van Weverberg et al., 2023; Zhang & Levin, 2019)	Thermodynamic cloud phase classifications	Classifications: clear sky, liquid, ice, mixed phase, snow, rain, drizzle, liquid + drizzle, unknown	---

Note: MICROBASE is the Atmospheric Radiation Measurement (ARM) Climate Research Facility baseline cloud microphysical properties (MICROBASE) value-added product.

2.2 Periods with Mesoscale Cellular Convection

Periods exhibiting open and closed MCC are determined through visual inspection of KAZR reflectivity time-height transects (see Figure 1) and satellite imagery. These periods of MCC are only determined during and shortly before or after the MCAO events at the AMF1 site, as defined in Geerts et al. (2022) and Lackner et al. (2023a). Short periods of MCC (< 3 hours), periods with disorganized clouds, periods with polar lows (Lackner et al., 2023b), and periods during which MCC was connected to mid/upper-tropospheric clouds with a different origin are not considered in the analysis. An example of the distinct vertical structure of open (closed) MCC in the KAZR data is shown in Figure 1a-c (1e-g). The closed-cell example appears to be mostly stratiform, with light precipitation often not reaching the ground (Figure 1e). Consistent with the literature, we still refer to these cloud patches as mesoscale cellular *convection* (MCC), even though the convective origin may not be apparent.

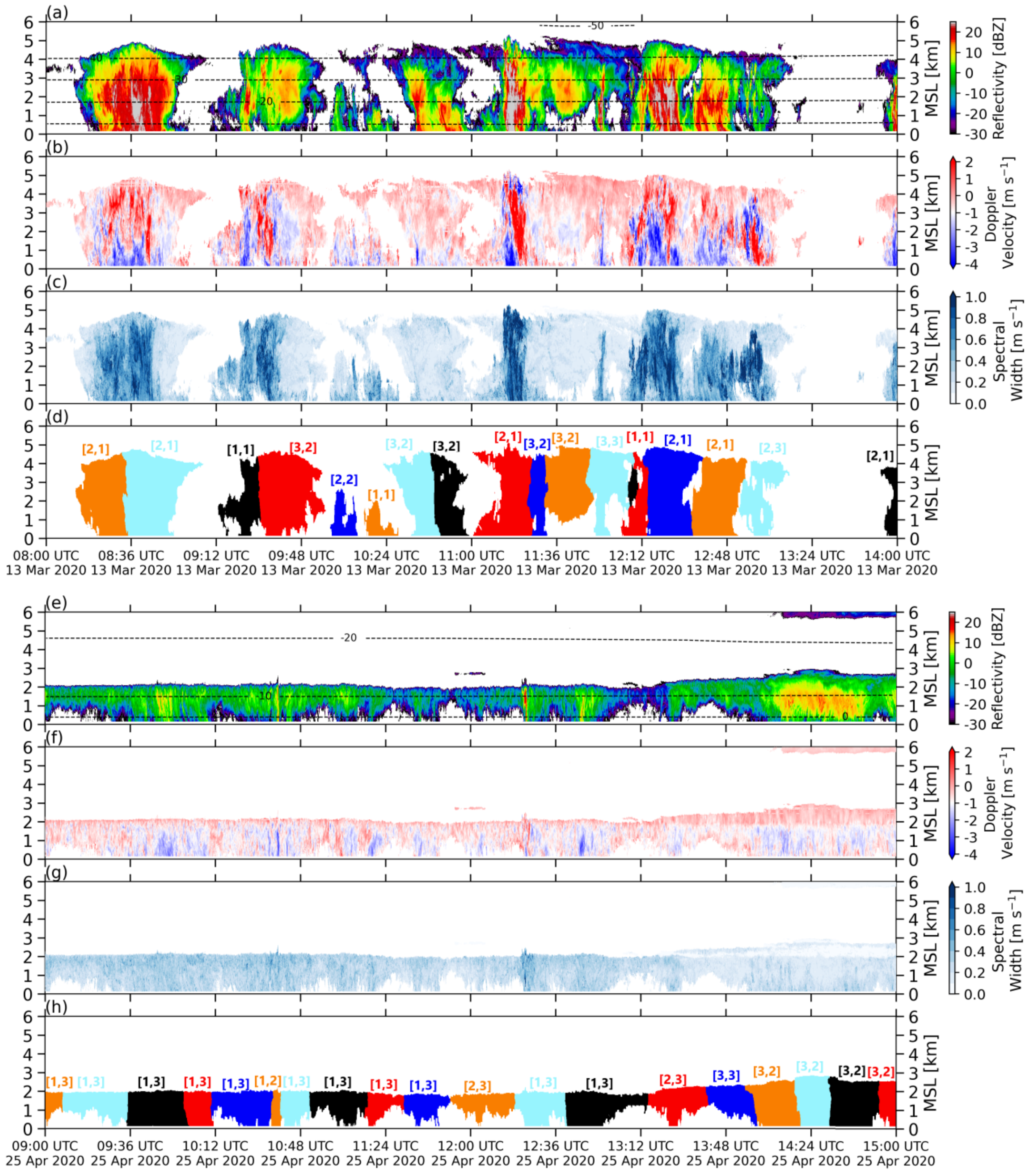

Figure 1. KAZR time-height transect for an open MCC period (a-d) and a closed MCC period (e-h) showing (a,e) KAZR reflectivity factor and temperature contours, (b,f) Doppler velocity, (c,g) spectral width, and (d,h) objectively identified cells via the watershed segmentation method (see Section 2.3). Note that the 5 colors in (d,h) merely are recycled and do not indicate any cell grouping. The numbers on top of each cell indicate the best-matched unit, introduced in Section 2.4.

Table 2 summarizes all MCC periods used in this study. The distinction between open and closed cells is based on KAZR data and satellite imagery and is unambiguous. A total of 13 periods with a total of ~249 hours of open MCC were identified at the AMF1 site, compared to 4 periods with a total of ~57 hours of closed MCC. In total, this accounts for ~37 % of the MCAO periods during COMBLE identified in Geerts et al. (2022) (which themselves cover 19% of the full 6-month field phase), mostly because of the previously mentioned reasons for which certain periods are not considered in the analysis. A broad spectrum of environmental conditions prevailed during these 17 periods. The *M* value is above the COMBLE MCAO mean of 4.1 K for most open MCC periods (10 of 13), indicating that open MCC tends to occur during more intense cold air outbreaks. Closed MCC on the other hand, tends to occur during weaker outbreaks, with mean *M* (1.2 K) being much lower than for open MCC (6.0 K). This difference in intensity likely explains why open MCC is observed more frequently than closed MCC at the AMF1 site: weaker MCAOs often lack the synoptic support to cover the large distance between the ice edge and the COMBLE observation site (~1000 km). Near-surface temperatures (winds) tend to be warmer (weaker) during closed MCC (Table 2). Mean cloud top heights (CTH) are typically higher during open MCC ranging from 2.2 – 3.3 km, compared to 1.9 – 2.4 km for closed MCC. However, it should be noted that CTHs are much more variable during individual open MCC periods and may exceed 4 km (see Figure 1).

Table 2. Periods of MCC used in this study, and mean environmental conditions during these MCC periods at the AMF1 site during COMBLE (Dec 2019 – May 2020).

Start Time	End Time	Type	Time	*M*	*T*	*U*	*WD*	*CTH*
			[min]	[K]	[°C]	[m s^{-1}]	[°]	[km]
01 Dec 2100 UTC	02 Dec 0330 UTC	Open	390	4.9	2.3	13.3	323.8	2.7
02 Dec 0700 UTC	02 Dec 1530 UTC	Open	510	4.1	2.2	12.1	323.3	2.5
31 Dec 0000 UTC	31 Dec 1915 UTC	Open	1155	3.7	2.1	10.5	285.6	2.5
04 Jan 0830 UTC	05 Jan 0200 UTC	Open	1050	6.3	0.3	10.6	329.6	2.6
21 Jan 2220 UTC	22 Jan 1830 UTC	Open	1210	3.1	2.3	9.2	328.3	2.4
02 Feb 1330 UTC	03 Feb 0615 UTC	Open	1005	5.0	0.8	10.9	319.8	2.2
03 Feb 1040 UTC	03 Feb 1845 UTC	Open	485	6.3	-1.0	7.3	307.6	2.6
04 Feb 1730 UTC	06 Feb 1545 UTC	Open	2775	7.2	-1.0	11.5	308.3	2.7
13 Mar 0800 UTC	14 Mar 0540 UTC	Open	1300	8.0	-3.0	9.3	336.1	3.3
28 Mar 0320 UTC	30 Mar 0945 UTC	Open	3265	7.1	-2.2	10.2	321.6	3.0
09 Apr 1420 UTC	10 Apr 0200 UTC	Open	700	4.8	-0.3	8.3	289.7	2.7
10 Apr 0810 UTC	10 Apr 1620 UTC	Open	490	5.0	0.3	6.8	303.7	2.8
10 Apr 1940 UTC	11 Apr 0530 UTC	Open	590	4.7	0.2	7.5	277.8	2.5
average		Open	14930	6.0	-0.4	10.1	314.4	2.7
17 Apr 0530 UTC	17 Apr 1630 UTC	Closed	660	1.1	3.3	8.8	290.1	1.9
24 Apr 1700 UTC	26 Apr 0010 UTC	Closed	1870	1.6	3.2	8.5	314.4	2.4
26 Apr 0300 UTC	26 Apr 1050 UTC	Closed	470	0.9	2.5	8.1	27.2	2.3
04 May 1050 UTC	04 May 1730 UTC	Closed	400	0.2	4.6	6.7	287.3	2.3
average		Closed	3400	1.2	3.3	8.3	316.6	2.3

Note: *M*: MCAO-index; *T*: 2-m Temperature; *U*: 10-m wind speed; *WD*: 10-m wind direction; *CTH*: Cloud top height. The time in the summary rows (‘average’) is the total cumulative time.

2.3 Cell Identification

To examine the characteristics of open and closed MCC, individual clouds are identified by applying a marker-based watershed segmentation to the KAZR reflectivity field (Vincent & Soille, 1991). This identification is necessary for the classification of the individual clouds described in Section 2.4. For the rest of the study, these individually identified clouds or cloud sections will be referred to as cells.

Watershed segmentation is a segmentation technique with its origins in image processing that has also found use in analyzing atmospheric data (e.g., Martini et al., 2014; Wu & Ovchinnikov, 2022). In the case of a marker-based watershed segmentation, all points in a dataset that are assigned to a predetermined marker comprise an identified object. Similar to how geographical locations are part of a water drainage basin (watershed), data points are assigned to the marker to which they have the steepest downward gradient.

For this study, the predetermined markers are local maxima in the smoothed KAZR reflectivity field. A Gaussian filter with a standard deviation of 1 is used for the smoothing. The smoothing and each following step use linear units of reflectivity (Z) instead of logarithmic units (dBZ). The local maxima are determined with two requirements: their value must exceed 2.5 dBZ (converted to Z) and they must be at least 80 KAZR profiles away in time from a larger value (320 seconds at 4 second time resolution). The corresponding minimum width depends on the advecting wind speed, e.g. 3.2 km for a 10 m s^{-1} layer-mean wind. Other requirements and thresholds were tested, but this setup was found to achieve the least amount of over- and under-segmentation of cells in a visual analysis. However, none of the tested setups were able to completely avoid bizarre segmentation outcomes. Clear segmentation errors are corrected by manually removing (identifying) specific maxima associated with over (under) segmented cells. After the local maxima are determined, the watershed segmentation is conducted on the inverted reflectivity field, since the algorithm determines objects based on downward gradients. Cells are confined to the -30 dBZ reflectivity threshold, which exceeds the KAZR sensitivity even at a largest relevant range (~5 km). Each object identified by the watershed segmentation is an individual cell. Lastly, small cells (< 45 KAZR profiles = 3 minutes) and cells with shallow cloud depth (< 50 KAZR vertical levels between the lowest and highest echo, or 1500 m) are ignored since they have too few data points for the method used to classify cells in Section 2.4. Furthermore, cells with high radar echo bases (> 1500 m above sea level) are ignored since these were typically not MCC but other types of clouds. In total, 780 cells are identified during the 17 periods, 612 (168) of which are during the open (closed) MCC periods, i.e. 78%. Examples of identified cells in a vertical transect with open MCC and closed MCC are shown in Figure 1d, h. Note that continuous cloud structures can be identified as multiple cells. Specifically, in the case of the closed MCC periods, the cells should not be seen as distinct clouds but rather as a continuous cloud field that is made up of separate precipitation cores.

2.4 Cell Classification

The next step is to classify the identified cells based on their vertical structure of KAZR reflectivity, Doppler velocity (i.e., hydrometeor vertical motion), and spectral width. Since the freezing level is usually at or very close to sea level, the radar echoes are dominated by snow. Spectral broadening in snow is generally largely due to atmospheric turbulence (Doviak & Zrnic, 1993). For instance, the red-colored cell around 11:20 UTC (Figure 1a-d) is marked by strong convective updrafts and high spectral width, due to high buoyantly-generated turbulent kinetic energy. The four closed MCC periods all occurred in late April – early May with freezing levels a few hundreds of meters above the surface, which impacts the profiles of all three radar moments.

For the purpose of classifying the 780 cells, we ignore the horizontal cloud structure (i.e., the time dimension of KAZR profiles, shown in Figure 1). Horizontal structure is highly dependent on deep-layer wind shear in the direction of cell motion, which is of little relevance to understanding the relation between cell vertical structure and its lifecycle. Instead, the vertical distribution of the three radar moments, i.e. the contoured frequency-by-altitude diagrams (CFADs) of reflectivity, Doppler velocity, and spectral width, provide a more standardized representation of the individual cells.

An unsupervised machine learning algorithm referred to as the self-organizing map (SOM) is utilized for this CFAD-based classification (Kohonen, 1982). The main objective of the SOM is to represent archetypical patterns in datasets using an interconnected node structure, with the user defining the structure arrangement. The basis of the SOM algorithm is the following equation, which updates the SOM nodes $\boldsymbol{m}_{i,j}$ in each training step t:

$$\boldsymbol{m}_{i,j}\,(t+1) = \boldsymbol{m}_{i,j}\,(t) + a(t) \times n(t) \times [\boldsymbol{x}(t) - \boldsymbol{m}_{i,j}\,(t)] \tag{1}$$

The subscripts i and j denote the location of the node in the grid and $\boldsymbol{x}(t)$ is the input vector. The learning rate $a(t)$ and neighborhood function $n(t)$ determine how much a node is updated in a training step. Different functions were tested and the ones that minimize the quantization (Kohonen, 2001) and topographic errors (Kiviluoto, 1996) of the SOM were selected. The learning rate $a(t)$ at each training step is the following function:

$$a(t) = \frac{a(t=\ 0)}{1 + \frac{5t}{t_n}} \tag{2}$$

The total number of training steps is denoted by t_n. The neighborhood function is a Gaussian distribution with a standard deviation $\sigma(t)$, which follows the same decrease at each training step as Equation 2. At the first training step, $a(t=\ 0)$ = 0.2 and $\sigma(t=\ 0)$ = 3. The center of $n(t)$ is always located on the node that has the smallest Euclidian distance to the input vector $\boldsymbol{x}(t)$. This node is referred to as the *best match unit* (BMU). Thus, values of $n(t)$ decrease the further a node is away from the BMU, leading to the arrangement of the SOM nodes by similarity. The decrease of $\sigma(t)$ with each training step continually reduces the value of $n(t)$ at nodes other than the BMU.

The input vector $\boldsymbol{x}(t)$ contains the data that is analyzed by the SOM algorithm, in this case the CFADs of all three KAZR moments, for a specific cell. The CFAD vertical structure are normalized by CTH (a cell's highest echo top), to focus the algorithm's outcome on the distribution of the radar variables within cloud, and not the variability of CTHs (which otherwise would dominate the SOM nodes). These CFADs provide a standardized and statistically comparable representation of each cell. Note that while cell width and CTH do not define the SOM nodes, they will be examined as we characterize each node.

Each training step of the SOM consists of calculating Equation 1 for all 780 cells at each node. In total, 10,000 training steps are conducted; in each training step, cells are selected in a random order. For this study, a grid of 3 by 3 nodes is selected. This setup produced nodes with distinct composite CFADs. Setups with more nodes showed little difference between some nodes, while setups with fewer nodes merges what we believe are distinct nodes. The final SOM resulting from the training process is referred to as the master SOM. The nodes of the master SOM are used to classify cells. Each node of the master SOM is characterized, first in terms of its composite radar moment vertical structure (e.g., reflectivity, shown in Figure 2) (Section 3.1), then in terms of cloud macroscale parameters (Section 3.2), and finally in terms of cloud microphysical and dynamical parameters (Section 3.3). Through this analysis, insights into the dynamical and microphysical characteristics of each node will be gained, and each node will be given a simple identifier. Note that the SOM technique treats each cell independently. It ignores the temporal succession of cells or cell nodes as illustrated in Fig. 1d, h, and thus any interaction that may occur between cells.

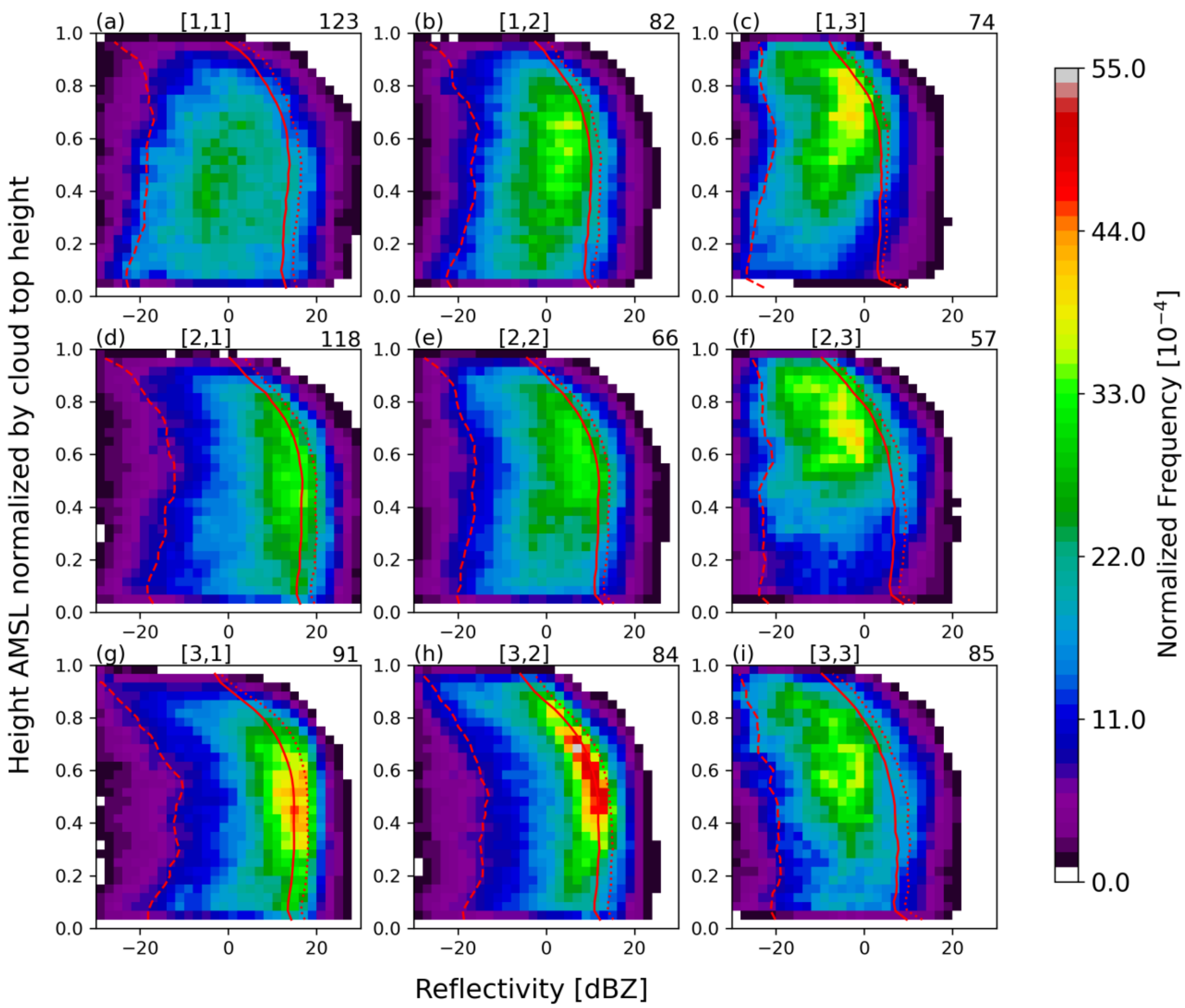

Figure 2. Reflectivity part of the master SOM. Each node (labeled as [row #, column #]) shows the normalized CFADs of reflectivity composited for all the node's members (cells). The numbers on the top right of each node indicate for how many cells (out of a total of 780) each node is the

BMU. The red dashed line indicates the 10th percentile, the solid line the mean, and the red dotted line the 90th percentile.

3 Results

3.1 Cell Classification

In this section, the individual nodes of the master SOM are used to classify the MCC cells observed during the identified periods (Table 2). For a comparison of the differences of the nodes, Figure 3 displays the fraction of cells in each node, which are from open and closed MCC periods, and the fraction of all open and closed cells that are in a specific node. For instance, 122 of the 123 cells in node [1,1] are from open MCC periods, accounting for 99.2 % of the cells in this node and 20.0% of all 612 cells during open MCC periods. Additionally, Figures 4 and 5 display the Doppler velocity and spectral width parts of the master SOM and Figure 6 shows the mean reflectivity profiles of each node.

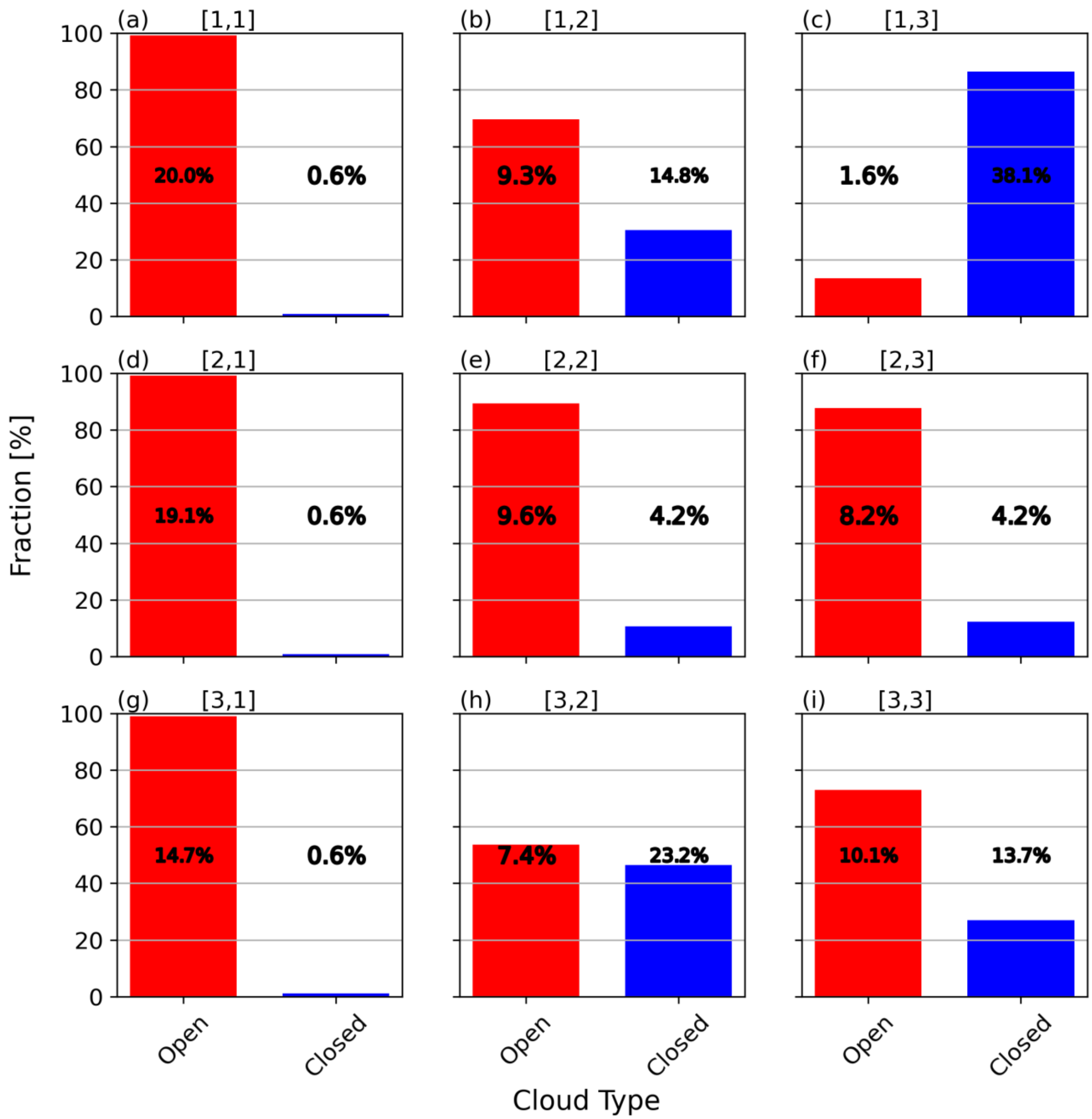

Figure 3. Fraction of cells, which are from open (red) and closed (blue) MCC periods at any node is displayed by the bars. The sum of the two bars in each node equals 100%. The numbers listed in each bar indicate the percentages of the total number of open (left) or closed (right) MCC cells that are in each node: the sum of nine numbers equals 100%.

3.1.1 *Intense convective cells*

The nodes shown in the left column ([1,1], [2,1], and [3,1]) contain almost exclusively cells found during open MCC periods (Figure 3a,d,g). Since these nodes have a large fraction of ascending hydrometeors (Figure 4a,d,g), some rising as fast as 2 m s^{-1} in node [1,1], these precipitation systems are convective. By definition, convection is characterized by ascending hydrometeors during part of the cloud lifecycle or in a cloud region (Houze, 2014). Since these nodes have the largest values of mean reflectivity profiles (Figure 6), they can be labelled *intense* *convective cells*. The intense cells undergo a lifecycle which becomes more apparent when cross-analyzing Figures 2, 4, 5 and 6. Node [1,1] describes rapidly *developing intense cells*, node [2,1] describes *mature intense cells*, and node [3,1] describes *weakening intense cells*. The developing intense cells are characterized by the strongest updrafts of any classification (Figure 4a). The distributions of reflectivity and Doppler velocity values are relatively broad at all levels (Figure 2a, 4a), i.e. clouds are highly heterogeneous with young updrafts and relatively short history of mixing. Furthermore, the most frequent reflectivity values are in the range of -10 – 0 dBZ (Figure 2a), indicating that hydrometeors in these cells generally are still quite small. The spectral width on average is the highest in node [1,1] (Figure 5a). When compared to the mature intense cells, the mean reflectivity profile of the developing intense cells is ~4 – 5 dBZ lower in most of the cloud (Figure 6).

The mature intense cells average the largest reflectivity values of all nodes, pointing to their intensity and maturity (Figure 6). Hydrometeors are mostly of precipitation size, with reflectivity maxima between 10 – 20 dBZ (Figure 2d). Hydrometeor vertical motions tend to be only slightly weaker than those of developing intense cells, either due to an increase in hydrometeor fall speeds, a weakening of the updrafts, or both (Figure 4d). Finally, the variation of hydrometeor vertical motions of the weakening intense cells is substantially lower (Figure 4g), and the mean Doppler velocity aloft is closest to zero (>-1 m s^{-1}), possibly indicating a residual upper-level mesoscale ascent late in the convective lifecycle, something often observed in more organized deep convection (Houze, 2014). Reflectivity values are more narrowly distributed between 10 – 20 dBZ (Figure 2g) indicating weak reflectivity gradients, consistent with more uniform vertical motions (Figure 4g), and much weaker turbulence (spectral width) than either the mature or developing intense cells (Figure 5a,d,g). The mean reflectivity profile for these weakening intense cells is ~2 dBZ lower compared to the mature intense cells (Figure 6). This indicates that the cells in this node generally contain much precipitation, especially at low levels, but are weakening. Since the intense cells and their associated lifecycle are found almost exclusively during open MCC periods (Figure 3), it can be inferred that such intense convection leads to the development of an open MCC cloud field. The dynamics driving the clearing (compensating subsidence over the depth of the convective cells) are not observed, since radar data are only collected in the cloudy sections. This coherent subsidence is at best only indirectly detected by radar: Figure 1a shows many echo-free regions. The relative lack of such regions in closed MCC (Figure 1e) suggests lack of coherent subsidence.

3.1.2 *Stratiform MCC*

Node [1,3] has the lowest mean reflectivity profile of any node with values ~3 dBZ in the lower half of the cloud (Figure 6), showing that precipitation associated with these cells is very light (also shown in Section 3.3), and, given the scarcity of low-level echoes, frequently evaporates before reaching the surface (Figure 1e, 2c). The frequent occurrence of (mostly weak) echoes near normalized cloud top indicates rather uniform cloud tops. Proximity soundings indicate that the cloud layer generally is capped by a stable layer (not shown). These cells have rather stratiform characteristics with weak vertical motions and only a small fraction of upward vertical motions, most of those near cloud top (Figure 4c). Spectral width values are generally low, with higher values near the surface (on account of high freezing levels and rain for some cells in this node) and near cloud top (Figure 5c). Cloud top turbulence may be attributed to generating cells (e.g., Hobbs & Locatelli, 1978; Rosenow et al., 2014) driven by cloud top radiative cooling (Keeler et al., 2016). Such cells, with a reflectivity and Doppler velocity structure similar that of generating cells at the top of frontal clouds (e.g., Plummer et al., 2015; Zaremba et al., 2022), become apparent in a zoom-in of these cells, e.g., around 10:00 UTC in Fig. 1e (not shown). These traits are indicative of stratiform clouds, with hydrometeors initiated at cloud top and slowly growing as they fall through the cloud. Thus, these cells can be classified as non- (or barely-) precipitating stratiform cells. These types of cells are mostly found during closed MCC periods (Figure 3c), which occurred mostly late in the season (Table 1). The SOM classifies very few cells from open MCC periods in this node because the relatively high M values, strong surface heating, and intense convective development characterizing such periods prevent the development of long-lasting stratiform clouds. These *non-precipitating stratiform cells* are likely the "base state" in a closed MCC environment, and they dominate in the closed MCC example in Figure 1h. Yet this cloud regime exists in a M>0 environment and contains occasional convective updrafts (Figure 1f), hence the use of the MCC terminology (consistent with McCoy et al., 2017 and others), even though stratiform processes may dominate clouds and precipitation.

3.1.3 *Other nodes*

Node [1,2] has a relatively broad distribution of hydrometeor vertical motions and spectral width at most levels of the cloud (Figures 4b, 5b). This indicates the presence of convective updrafts. The most frequent reflectivity values are relatively low (0 – 5 dBZ) and distributed throughout the cloud (Figure 2b). Furthermore, the mean reflectivity profile is very similar to the developing intense cells, except that values are ~4 dBZ lower (Figure 6), i.e., the convection is weaker. Combined, these characteristics indicate that cells in node [1,2] are developing. These cells can be labelled *developing moderately intense cells*. Node [2,2] can be labelled *mature moderately intense cells* since the comparison between this node and the developing moderate cells is very similar to the comparison between the developing and mature intense cells in terms of the reflectivity (Figure 2b,e), hydrometeor vertical velocity distributions (Figure 4b,e), and spectral width (Figure 5b,e). Moreover, the mature moderate cells have larger mean reflectivity values than the developing moderate cells (Figure 6).

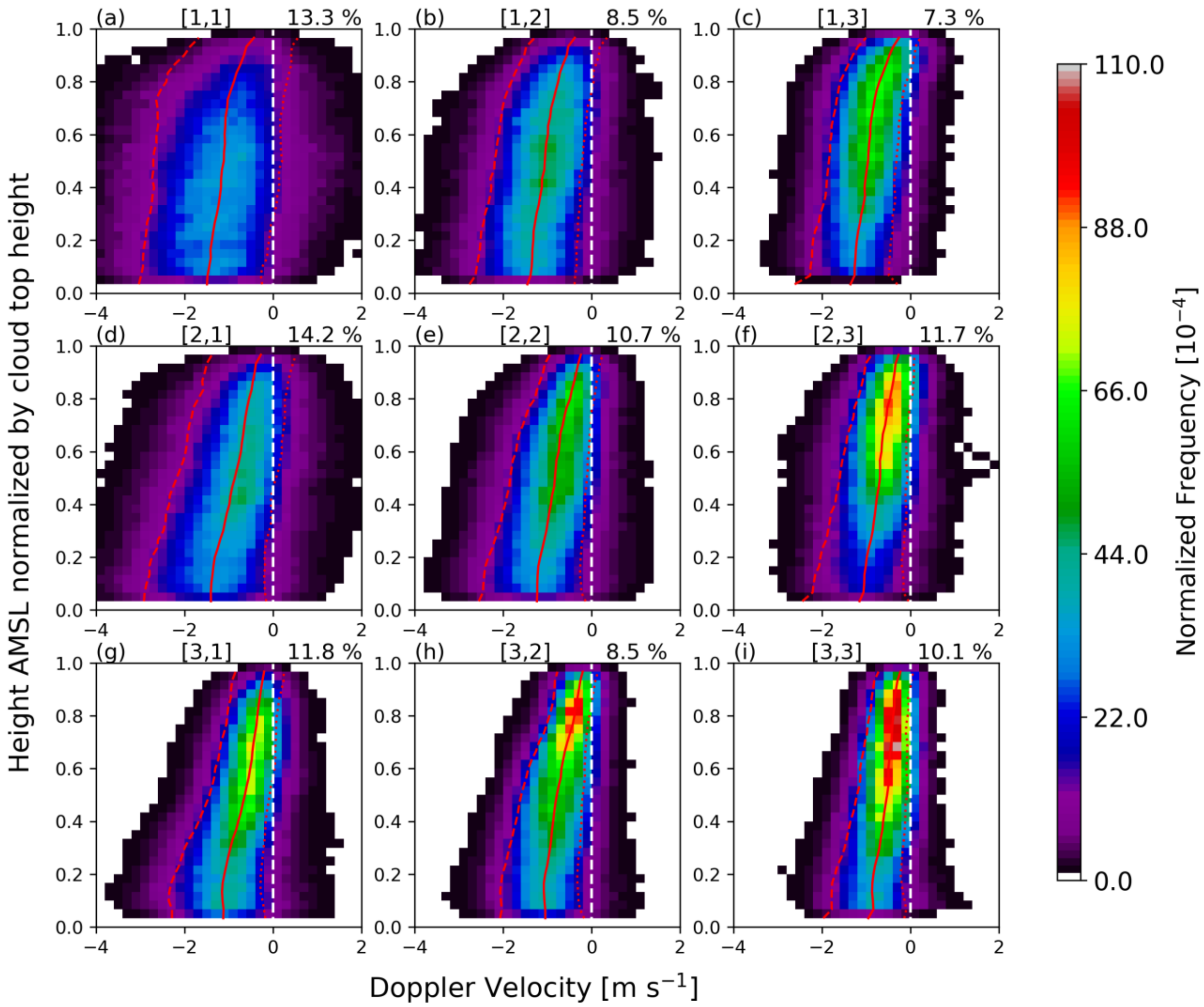

Figure 4. As Figure 2 but for Doppler velocity (hydrometeor vertical motion). The white vertical dashed line indicates 0 m s^{-1} Doppler velocity. The percentages on top indicate the percentage of positive Doppler velocities in each CFAD.

Cells classified as moderate convection (middle column) can be found during open and closed MCC periods (middle column in Figure 3), whereas intense cells (left columns) almost exclusively occur during open MCC periods. This indicates that moderate convection can occasionally occur in the closed MCC environment. Node [3,2] has a rather narrow vertical velocity distribution (Figure 4h), very low spectral width values (Figure 5h), and narrowly distributed, high reflectivity values, although a few dB lower than the weakening intense cells (Figure 2g, h). Therefore, the node is labelled *weakening moderately intense cells.* This node is notable because almost half of its cells occur during closed MCC periods, much more than the two other moderate intensity nodes. While the cells in node [3,2] from open MCC periods may descend from purely convective cells (i.e., weakening moderate cells), the cells from closed MCC periods may be relatively intense stratiform cells. For instance, some cells in Figure 1h fall within this node but are rather stratiform, while two adjacent cells observed around 13 March 2020 1040 UTC are also classified in this node and appear to be the remnants of active convection (Figure 1d). This overlap is remarkable: cells with very different development mechanisms (convective vs.

stratiform) result in very similar vertical cloud structures (reflectivity, hydrometeor vertical velocity, and spectral width) at some point during their development. Thus, we call node [3,2] *mixed weakening moderate + stratiform*.

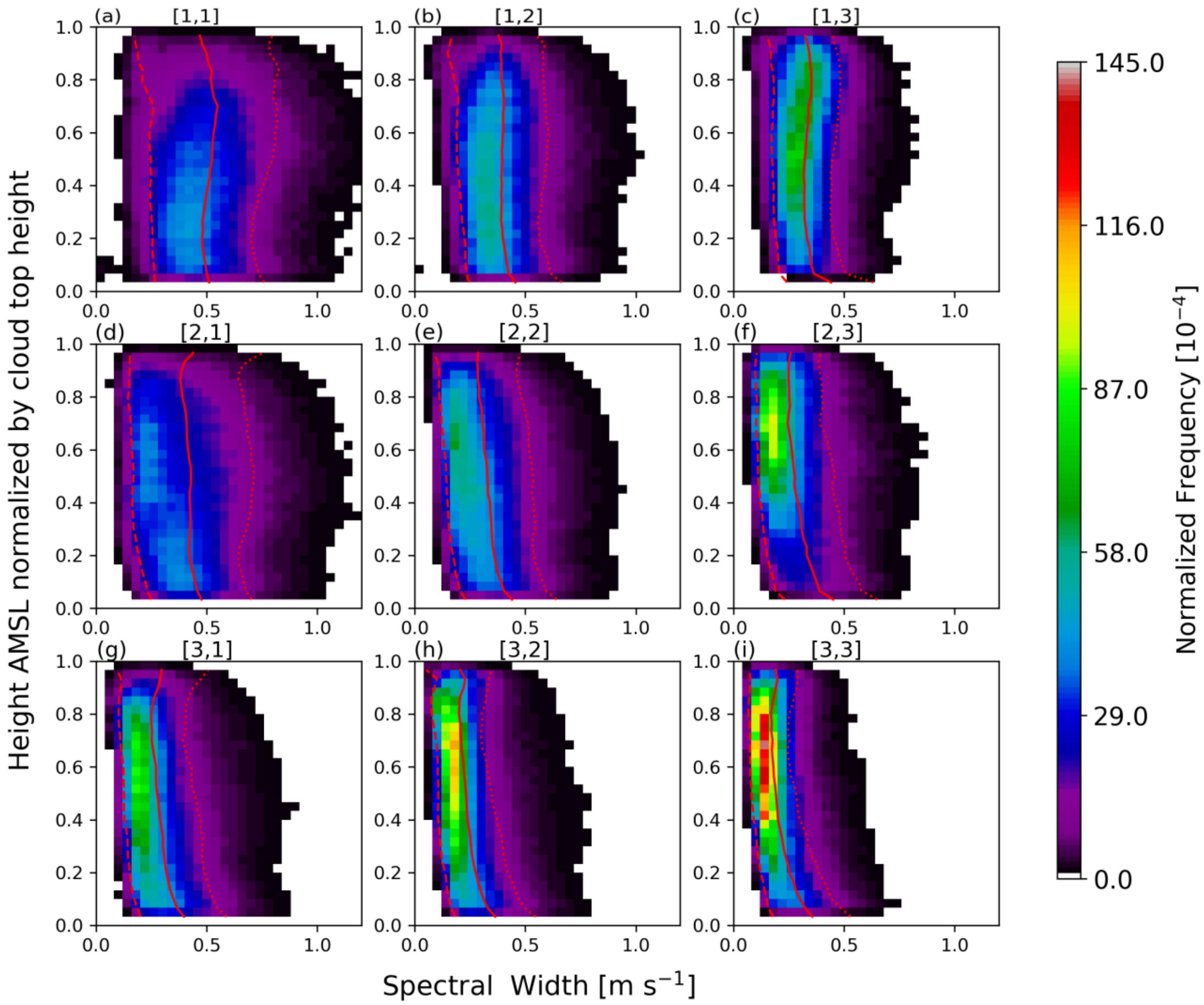

Figure 5. As Figure 2 but for spectral width.

This leaves nodes [2,3] and [3,3] to be analyzed. In terms of the reflectivity distribution (Figure 2f,i) and the mean reflectivity profile (Figure 6) these two nodes are very similar. Only the non-precipitating stratiform cells (node [1,3]) have lower mean reflectivity values. As in node [1,3], reflectivity values in these two nodes are most frequently found in the upper half of the clouds, with few low-level echoes, indicating that they are often non-precipitating. The low-level echoes are mostly below 0 dBZ for both nodes, indicating light surface precipitation when it occurs. Of all nine nodes, node [3,3] has the narrowest distribution of vertical velocity, generally slow hydrometeor descent (Figure 4i), and the narrowest distribution of spectral width with generally very low values (Figure 5i). An important difference between the non-precipitating stratiform cells (node [1,3]) and the cells in nodes [2,3] and [3,3] can be seen in the spectral width (Figure 5). While for nodes [2,3] and [3,3] the mean spectral width steadily decreases with height, for the stratiform cells node [1,3] it is higher aloft with a maximum near cloud top. This is consistent with the observation, in the time-height transects of individual cells, that fine-scale cloud top generating cells with high spectral width, high variability in vertical velocity, and reflectivity virga (Figure 1h) are common in node [1,3], and less common in the two other nodes on the right, nodes [2,3] and [3,3]. This indicates that the latter two nodes are associated relatively more with convection than long-lived stratiform cloud, since the radiative cooling needed for generating cells may be a slow process (Keeler et al., 2016). These two nodes occur mostly during the open-cellular periods (Figure 3). Node [2,3] has slightly lower mean reflectivity compared to node [3,3] (Figure 6), and very few low-level echoes, indicating that it is mostly non-precipitating.

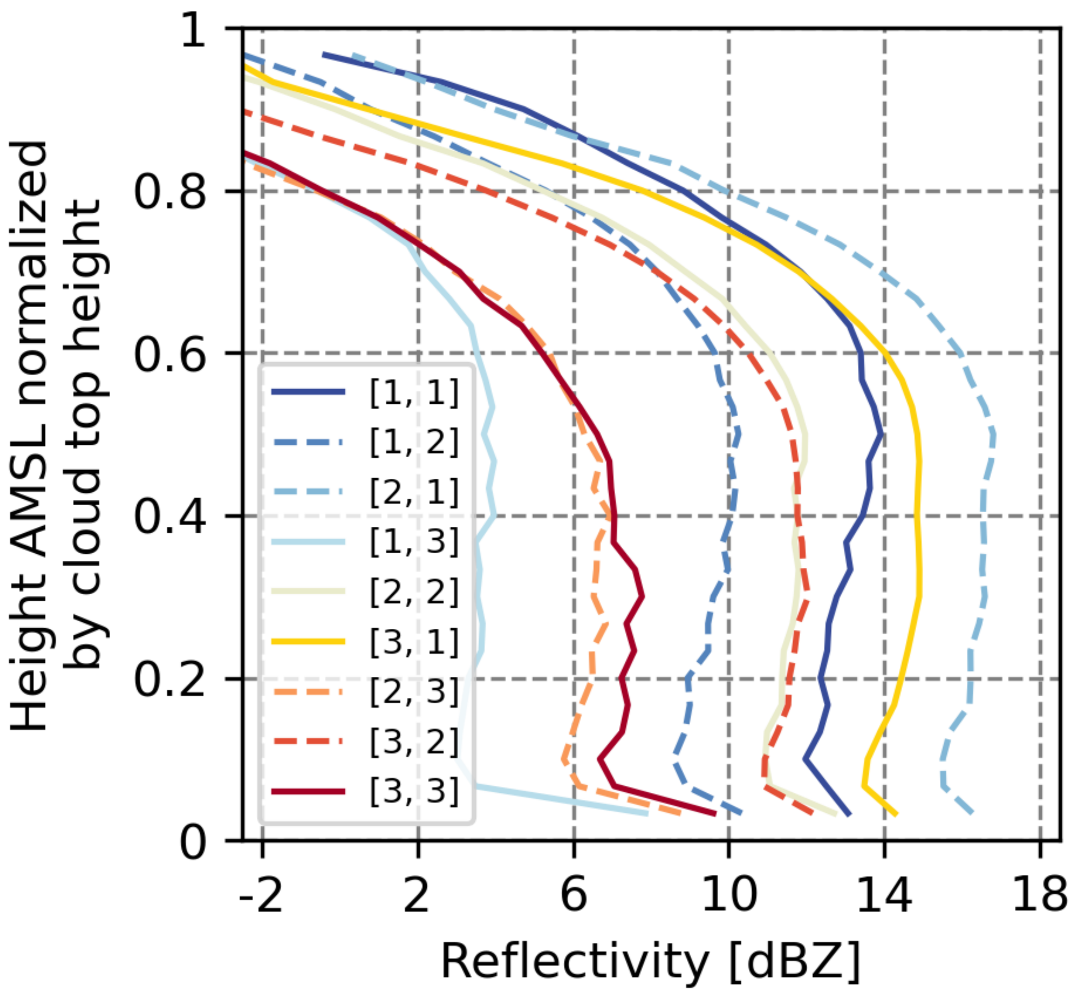

Figure 6. Mean profiles of reflectivity at each node of the master SOM. Dashed lines are used to help differentiate nodes.

Based on our analysis, node [3,3] can be labelled *dissipating anvils*. The cells included in this node are at the very end of the convective lifecycle of either intense or moderate cells. Node [2,3] has more diverse vertical motions with a rather high fraction of hydrometeors carried upward and some localized high spectral width values. This may be explained by the development of new updrafts at the outflow boundary of dissipating cells. Thus, we call this node *mixed anvils + updrafts*.

In summary, seven nodes are purely convective or of convective origin, i.e., the three nodes on the left (intense convection), 2 in the middle (moderate convection), and two on the right: node

[3,3] as dissipating anvil, and node [2,3] as mixed anvils + new convective updrafts. One node (the upper right, [1, 3]) is purely stratiform, occurring during weak MCAOs. And one node, [3,2], is mixed convective-stratiform, combining the weakening moderate convective cells with more intense stratiform cells.

3.2 Cloud Morphological Characteristics

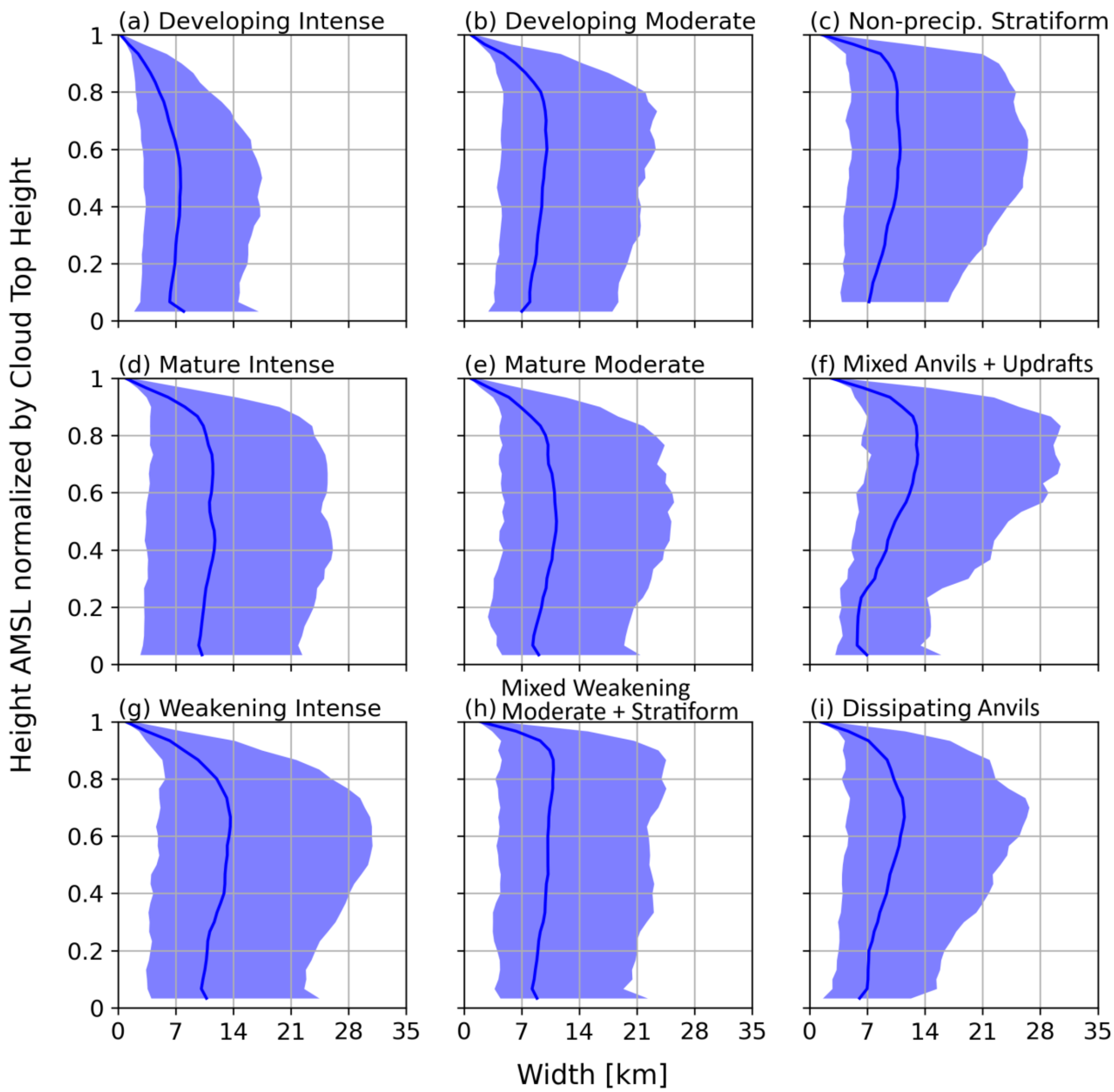

Figure 7. Width of the identified cells at each node of the master SOM. The solid line indicates the mean and the shaded area the interquartile range. The titles indicate the node nomenclature.

To examine the characteristics of the cell classifications, other observations collected during the identified cells are collectively analyzed for each classification. In this section, cell morphological characteristics such as the cloud width and CTH are analyzed. The horizontal width of the cells is determined by multiplying the temporal width of the cells at normalized height levels with the closest available *INTERPOLATEDSONDE* wind speed (Figure 7), i.e., it is the along-wind width only. Note that this width is not fully independent from the classification as it is determined using the same radar data. Due to the standardized CFADs, the classification contains only height-relative information about the cloud width, e.g., upper vs lower levels. The absolute cloud width matters less, it depend in part on a choice in the watershed segmentation method. What matters more is the vertical structure of the cell width.

There are substantial differences between the cell classifications in terms of the vertical structure of their width. The intense cells (left column) show a widening as they develop. On average, developing intense cells are narrower than any other classification and are narrow-peaked, since the cells are in an upward development. At the mature stage, cells are wider than at the developing stage and the width of cells is relatively similar throughout the cloud except very close to cloud top (Figure 7d). The weakening intense cells are wider in the upper half, the result of divergent flow aloft (Figure 7g). Anvil formation is typical for capped deep convection, but anvils are not well-developed in the weakening intense cells.

This same pattern is not as clear for the moderate cells (middle column, Figure 7). This is likely due to some of these cells occurring during closed MCC periods. The moderate convection during closed MCC periods are generally surrounded by preexisting clouds, namely those that are classified as stratiform cells (Figure 7c,h). Therefore, the average width of moderate cells shows similarities to the stratiform cells. The non-precipitating stratiform cells have a constant width in the upper half of the cloud, but they become narrower in the lower parts of the cloud, indicating their frequent lack of surface precipitation. If Figures 7b,e,h are created for the open MCC periods only, these moderately-intense cells show a widening as they mature, comparable to the intense cells with a similar indication of cloud top divergence for the weakening cells (node [3,2]) (not shown). The dissipating anvils (Figure 7f,i), have an especially large contrast between width near cloud base (narrow) and cloud top (wide). This is likely related to the cells dissipating from the base upwards, leaving behind anvils.

Next, CTH, cloud top temperature (CTT), and cloud base height (CBH) are analyzed (Figure 8). A distribution of CTHs, CBHs and CTTs is available for each cell. Here, we use the 10% highest CTHs (10% lowest CBHs and CTTs) of each cell to avoid cells being represented by outliers of these variables. The CTH of a cell is determined based on the cell mask from the watershed segmentation (see Figure 1d,h). The CTT is the *INTERPOLATEDSONDE* temperature at this height at the matching time. CBH estimates are from the *ARSCLKAZRBND1KOLLIAS* product, which is derived from ceilometer and lidar data. Since cells can overlap in time, cloud base heights are assigned to the cell that is the closest to the surface in any given KAZR profile.

The distributions of CTH and CTT show a large spread within individual nodes (Figure 8), which can be explained by the variability of the mean CTHs of the MCC periods (see Table 2). However, the distributions also reveal differences between the nodes. For the intense cells, the mean CTH is the shallowest (2.8 km) for the developing node (Figure 8a). This can be explained with the cells developing upwards and not having reached the maximum CTH of their lifecycle. This maximum CTH (3.6 km) is reached at the mature stage (Figure 8d). At the weakening stage, CTHs are slightly lower on average (3.3 km) (Figure 8g). As a result, very few developing intense cells (5 %) have CTTs at which homogeneous freezing can occur [-38 °C (Ickes et al., 2015)], while many more mature intense cells (25 %) reach such cold temperatures (Figure 8a,d).

This indicates that primary nucleation may contribute to an abundance of ice crystals in the cells, which may play a role in the lifecycle of these intense cells. Cloud phase will be examined in more detail later (Section 3.3). CBHs are not very different between the lifecycle stages, nor between MCC periods; they are almost always between 300-1000 m above sea level, consistent with the surface-based lifting condensation level.

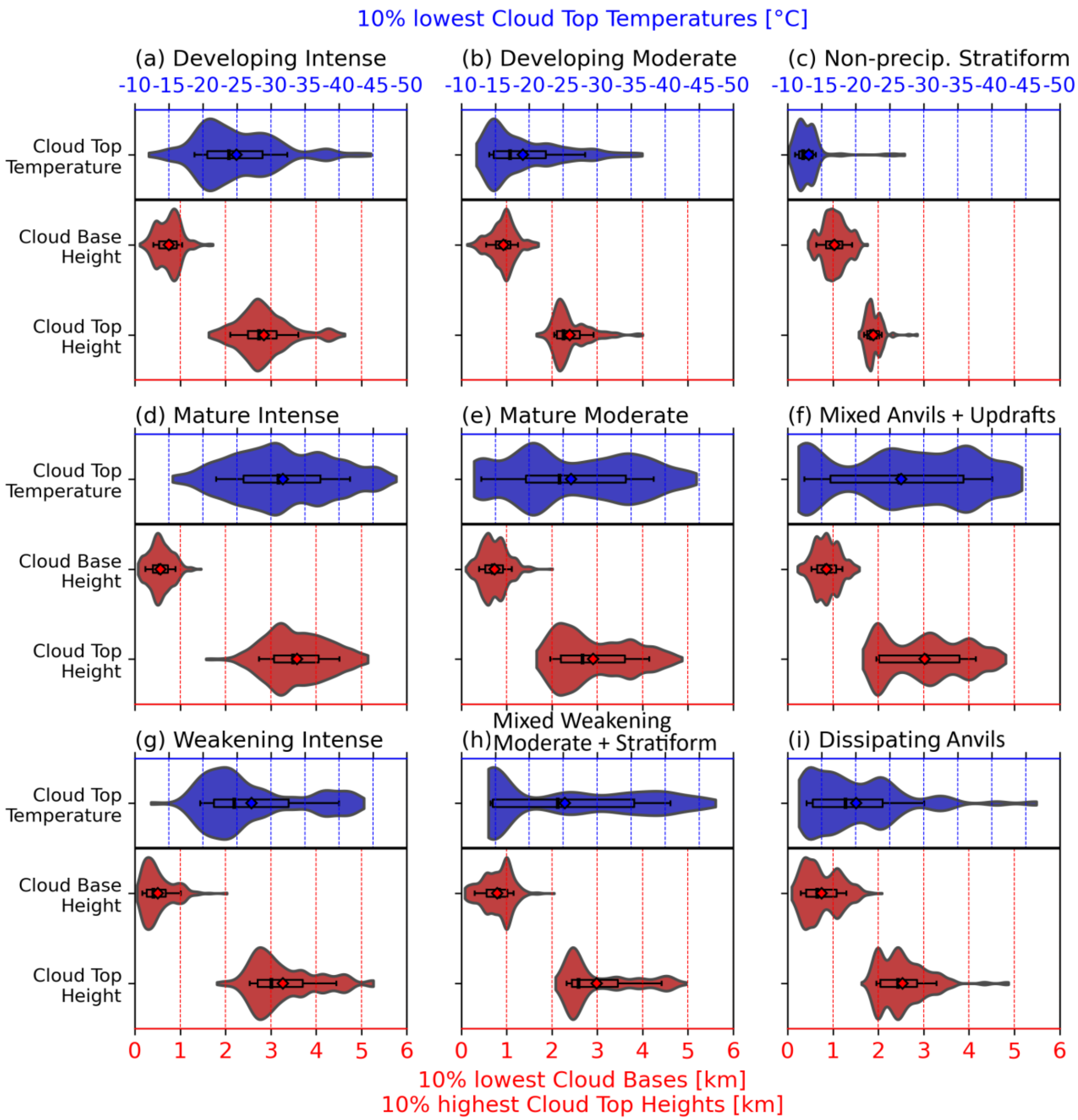

Figure 8. Distributions of cloud top temperature, cloud top height, and cloud base height of the identified cells at each node of the master SOM. The box plots inside the violin plots indicate mean (diamond), median (solid black line), interquartile range (box), and 10th and 90th percentile (whiskers).

The distributions of CTH, CTT, and CBH for the moderate cells do not directly resemble the same pattern as the intense cells. The closed MCC periods within these moderate cells generally have rather shallow cloud tops (CTH between 2 – 2.5 km) and high cloud bases (CBH ~1 km) (not shown). The mean CTH increases steadily with the development stages from 2.4 km (developing), to 2.9 km (mature), to 3.0 km (mixed weakening moderate + stratiform) (Figure 8b,e,h). This deepening is more pronounced when analyzing only moderate cells from open MCC periods (2.5 km, 3.2 km, and 3.5 km; not shown). The reasons for the continued deepening in the late-convective stage are not clear. It might be related to the node [3,2] also containing some cells of a more intense convective origin (left column in Figures 2, 4, and 5) that have further weakened compared to the weakening intense classification. Such cells tend to have a high cloud top, explaining the tail of CTH values around 4-5 km in this node (Figure 8h).

The non-precipitating stratiform cells have comparatively warm CTT (> -15 °C), and a narrow distribution of CTH with ~80 % of cloud tops between 1.7 – 2.1 km, indicating their homogeneous cloud top, and a CBH around 1 km. This highlights the shallow depth of the closed cells, and explains the lack of surface precipitation. The two other nodes in the right column (dissipating anvils) also have a large fraction of CBHs above 1.0 km. The mixed node (dissipating anvils + updrafts) has a broad range of CTHs, with new updrafts on the shallow side and dissipating anvils on the deep side.

3.3 Cloud Microphysical and Dynamical Characteristics

In this section, microphysical and dynamical characteristics of the cells are analyzed, specifically five variables: microwave radiometer liquid water path (LWP) (Figure 9), radar-retrieved ice water path (IWP) from the *MICROBASEKAPLUS* product (Figure 9), cloud phase classification from the *THERMOCLDPHASE* product (Figure 10), surface precipitation rate, and surface temperature anomaly (cold pools) (Figure 11). Since multiple cells can be identified in a single vertical column, it is not always possible to unambiguously assign a LWP measurement (a vertically integrated quantity) to a single cell. Therefore, LWP are only assigned to a cell if that cell makes up more than 80 % of cloud in a vertical profile and is the closest cell to the surface. The *MICROBASEKAPLUS* product has the same time-height resolution as the ARSCL product (see Table 1) used to identify individual cells. Ice water content retrievals from this product can be directly used to calculate the IWP associated with a specific cell, using the 2D watershed segmentation. Data points from the 2D *THERMOCLDPHASE* product are assigned to the cell they cover. Precipitation measurements and the surface temperature anomaly are assigned to the cell closest to the surface. The minimum (including negative values) difference between the instantaneous and the six-hour mean surface virtual potential temperature (centered on the instantaneous value) during each cell is used as a measure for cold pool strength.

The cell classifications show strong contrasts in terms of presence of liquid water (Figure 9). For the intense cells, LWP is much larger in the developing stage with a mean value (90th percentile) of 0.33 kg m^{-2} (0.86 kg m^{-2}) (Figure 9a). As the convection matures, LWP values decrease substantially: the mean value (90th percentile) for the mature intense cells is 0.21 kg m^{-2} (0.62 kg m^{-2}) and for the weakening intense cells 0.08 kg m^{-2} (0.20 kg m^{-2}) (Figure 9d,g). Furthermore, the LWP distributions of the intense cells have their mode at relatively low values (< 0.1 kg m^{-2}), indicating that the presence of larger amounts of liquid is confined to small pockets, often coincident with deep updrafts (not shown).

IWP increases with the development of the intense cells reaching the maximum at the mature stage and is only slightly lower in the weakening stage, but still substantially larger than at the developing stage (Figure 9a,d,g). This shift in phase with the lifecycle of the intense cells is also present in the *THERMOCLDPHASE* product. Slightly over 50 % of the phase identified in developing intense cells is mixed-phase or even pure liquid (Figure 10a). In contrast, mixed-phase makes up just ~15 % of the phase identified in the mature and weakening intense cells, the rest being classified as snow (or ice) (Figure 10d,g). These findings indicate that the conversion of liquid to ice is a key component in the evolution of the intense convection.

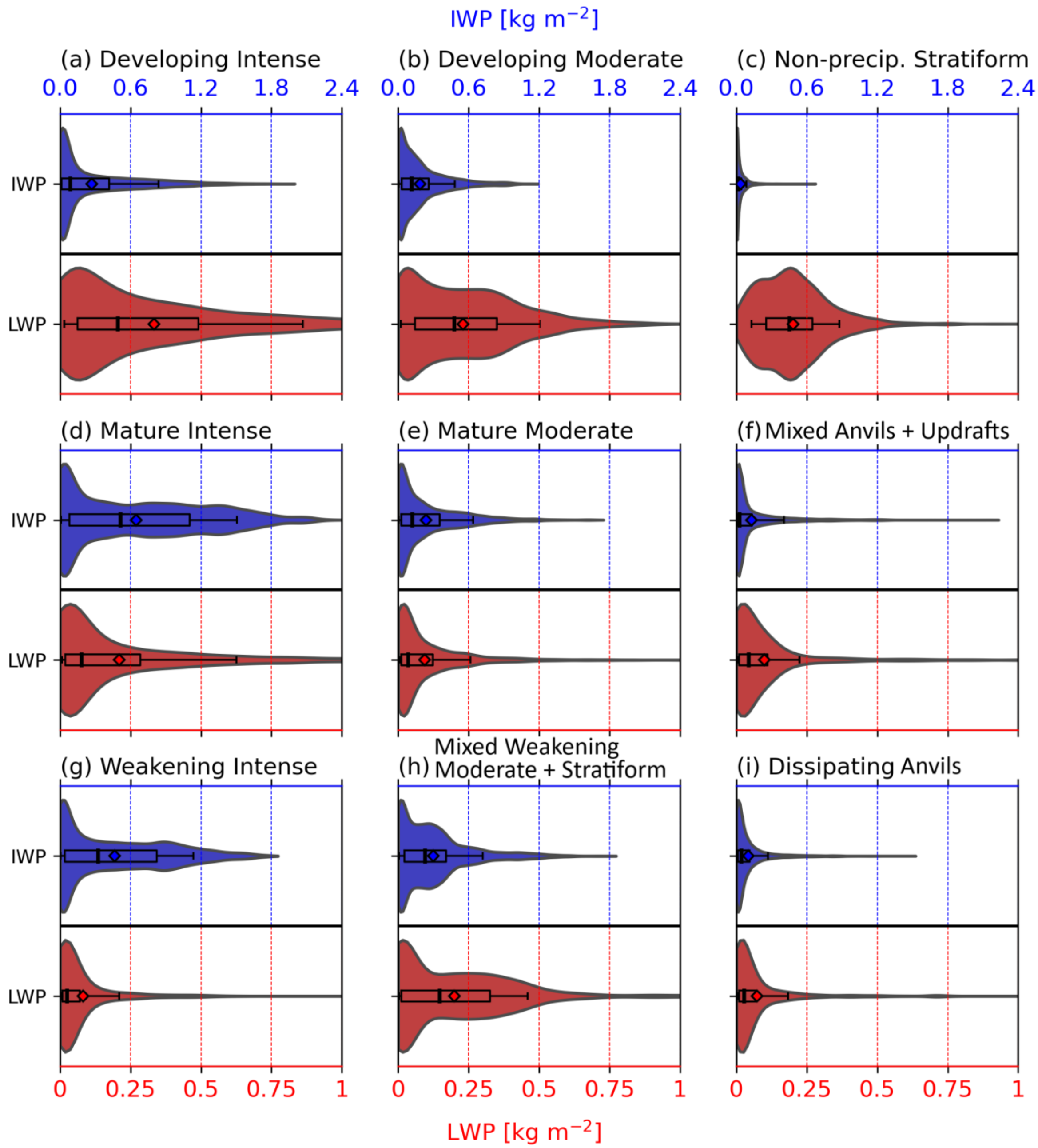

Figure 9. Distributions of liquid water path (LWP) and radar retrieved ice water path (IWP) of the identified cells at each node of the master SOM. Box plots are as in Figure 8.

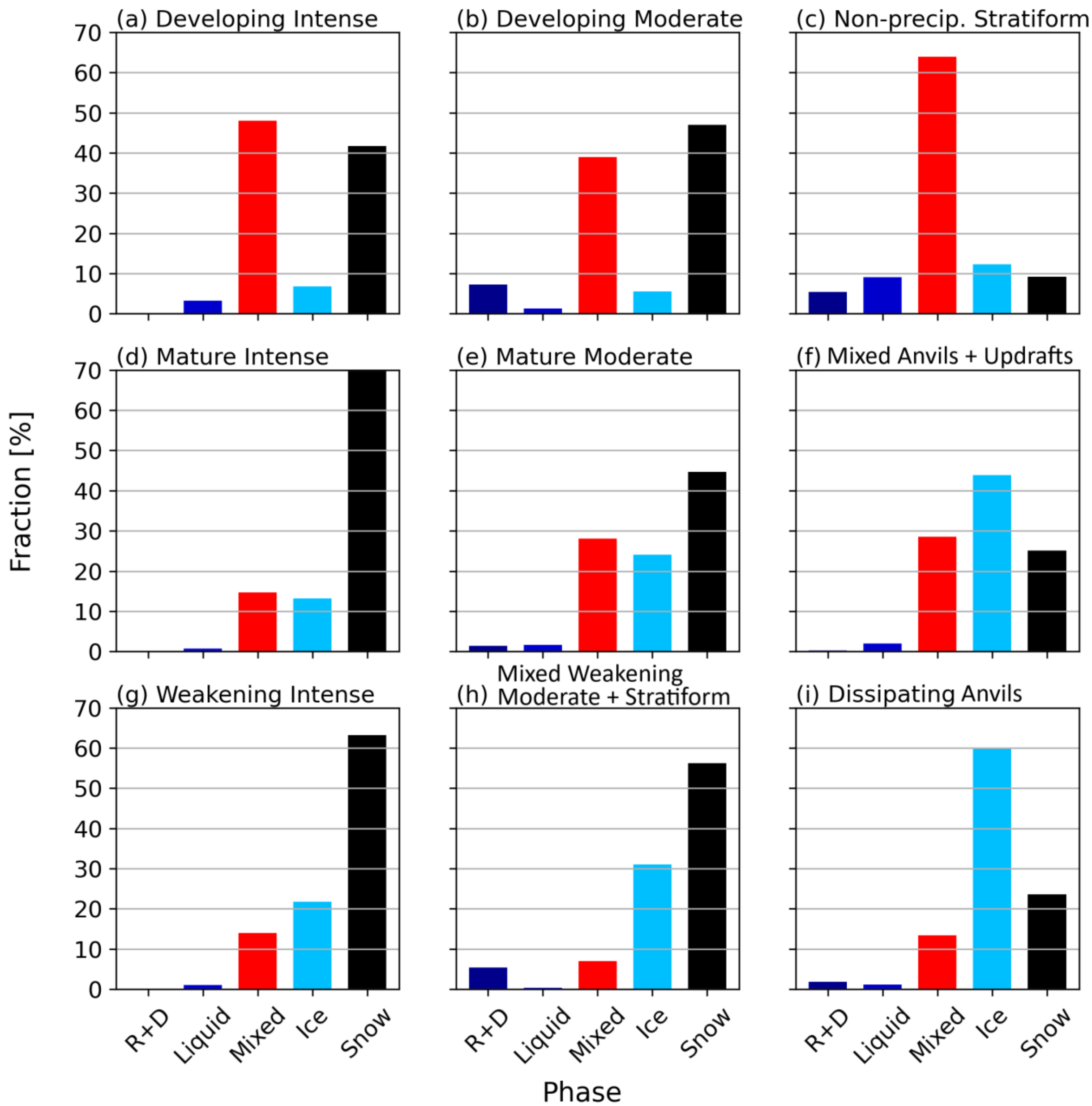

Figure 10. Fraction of retrieved cloud phase of the identified cells at each node of the master SOM. Note that R+D refers to rain and drizzle combining the rain, drizzle, and liquid + drizzle classification of the *THERMOCLDPHASE* product (see Table 1).

The moderate cells follow the same pattern, with comparatively more liquid and mixed-phase during the developing stage and more ice during the mature stage (Figures 9b,e; 10b,e). If only the contributions from open MCC periods to the mixed node [3,2] (weakening moderate + stratiform) are considered, this pattern continues and there is even less liquid (i.e., LWP mode close to 0 kg m^{-2} in Figure 9h) (not shown). However, this mixed node includes cells with rather large LWP values with a mean of 0.2 kg m^{-2}. This can be mostly attributed to a high freezing level and rain, associated with the closed stratiform cells. Above the freezing level these stratiform cells are mostly ice phase, with very little mixed-phase (Figure 10h). In contrast, the non-precipitating stratiform cells consist mostly (~75 %) of either mixed-phase or pure liquid cloud (Figure 10c).

Liquid is rather normally distributed in this node. This can be inferred from the median, mean, and mode of LWP being almost the same (~0.19 kg m^{-2}) for this node. The dissipating anvils consist mostly of ice particles smaller than snow and have very low LWP values (Figures 9f,i; 10f,i). The mixed node with dissipating anvils + updrafts has a larger spike of high LWP values and more liquid and mixed-phase cloud pixels than the dissipating anvil node (Figures 9f,i; 10f,i). This is likely due to the new updrafts and associated supercooled liquid water.

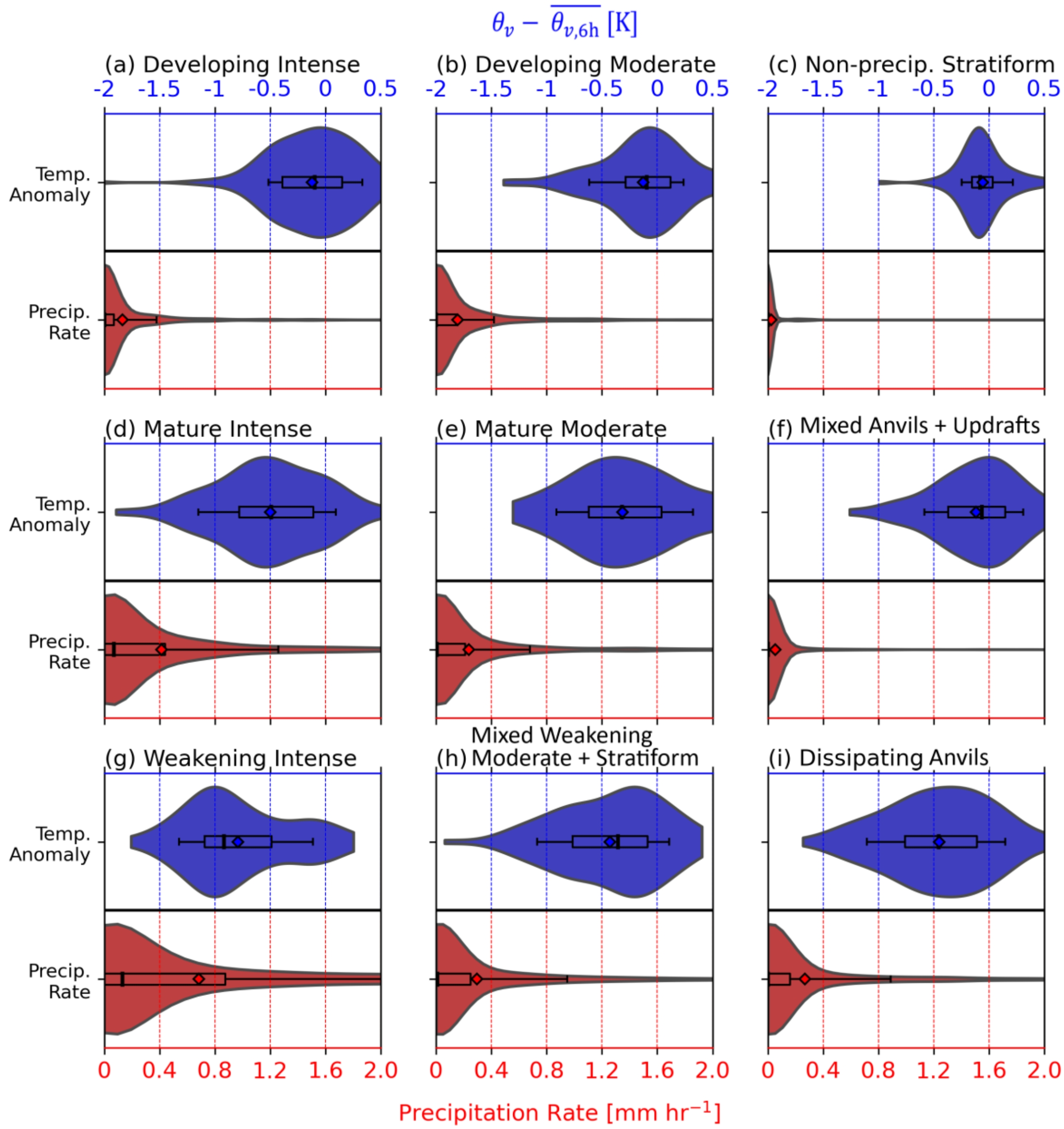

Figure 11. Distribution of precipitation rate and virtual potential temperature anomaly of the identified cells at each node of the master SOM. Box plots are as in Figure 8.

Lastly, surface precipitation rates and cold pools are analyzed. Precipitation rates and (cold) temperature anomalies increase steadily in the lifecycle of the intense cells (Figure 11a,d,g). Maxima are reached at the weakening stage, confirming the convective nature of the cold pools, i.e. driven by precipitation. For the moderate cells, the same pattern is observed with overall weaker precipitation rates (Figure 11b,e,h). The mixed weakening moderate + stratiform category shows little difference in precipitation rate between open and closed MCC periods (not shown). The non-precipitating stratiform node has the lowest precipitation rates confirming this nomenclature, which was based on the low reflectivity values (Figures 2c, 11c). Furthermore, these cells display no significant temperature anomalies at the surface, indicating that the precipitation cores are driven not by cold pool dynamics, but by dynamical or microphysical processes in the cloud layer. For the dissipating anvil nodes, there is a contrast in precipitation rate and temperature anomaly (Figure 11f,i). While both nodes have low precipitation rates, the mixed dissipating anvils + updrafts node has even less precipitation and are not associated with cold pools. This is consistent with the development of new updrafts outside (or at the edges of) cold pools.

4 Discussion

Our analysis shows that a SOM-based classification can determine distinct characteristics of open and closed MCC clouds observed during MCAOs. This classification reveals a lifecycle of convection in which cells rapidly develop (with strong updrafts, high spectral width, growing cloud tops, high LWP but low IWP, no cold pool) and then weaken through mixed-phase precipitation processes, which in turn produce a cold pool. This study does not explicitly observe the coupled microphysical-dynamical process that results in the open-cellular cloud morphology, but the existence of objectively defined nodes corresponding to lifecycle stages strongly suggest it. This cycling contributes to the short residence times of water vapor (~1 day) described for MCAOs (Papritz & Sodemann, 2018).

A few things need to be considered when interpreting the results of the cloud classification. While discrete classifications are determined, it should be remembered that these clouds are found on a continuous spectrum in terms of their development stage, the intensity of convection, or stratiform precipitation growth. For instance, the transition of cells from developing to mature is fluid. Also, an individual cell defined objectively (watershed segmentation method) may consist of multiple smaller convective turrets at different life stages, e.g., the blue cell around 12:20 UTC in Figure 1a-d. Furthermore, the analysis is limited to a two-dimensional snapshot of each cell in time and height. The large number of different cells that were observed enables the classification, but it is not possible to determine how the characteristics relate to the three-dimensional structure of individual cells. For instance, it is not possible to determine through which parts of a cell the vertical scans of the KAZR were made (center vs. edge of cell). Another limitation of this study is that it does not examine cell interaction, e.g., does a cold pool associated with a dissipating cell trigger a new cell?

Nevertheless, the approach to objectively classify clouds helps to conceptualize the defining characteristics of convective cells during MCAOs, and to elucidate the driving dynamical-microphysical processes. While buoyant ascent (possibly forced by outflow boundary convergence) clearly is the driving mechanism during relatively intense MCAOs with strong surface heating, stratiform cloud and precipitation processes play a role as well, especially under low M value conditions. This may result in a closed MCC organization, which was less common during COMBLE, and observed only late in the season (April-May). Remarkably, the cloud vertical structure of late-stage moderately-intense convective cells, as probed by the KAZR, is very similar to that of more intense stratiform precipitation cores, resulting in a single SOM node that combines both.

Our analysis is limited to the comparatively short data record of COMBLE. While a certain range of MCAO intensities is captured, it is not necessarily representative of the long-term climatological conditions over the high-latitude oceans, and certainly not elsewhere. The approach presented here, using the COMBLE dataset, could be used to analyze MCC clouds at observational sites with a much longer data record such as ARM's North Slope of Alaska and Eastern North Atlantic (ENA) sites. On average, M values are lower in the area of the ENA site during MCAOs (or post-frontal periods) (Fletcher et al., 2016), and the amount of time during which open and closed MCC occurs is much more equal (Jensen et al., 2021). A comparison with the ENA site allows examination of the importance of the ice phase in MCAO convection dynamics: the ENA site is considerably warmer, resulting in less common mixed-phase MCC.

In future work, we will analyze the MCAOs listed in Table 2 using a large-eddy simulation (LES), starting with the 13 March 2020 case (Kosović et al., 2023), and a similar approach as presented in this study. A radar simulator such as the Cloud Resolving Model Radar Simulator (Oue et al., 2020) will be used to create synthetic radar data from the LES that is comparable to KAZR observations. A comparison of the observational SOM classification to an LES-output based SOM classification may reveal whether the model reproduces MCC with the characteristics observed in this study. Another benefit of conducting this analysis with LES is that the three-dimensional spatial structure and the temporal evolution of the variables presented in this study could be analyzed. The LES further allows analysis of a large number of cells given that the simulation covers a large domain, stretching from north of the ice edge to south of the COMBLE observation site (~1,200 km distance).

5 Conclusions

In this study, the characteristics of open and closed mesoscale cellular convection (MCC) during marine cold-air outbreaks (MCAO) are investigated. The data were collected as part of the COMBLE field campaign at a coastal site in northern Norway. Unambiguous periods of open and closed MCC are determined within MCAOs, based on radar and satellite data. Periods of open and closed MCC are associated with MCAO intensity and display distinct cloud top height differences. Watershed segmentation and self-organizing map (SOM) algorithms are used to objectively identify and classify individual cells observed during these periods. The identification and classification of cells are based on the vertical structure of the cells as observed by a vertically scanning Ka-band radar. The main findings from the classification of individual cells are as follows:

- The SOM algorithm yields nine nodes with distinct vertical distributions of the three radar moments, i.e., reflectivity, Doppler velocity, and spectral width. These nine nodes are given descriptive identifiers that refer to the intensity of their convection, their lifecycle stage, and whether they have stratiform characteristics. The nine classifications are: developing, mature and weakening intense cells, developing and mature moderate cells, mixed weakening moderate cells and precipitating stratiform cells, non-precipitating stratiform cells, anvils with updrafts, and dissipating anvils.

- Cells classified as intense convection have the largest reflectivity values of any classification, and only occur during open MCC periods. They are separated into developing, mature, and weakening intense cells. Developing cells are associated with strong updrafts, high spectral width, shallower cloud depth, pockets of large liquid water path, and minimal precipitation. As the convection matures and then decays, updrafts decrease in strength, cloud tops are higher and colder, the amount of liquid decreases while the amount of ice increases, surface precipitation intensifies and cold pools develop below the cells. At the very end of the lifecycle dissipating anvils remain. New updrafts can develop below these anvils if they are not located over cold pools.

- Cells with more moderate convection compared to the intense cells are observed during open MCC and occasionally during closed MCC. This moderate convection shows a similar lifecycle as the intense cells.

- During closed MCC periods, cells with stratiform characteristics are observed most frequently. These stratiform cells are separated into non-precipitating and precipitating cells. Non-precipitating stratiform cells are relatively shallow with cloud tops mostly >-15°C, have weak vertical drafts, somewhat enhanced spectral width near cloud top, and mostly mixed-phase cloud. The precipitating stratiform cells are dominated by ice, and are structurally very similar to late-phase (weakening) moderately intense convective cells. Neither reveal distinct cold pools.

Acknowledgments: The COMBLE campaign was enabled by an ARM Mobile Facility deployment proposal to the Office of Science of the U.S. Dept. of Energy (DOE). This research was supported by DOE Atmospheric System Research (ASR) Grants DE-SC0018927 and DE-SC0021151.We appreciate the assistance from several ARM instrument and data mentors, especially Scott Giangrande. The National Center for Atmospheric Research is a major facility sponsored by the National Science Foundation under Cooperative Agreement 1852977.

Open Research: COMBLE campaign data at the Andenes site (ANX) used in this study are available through the references found in Table 1. The watershed segmentation was implemented in python code using the sci-kit image package (https://scikit-image.org/) and the self-organizing map algorithm was implemented with the minisom package (https://github.com/JustGlowing/minisom).

References

Agee, E. M. (1987). Mesoscale cellular convection over the oceans. *Dynamics of Atmospheres and Oceans*, *10*(4), 317–341. https://doi.org/10.1016/0377-0265(87)90023-6

Atkinson, B. W., & Zhang, J. W. (1996). Mesoscale shallow convection in the atmosphere. *Reviews of Geophysics*, *34*(4), 403–431. https://doi.org/10.1029/96RG02623

Brümmer, B. (1999). Roll and Cell Convection in Wintertime Arctic Cold-Air Outbreaks. *Journal of the Atmospheric Sciences*, *56*(15), 2613–2636. https://doi.org/10.1175/1520-0469(1999)056<2613:RACCIW>2.0.CO;2

Clothiaux, E. E., Miller, M. A., Perez, R. C., Turner, D. D., Moran, K. P., Martner, B. E., Ackerman, T. P., Mace, G. G., Marchand, R. T., Widener, K. B., Rodriguez, D. J., Uttal, T., Mather, J. H., Flynn, C. J., Gaustad, K. L., & Ermold, B. (2001). *The ARM Millimeter Wave Cloud Radars (MMCRs) and the Active Remote Sensing of Clouds (ARSCL) Value Added Product (VAP)* (DOE/SC-ARM/VAP-002.1). ARM user facility, Pacific Northwest National Laboratory, Richland, WA. https://doi.org/10.2172/1808567

Davy, R., & Outten, S. (2020). The arctic surface climate in CMIP6: Status and developments since CMIP5. *Journal of Climate*, *33*(18), 8047–8068. https://doi.org/10.1175/JCLI-D-19-0990.1

Doviak, R., & Zrnic, D. S. (1993). *Doppler Radar and Weather Observations (2nd ed.).* San Diego: Academic Press.

Dunn, M., Johnson, K., & Jensen, M. (2011). *The Microbase Value-Added Product: A Baseline Retrieval of Cloud Microphysical Properties* (DOE/SC-ARM/TR-095). PNNL; Richland, WA. https://doi.org/10.2172/1015189

Eastman, R., McCoy, I. L., & Wood, R. (2021). Environmental and Internal Controls on Lagrangian Transitions from Closed Cell Mesoscale Cellular Convection over Subtropical Oceans. *Journal of the Atmospheric Sciences*, *78*(8), 2367–2383. https://doi.org/10.1175/JAS-D-20-0277.1

Eastman, R., McCoy, I. L., & Wood, R. (2022). Wind, Rain, and the Closed to Open Cell Transition in Subtropical Marine Stratocumulus. *Journal of Geophysical Research: Atmospheres*, *127*(20), e2022JD036795. https://doi.org/10.1029/2022JD036795

Fairless, T., Jensen, M., Zhou, A., & Giangrande, S. (2021). *Interpolated Sounding and Gridded Sounding Value-Added Products*. DOE Office of Science Atmospheric Radiation Measurement (ARM) Program (United States). https://doi.org/10.2172/1248938

Fletcher, J. K., Mason, S., & Jako, C. (2016). A climatology of clouds in marine cold air outbreaks in both hemispheres. *Journal of Climate*, *29*(18), 6677–6692. https://doi.org/10.1175/JCLI-D-15-0783.1

Geerts, B., Giangrande, S. E., McFarquhar, G. M., Xue, L., Abel, S. J., Comstock, J. M., Crewell, S., DeMott, P. J., Ebell, K., Field, P., Hill, T. C. J., Hunzinger, A., Jensen, M. P., Johnson, K. L., Juliano, T. W., Kollias, P., Kosovic, B., Lackner, C., Luke, E., … Wu, P. (2022). The COMBLE Campaign: A Study of Marine Boundary Layer Clouds in Arctic Cold-Air Outbreaks. *Bulletin of the American Meteorological Society*, *103*(5), E1371–E1389. https://doi.org/10.1175/BAMS-D-21-0044.1

Hobbs, P. V., & Locatelli, J. D. (1978). Rainbands, Precipitation Cores and Generating Cells in a Cyclonic Storm. *Journal of the Atmospheric Sciences*, *35*(2), 230–241. https://doi.org/10.1175/1520-0469(1978)035<0230:RPCAGC>2.0.CO;2

Houze, R. A. (2014). *Cloud Dynamics*. Academic Press.

Ickes, L., Welti, A., Hoose, C., & Lohmann, U. (2015). Classical nucleation theory of homogeneous freezing of water: Thermodynamic and kinetic parameters. *Physical Chemistry Chemical Physics*, *17*(8), 5514–5537. https://doi.org/10.1039/C4CP04184D

Intergovernmental Panel On Climate Change. (2021). *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (1st ed.). Cambridge University Press. https://doi.org/10.1017/9781009157896

Jensen, M. P., Ghate, V. P., Wang, D., Apoznanski, D. K., Bartholomew, M. J., Giangrande, S. E., Johnson, K. L., & Thieman, M. M. (2021). Contrasting characteristics of open-and closed-cellular stratocumulus cloud in the eastern North Atlantic. *Atmospheric Chemistry and Physics*, *21*(19), 14557–14571. https://doi.org/10.5194/acp-21-14557-2021

Jensen, M. P., Giangrande, S., Fairless, T., & Zhou, A. (2019). *Interpolated sonde (INTERPOLATEDSONDE)* [dataset]. Atmospheric Radiation Measurement (ARM) User Facility. https://doi.org/10.5439/1095316

Johnson, K., Giangrande, S., & Toto, T. (2019). *Active Remote Sensing of CLouds (ARSCL) Product Using Ka-Band ARM Zenith Radars (ARSCLKAZRBND1KOLLIAS)* [dataset]. Atmospheric Radiation Measurement (ARM) User Facility. https://doi.org/10.5439/1393438

Johnson, K., & Jensen, M. (2019). *Active Remote Sensing of CLouds (ARSCL) Product Using Ka-Band ARM Zenith Radars (ARSCLKAZR1KOLLIAS)* [dataset]. Atmospheric Radiation Measurement (ARM) User Facility. https://doi.org/10.5439/1228768

Keeler, E., Kyrouac, J., & Ermold, B. (2019). *Automatic weather station (MAWS)* [dataset]. Atmospheric Radiation Measurement (ARM) User Facility. https://doi.org/10.5439/1182027

Keeler, J. M., Jewett, B. F., Rauber, R. M., McFarquhar, G. M., Rasmussen, R. M., Xue, L., Liu, C., & Thompson, G. (2016). Dynamics of cloud-top generating cells in winter cyclones. Part I: Idealized simulations in the context of field observations. *Journal of the Atmospheric Sciences*, *73*(4), 1507–1527. https://doi.org/10.1175/JAS-D-15-0126.1

Kiviluoto, K. (1996). Topology preservation in self-organizing maps. *Proceedings of International Conference on Neural Networks (ICNN'96)*, *1*, 294–299 vol.1. https://doi.org/10.1109/ICNN.1996.548907

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, *43*(1), 59–69. https://doi.org/10.1007/BF00337288

Kohonen, T. (2001). *Self-Organizing Maps* (Vol. 30). Springer. https://doi.org/10.1007/978-3-642-56927-2

Kolstad, E. W., & Bracegirdle, T. J. (2008). Marine cold-air outbreaks in the future: An assessment of IPCC AR4 model results for the Northern Hemisphere. *Climate Dynamics*, *30*(7–8), 871–885. https://doi.org/10.1007/s00382-007-0331-0

Kosović, B., Juliano, T. W., Lackner, C. P., Xue, L., Geerts, B., & Oteng, N. (2023). *High-resolution Multiscale Simulations of Mixed-Phase Convective Clouds Observed during the Cold-Air Outbreaks in the Marine Boundary Layer Experiment (COMBLE)*. AGU Fall meeting, San Francisco.

Kyrouac, J., Shi, Y., & Tuftedal, M. (2019). *Surface meteorological instrumentation (MET)* [dataset]. Atmospheric Radiation Measurement (ARM) User Facility. https://doi.org/10.5439/1786358

Lackner, C. P., Geerts, B., Juliano, T. W., Xue, L., & Kosovic, B. (2023). Vertical Structure of Clouds and Precipitation During Arctic Cold-Air Outbreaks and Warm-Air Intrusions: Observations From COMBLE. *Journal of Geophysical Research: Atmospheres*, *128*(13), e2022JD038403. https://doi.org/10.1029/2022JD038403

Lackner, C. P., Geerts, B., Wang, Y., Juliano, T. W., Xue, L., Kosović, B., & Turner, D. D. (2023). Insights into the relation between vertical cloud structure and dynamics of three polar lows: Observations from COMBLE. *Quarterly Journal of the Royal Meteorological Society*, qj.4543. https://doi.org/10.1002/qj.4543

Martini, M. N., Gustafson, W. I., Yang, Q., & Xiao, H. (2014). Impact of resolution on simulation of closed mesoscale cellular convection identified by dynamically guided watershed segmentation. *Journal of Geophysical Research: Atmospheres*, *119*(22), 238. https://doi.org/10.1002/2014JD021962

McCoy, I. L., McCoy, D. T., Wood, R., Zuidema, P., & Bender, F. A. -M. (2023). The Role of Mesoscale Cloud Morphology in the Shortwave Cloud Feedback. *Geophysical Research Letters*, *50*(2). https://doi.org/10.1029/2022gl101042

McCoy, I. L., Wood, R., & Fletcher, J. K. (2017). Identifying Meteorological Controls on Open and Closed Mesoscale Cellular Convection Associated with Marine Cold Air Outbreaks. *Journal of Geophysical Research: Atmospheres*, *122*(21), 11,678-11,702. https://doi.org/10.1002/2017JD027031

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

Mohrmann, J., Wood, R., Yuan, T., Song, H., Eastman, R., & Oreopoulos, L. (2021). Identifying meteorological influences on marine low-cloud mesoscale morphology using satellite classifications. *Atmospheric Chemistry and Physics*, *21*(12), 9629–9642. https://doi.org/10.5194/acp-21-9629-2021

Muhlbauer, A., McCoy, I. L., & Wood, R. (2014). Climatology of stratocumulus cloud morphologies: Microphysical properties and radiative effects. *Atmospheric Chemistry and Physics*, *14*(13), 6695–6716. https://doi.org/10.5194/acp-14-6695-2014

Oue, M., Tatarevic, A., Kollias, P., Wang, D., Yu, K., & Vogelmann, A. M. (2020). The Cloud-resolving model Radar SIMulator (CR-SIM) Version 3.3: Description and applications of a virtual observatory. *Geoscientific Model Development*, *13*(4), 1975–1998. https://doi.org/10.5194/gmd-13-1975-2020

Papritz, L., & Sodemann, H. (2018). Characterizing the Local and Intense Water Cycle during a Cold Air Outbreak in the Nordic Seas. *Monthly Weather Review*, *146*(11), 3567–3588. https://doi.org/10.1175/MWR-D-18-0172.1

Pithan, F., Svensson, G., Caballero, R., Chechin, D., Cronin, T. W., Ekman, A. M. L., Neggers, R., Shupe, M. D., Solomon, A., Tjernström, M., & Wendisch, M. (2018). Role of air-mass transformations in exchange between the Arctic and mid-latitudes. *Nature Geoscience*, *11*(11), 805–812. https://doi.org/10.1038/s41561-018-0234-1

Plummer, D. M., McFarquhar, G. M., Rauber, R. M., Jewett, B. F., & Leon, D. C. (2015). Microphysical Properties of Convectively Generated Fall Streaks within the Stratiform Comma Head Region of Continental Winter Cyclones. *Journal of the Atmospheric Sciences*, *72*(6), 2465–2483. https://doi.org/10.1175/JAS-D-14-0354.1

Rantanen, M., Karpechko, A. Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., Vihma, T., & Laaksonen, A. (2022). The Arctic has warmed nearly four times faster than the globe since 1979. *Communications Earth and Environment*, *3*(1), 1–10. https://doi.org/10.1038/s43247-022-00498-3

Ritsche, M. T., & Prell, J. (2011). *ARM Surface Meteorology Systems Instrument Handbook* (DOE/SC-ARM/TR-086). PNNL; Richland, WA. https://doi.org/10.2172/1007926

Rosenow, A. A., Plummer, D. M., Rauber, R. M., McFarquhar, G. M., Jewett, B. F., & Leon, D. (2014). Vertical velocity and physical structure of generating cells and convection in the comma head region of continental winter cyclones. *Journal of the Atmospheric Sciences*, *71*(5), 1538–1558. https://doi.org/10.1175/JAS-D-13-0249.1

Serreze, M. C., Barrett, A. P., Stroeve, J. C., Kindig, D. N., & Holland, M. M. (2009). The emergence of surface-based Arctic amplification. *Cryosphere*, *3*(1), 11–19. https://doi.org/10.5194/tc-3-11-2009

Serreze, M. C., & Francis, J. A. (2006). The arctic amplification debate. *Climatic Change*, *76*(3–4), 241–264. https://doi.org/10.1007/s10584-005-9017-y

Shupe, M. D. (2007). A ground-based multisensor cloud phase classifier. *Geophysical Research Letters*, *34*(22), 1–5. https://doi.org/10.1029/2007GL031008

Taylor, P. C., Cai, M., Hu, A., Meehl, J., Washington, W., & Zhang, G. J. (2013). A decomposition of feedback contributions to polar warming amplification. *Journal of Climate*, *26*(18), 7023–7043. https://doi.org/10.1175/JCLI-D-12-00696.1

Turner, D. D., Clough, S. A., Liljegren, J. C., Clothiaux, E. E., Cady-Pereira, K. E., & Gaustad, K. L. (2007). Retrieving liquid water path and precipitable water vapor from the atmospheric radiation measurement (ARM) microwave radiometers. *IEEE Transactions on Geoscience and Remote Sensing*, *45*(11), 3680–3689. https://doi.org/10.1109/TGRS.2007.903703

Van Weverberg, K., Giangrande, S., Zhang, D., Morcrette, C. J., & Field, P. R. (2023). On the Role of Macrophysics and Microphysics in Km-Scale Simulations of Mixed-Phase Clouds during Cold Air Outbreaks. *Journal of Geophysical Research: Atmospheres*, 1–37. https://doi.org/10.1029/2022JD037854

Vincent, L., & Soille, P. (1991). Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *13*(6), 583–598. https://doi.org/10.1109/34.87344

Wang, M., Giangrande, S., Johnson, K., & Jensen, M. (2019). *Improved MICROBASE Product with Uncertainties (MICROBASEKAPLUS)* [dataset]. Atmospheric Radiation Measurement (ARM) User Facility. https://doi.org/10.5439/1768890

Wood, R., & Hartmann, D. L. (2006). Spatial variability of liquid water path in marine low cloud: The importance of mesoscale cellular convection. *Journal of Climate*, *19*(9), 1748–1764. https://doi.org/10.1175/JCLI3702.1

Wu, P., & Ovchinnikov, M. (2022). Cloud Morphology Evolution in Arctic Cold-Air Outbreak: Two Cases During COMBLE Period. *Journal of Geophysical Research: Atmospheres*, *127*(10), 1–20. https://doi.org/10.1029/2021JD035966

Zaremba, T. J., Heimes, K., Rauber, R. M., Geerts, B., French, J. R., Grasmick, C., Tessendorf, S. A., Xue, L., Friedrich, K., Rasmussen, R. M., Kunkel, M. L., & Blestrud, D. R. (2022). Vertical Motions in Orographic Cloud Systems over the Payette River Basin. Part II: Fixed and Transient Updrafts and Their Relationship to Forcing. *Journal of Applied Meteorology and Climatology*, *61*(11), 1733–1751. https://doi.org/10.1175/JAMC-D-21-0229.1

Zhang, D. (2019). *MWR retrievals (MWRRET1LILJCLOU)* [dataset]. Atmospheric Radiation Measurement (ARM) User Facility. https://doi.org/10.5439/1027369

Zhang, D., & Levin, M. (2019). *Thermodynamic Cloud Phase (THERMOCLDPHASE).* [dataset]. Atmospheric Radiation Measurement (ARM) user facility. https://doi.org/10.5439/1871014

Zhao, C., Xie, S., Chen, X., Jensen, M. P., & Dunn, M. (2014). Quantifying uncertainties of cloud microphysical property retrievals with a perturbation method. *Journal of Geophysical Research: Atmospheres*, *119*(9), 5375–5385. https://doi.org/10.1002/2013JD021112