Objective Classification of Tropical Mesoscale Convective Systems

MICK POPE

Bureau of Meteorology Training Centre, Bureau of Meteorology, Melbourne, Victoria, Australia

CHRISTIAN JAKOB AND MICHAEL J. REEDER

School of Mathematical Sciences, Monash University, Melbourne, Victoria, Australia

(Manuscript received 7 August 2008, in final form 8 June 2009)

ABSTRACT

A cluster analysis is applied to the mesoscale convective systems (MCSs) that developed in northern Australia and the surrounding oceans during six wet seasons (September–April) from 1995/96 to 2000/01. During this period, 13 585 MCSs were identified and tracked using an infrared channel (IR1) on the Japanese Meteorological Agency Geostationary Meteorological Satellite 5 (*GMS5*). Based on the lifetimes of the MCSs, the area covered by cloud, the expansion rate of the cloud, the minimum cloud-top temperature, and their zonal direction of propagation, the MCSs are grouped objectively into four classes. One of the strengths of the analysis is that it objectively condenses a large dataset into a small number of classes, each with its own physical characteristics.

MCSs in class 1 (short) are relatively short lived, with 95% having lifetimes less than 5 h, and they are found most frequently over the oceans during the early and late parts of the wet season. MCSs in classes 2 and 3 [long and intermediate west (Int-West)] are longer lived and propagate to the west, developing over continental northwest Australia in deep easterly flow during breaks in the monsoon. These two classes are distinguished principally by their lifetime, with 95% of MCSs in the long class having lifetimes exceeding 4 h. Class 4 (Int-East) comprises MCSs that form over the subtropical latitudes of eastern Australia and in the deep westerly flow over northern parts of the continent during the monsoon and active phases of the MJO.

1. Introduction

Satellite-based cloud climatologies of mesoscale convective systems (MCSs) have been constructed using infrared (IR) observations from geostationary satellites (Williams and Houze 1987; Miller and Fritsch 1991; Zuidema 2003; Machado and Laurent 2004; Kondo et al. 2006; Pope et al. 2008) and used to study various aspects of convection in the tropics. For instance, studies such as these have examined the seasonal and diurnal cycles of tropical convection and the relationship between the initial development of MCSs and their lifetimes. One advantage of the IR observations taken from geostationary satellites is that the time resolution is higher than that of instruments mounted on polar-orbiting satellites; consequently, these observations can be used to track individual convective systems. One disadvantage of IR observations is the relatively poor relationship between cloud-top temperature (CTT) and the underlying precipitation. Yuter and Houze (1998) showed that, over the western tropical Pacific, there was a poor relationship between cloud-top temperatures less than 235 K and precipitating area and between cloud top temperatures less than 208 K and convective rainfall. Liu et al. (2007) also showed that the relationship between cloud top temperatures less than 210 K and the height of the 20-dBZ echo, an indicator of precipitation intensity, shows a large geographical variation, being poorer over the oceans compared to over continents.

The volume of information generated by climatological studies of MCSs can be overwhelming. For this reason, many studies have sought to classify and hence condense the information so as to analyze and interpret the physical significance of the data. In most cases, however, the classification schemes used have been defined subjectively. For example, Maddox (1980) defined a mesoscale convective complex (MCC) based on

Corresponding author address: Mick Pope, Bureau of Meteorology Training Centre, Bureau of Meteorology, GPO Box 1289, Melbourne, VIC 3001, Australia.
E-mail: m.pope@bom.gov.au

DOI: 10.1175/2009JCLI2777.1

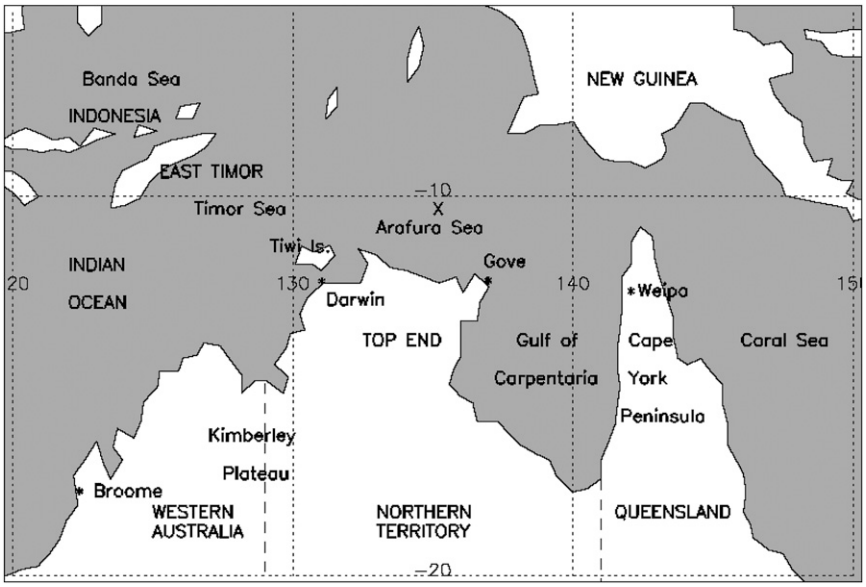

FIG. 1. Map of the study region for the north Australian monsoon.

size thresholds for two CTT thresholds, lifetime, and morphology (eccentricity at time of maximum areal extent). The lower CTT threshold (221 K) was chosen to ensure that a large area was precipitating, and the size and lifetime criteria of Maddox (1980) were chosen to ensure MCCs identified the largest and longest-lived MCSs. However, Youter and Houze (1998) have shown that the relationship between CTT and precipitating area is poor (Chen and Houze 1998), and both lifetime and size follow power-law distributions (Pope et al. 2008), suggesting that convection is scale invariant and hence highlight the arbitrary nature of the thresholds of Maddox (1980). Mathon et al. (2002) identified propagating MCSs using IR data, which they called organized convective systems. The organized convective systems were found to produce most of the observed rainfall over the Sahel. This close relationship between the organized convective systems of Mathon et al. (2002) and rainfall was achieved by a subjective a posteriori choice of characteristics of MCSs.

Using the dataset of Pope et al. (2008), the principal objective of the present study is to define an objective and physically meaningful classification of the MCSs that develop in the north Australian monsoon. The physical significance of this objective classification is examined by comparing aspects of the environment in which the derived MCSs classes occur, in particular the seasonal cycle, geographical distribution, and state of the monsoon circulation.

The dataset used in this study, together with the technique employed for the objective classification and the resulting classes, are described in section 2. Section 3 examines the seasonal cycle of the number of MCSs for each class. Section 4 examines the geographical distribution of the genesis of MCSs for each class and relates differences in genesis location to the mean zonal wind field. Section 5 relates intraseasonal changes in the relative numbers of each of the MCS classes to the state of the Madden–Julian oscillation (MJO). A discussion and summary of the major results are presented in section 6.

2. Objectively defined mesoscale convective system classes

a. Dataset

The present study uses the MCS dataset of Pope et al. (2008). Mesoscale convective systems making up the dataset were identified and tracked using the IR1 channel (10.5–11.5 μm) of the Japanese Meteorological Agency Geostationary Meteorological Satellite 5 (*GMS5*). The geographical region of the analysis is 5°–25°S, 120°–150°E (Fig. 1), with a pixel resolution of 5×5 km^2 at the satellite subpoint. The period 1995–2001 was chosen since hourly imagery is generally available, the main exceptions being a small gap in the 1998/99 season and the 1330 and 1430 UTC images, of which approximately 38% are missing. Each season consists of the months September–April. The MCSs tracking procedure followed the algorithm of Williams and Houze (1987), which was developed further by Mapes and Houze (1993). A convective element was taken to be a MCS if it covered at least 5000 km^2 for at least 2 h. The size threshold is arbitrary, but the minimum area needs to be sufficiently large to ensure that the MCS tracking technique correctly identifies the

motion of a MCS through its lifetime. Two temperature thresholds were used to identify MCSs: 208 and 235 K. The choice of temperature threshold is also arbitrary, but these two values are commonly used [for a discussion of the choice of temperature thresholds, see Mapes and Houze (1993) and Pope et al. (2008)]. The discussion that follows only considers the results for the 208-K threshold. The 235-K results were used to check the consistency of the clustering technique.

b. Clustering technique

Cluster analysis, which is a technique used to group data by identifying the centers of natural clusters in the dataset, has been used in meteorological applications (e.g., Gong and Richman 1995; Jakob and Tselioudis 2003; Arnott et al. 2004; Rossow et al. 2005). Cluster analysis objectively partitions the data by identifying the relative constellations of contiguous points in the dataset. Given that meteorological data is typically continuously but not necessarily uniformly distributed, cluster analysis does not identify discrete structures but those regions of the data that are relatively externally isolated and internally cohesive (Gong and Richman 1995). It is therefore inappropriate to discuss divisions between the clusters; rather, we consider the distribution around the cluster centroid. The particular algorithm used here is the iterative *K*-means cluster analysis algorithm (Afifi et al. 2004). In this algorithm, each element of the dataset is assigned to one of *K* clusters. Convergence of the algorithm is achieved when the membership of each cluster does not change with successive iterations.

One shortcoming of the *K*-means algorithm is that that number of clusters searched for needs to be prescribed. Several quasi-objective measures to decide on the optimal cluster number have been proposed by Rossow et al. (2005). Here, an objective technique for choosing the appropriate number of clusters is developed using ensemble runs of the clustering algorithm. For each ensemble run, the first-guess cluster centroids are randomly chosen. The resulting values of the cluster centroids for each ensemble member are then plotted and visually compared (not shown). A large spread between the cluster centroids for each ensemble run indicates that the solutions are sensitive to the initial first-guess cluster centroid, whereas a small spread in cluster centroids indicates that the solution is relatively insensitive to the initial choice and represents the natural number of clusters in the dataset. The results for the ideal number of clusters in this study showed essentially no spread in the values of the cluster centroids.

The physical properties used define the clusters are the lifetime of the MCS, the maximum equivalent radius during the life of the MCS [defined as $R_{\mathrm{MAX}} = (\text{lifetime maximum cloudy area}/\pi)^{1/2}$], the minimum single pixel temperature during the life of the MCS T_{MIN}, the lifetime average zonal propagation speed U_p, and the normalized cloud area expansion rate during the first hour after appearance above the minimum size threshold [defined as $A_E = (dA/dt)/\bar{A}$, where $\bar{A}$ is the mean area over the first hour and dA/dt is the change in area during the first hour; Machado and Laurent (2004)]. The zonal propagation speed is chosen to reflect differences in the broadscale environment. During the months from September to April, the synoptic environment changes from southeasterly trades to alternating periods of westerly and easterly monsoon conditions (Drosdowsky 1996); consequently, the magnitude and direction of the deep tropospheric winds should be related to the propagation direction of the center of the cloudy area of a MCS. The normalized cloudy expansion rate over the first hour of a MCS's life has been shown to be a useful predictor for the lifetime of the MCS (Machado and Laurent 2004; Pope et al. 2008). Four of the MCS parameters—lifetime, minimum lifetime single pixel temperature, maximum lifetime equivalent radius, and initial expansion rate—are correlated with each other (Pope et al. 2008). However, each of the four parameters is not entirely dependent on the others. For example, Fig. 17a of Pope et al. (2008) shows the relationship between the initial expansion rate and lifetime of the MCS. The spread in values of initial expansion rate increases with increasing lifetime, suggesting that the lifetime of a MCS is not solely determined by the initial expansion of the cloudy area. This further suggests that the inclusion of correlated parametersis not adding redundant information to the clustering algorithm. Hence, all five of the parameters selected were used in the clustering. Each of the five input parameters spans a different range and needs to be normalized. This is achieved by fitting the probability distribution function of the parameters with either a normal or lognormal distribution and then rescaling the data to a normal distribution with a mean of zero and standard deviation of one.

c. MCS classes

The groups found by the *K*-means cluster analysis of the MCS dataset are referred to as MCS classes. The clustering was performed on MCSs identified using 208- and 235-K thresholds. Using the randomly seeded ensemble technique described above, the number of classes chosen for both temperature thresholds is found to be four. The average value for a given parameter does not change much between temperature thresholds; hence, in the discussion that follows, only the 208-K classes will be shown.

Of the 13 585 MCSs at the 208-K threshold used in the clustering, 3873 occur in class 1, 2588 in class 2, 3151 in

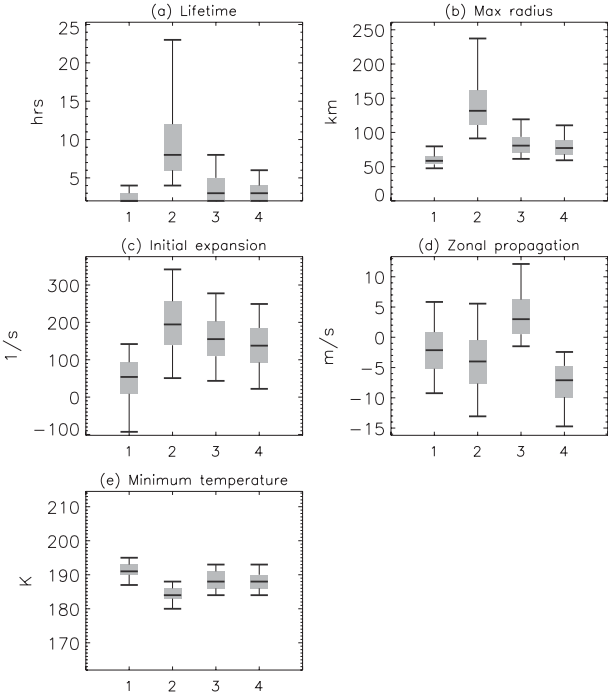

FIG. 2. Box and whisker plots showing the properties of four convective system classes of mesoscale convective systems identified using a 208-K cloud top temperature threshold. Properties shown are (a) system lifetime (hours); (b) maximum lifetime size, as measured by an equivalent radius, assuming a circular cloudy area (km); (c) the normalized expansion rate of the cloudy area over the first hour after system genesis (s^{-1}); (d) the mean lifetime zonal propagation speed ($m\ s^{-1}$); and (e) lifetime minimum single pixel temperature (K). The numbers along the abscissa are the four classes of MCSs: 1) short-lived, 2) long-lived, 3) intermediate-lifetime eastward propagating, and 4) intermediate-lifetime westward propagating.

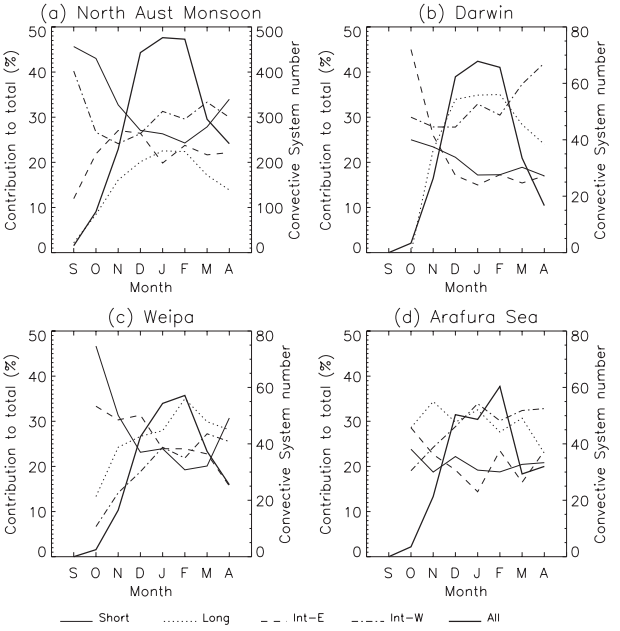

FIG. 3. Seasonal cycle of the relative contribution of each of the four MCS classes to the total number of MCSs observed per month (left-hand ordinate and thin lines) and the total number of MCSs observed per month (right-hand ordinate and thick solid line) (a) in the north Australian monsoon region defined by Fig. 1 and in a 5° × 5° box centered on (b) Darwin, (c) Weipa, and (d) 10°S, 135°E in the Arafura Sea.

class 3, and 3973 in class 4. The characteristics of the four classes are shown in Fig. 2. Class 1 has the shortest lifetimes (with 95% having lifetimes less than 5 h), smallest areas (95% have $R_{MAX} \leq 80$ km), largest lifetime minimum pixel temperatures (95% have $T_{MIN} \geq 187$ K), and weakest expansion rates. Negative expansion rates imply that some MCSs in this class reach their maximum size at their time of first appearance above the minimum size threshold (referred to as genesis) and thereafter contract. Members of this class propagate mostly westward ($U_p < 0$). This class is referred to as the short class.

Class 2 has the longest lifetimes (95% have lifetimes ≥4 h and 50% ≥6 h), largest lifetime maximum sizes (95% have $R_{MAX} \geq 92$ km and 50% have $R_{MAX} \geq 130$ km), most rapid initial expansion, and lowest lifetime minimum pixel temperatures. Members of this class also propagate mostly westward. This class is referred to as the long class. Note that the long lifetimes, large maximum radius, and direction of propagation identify approximately half of the MCS in class 2 with MCCs as defined by Miller and Fritsch (1991).

Classes 3 and 4 have lifetimes, sizes, expansion rates, and minimum lifetime single pixel temperatures intermediate between the short and long classes. Class 3 comprises almost entirely eastward propagating MCSs ($U_p > 0$), whereas class 4 is comprises almost entirely westward propagating MCSs ($U_p < 0$). These two classes are referred to as intermediate east (Int-East) and intermediate west (Int-West) classes, respectively.

The cluster analysis objectively condenses a dataset of 13 585 MCSs into four classes. The extent to which each class has physical meaning is examined next.

3. Seasonal cycle of MCS classes

Tropical convection over northern Australia is closely linked to the seasonal cycle of the north Australian monsoon, with over 90% of the annual rainfall falling between the months from November to April (Nicholls et al. 1982). Figure 3 shows the number of MCSs from September to April. The relative number of MCSs in each class and the total number of MCSs are plotted for the entire study region and for 5° × 5° boxes centered on Darwin, Weipa, and the Arafura Sea (10°S, 135°E; marked with an X in Fig. 1). Few MCSs form over the

study area during September and October; hence, the relative contributions to each class probably has little significance at this time. The total MCS number reaches its peak during the months from December to February, with a sharp rise during November and a sharp decrease during March.

The relative contribution from the long class follows the seasonal cycle of total number of MCSs, reaching a peak during December–February over continental Australia, in broad agreement with the results of Miller and Fritsch (1991) for their MCCs. Over the Arafura Sea (Fig. 3d) the seasonal cycle of the long class is relatively flat with a minimum contribution during October and April. In contrast, the short class makes its major contribution to the total number of MCSs outside of the peak months of December–February, the main exception being Broome (Fig. 1) where the short class also peaks during January (not shown). Broome is located near 17°S, and this difference between the seasonal cycle of the short class here and other parts of the north Australian region may be related to the movement of the monsoon trough (see below).

The relative contribution of the Int-West class increases throughout the season, typically reaching a maximum during March and April, although a peak is found during January also, particularly over the Arafura Sea (Fig. 3d) where the Int-West class makes the largest contribution to the total number of MCSs. The Int-East class makes its largest contribution during November and December with a secondary maximum during February (Fig. 3a). Typically, the maximum contribution from the Int-East class occurs outside of the monsoon season, which begins during mid to late December at Darwin (Drosdowsky 1996). During the active monsoon, deep westerly winds prevail, suggesting that MCSs should propagate mostly eastward. A peak in the relative contribution from the Int-East class for the entire region (Fig. 3a), Darwin (Fig. 3b), and the Arafura Sea (Fig. 3d) during February confirms some relationship between the monsoon and the occurrence of the eastward propagating MCSs. However, a large number of Int-East class MCSs occur before December, implying that the Int-East class cannot be identified solely with a monsoon environment; this observation is investigated in more detail in the next section.

4. Geographical distribution of MCS classes

In section 3 it was shown that the four MCS classes have seasonal cycles that are different from each other. A simple identification of the intermediate lifetime classes with changes in the state of the monsoon is not straightforward, given the dominance of the Int-West over the Int-East during the monsoon months from December to February. The relationship between the MCS classes and the synoptic environment in which they form is examined now.

Figure 4 shows the geographical distribution of genesis of MCSs for all four classes. The distributions are plotted as the relative contribution to the total number of MCSs on a $2° \times 2°$ grid for September–April. Figure 3 shows that most of these occur during December–February. The geographical locations at which the MCSs form during December–February (not shown) are similar to those for the entire season (Fig. 4). MCSs in the short class form more commonly over water than over land. The long class forms most commonly over the northwest coast of Australia, consistent with the results of Miller and Fritsch (1991) for MCCs. One of the most striking features of Fig. 4 is that the most favored regions for the genesis of MCSs, the Int-East and Int-West classes, are sharply separated: the Int-East class forms most commonly over the east coast of Cape York Peninsula, whereas the Int-West class forms mainly over the western half of northern Australia and surrounding waters.

Figure 5 shows the geographical distribution of all occurrences of the four MCS classes from September to April as the relative contribution to the total number of MCSs observed. The short class makes its largest contribution to MCS numbers with a similar distribution to the genesis distribution (Fig. 4). MCSs in the short class are short lived and weakly propagating and hence do not move far from their genesis location. The long class makes its largest contribution to MCS numbers to the west of its genesis locations, consistent with the long lifetime and westward propagation of MCSs in this class. The Int-East class makes its largest contribution to MCS numbers slightly to the east of its genesis locations, consistent with the eastward propagation and intermediate lifetimes of MCSs in this class. Likewise, the results for the Int-West are consistent with the main genesis locations and their propagation.

The east–west division in the most common regions for the genesis of MCSs in intermediate classes is related to differences in the synoptic wind regime. Figure 6 shows the correlation between the daily mean zonal wind from the 40-yr European Centre for Medium-Range Weather Forecasts Re-Analysis (ERA-40) (Uppala et al. 2005) at 700, 500, and 200 hPa and the daily mean number of MCSs for each class. All calculations are made on a 2.5×2.5 grid for the months December–February. An equivalent calculation using the deep-layer pressure-weighted mean zonal wind is plotted also in Fig. 6, where the deep-layer mean is defined using the levels 925, 850, 700, and 500 hPa. There is a consistent positive (negative) correlation between the zonal wind and number of MCSs in the

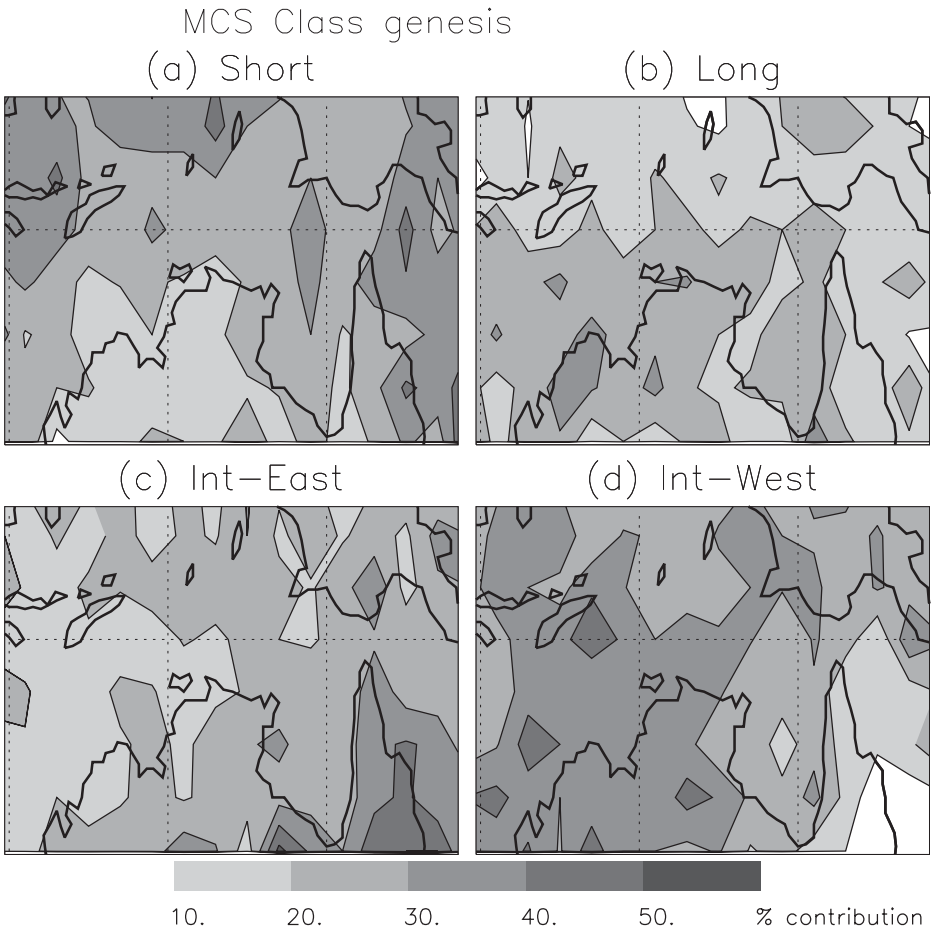

FIG. 4. Genesis locations of the convective system classes: (a) short-lived, (b) long-lived and intermediate lifetime, (c) eastward propagating, and (d) westward propagating for September–April 1995/96 to 2000/01 of lifetime ≥2 h plotted as the relative contribution to the total number of convective systems generated. Results are plotted on a 2° × 2° grid.

Int-East class (Int-West class), excluding September for which the number of MCSs is small. These correlations are largest for the 500-hPa zonal wind in all months, excluding September. The correlations between the number of MCSs in each class and the deep-layer mean zonal wind are small because the correlation with the zonal winds at the lowest level is poor (not shown). In particular, the mean December–February 925-hPa zonal wind over the northwest coast of Australia (not shown) is westerly owing to the heat low/trough system (Arnup and Reeder 2007). The correlations for the short class are typically small. They vary throughout the season from near zero or negative in the period to November and become positive after that, reflecting the change in mean winds associated with the monsoon.

The correlations for the long class vary seasonally as well. Again, with the exception of September, the correlations in the midtroposphere are mostly negative, implying that these systems form in the break monsoon periods and propagate westward. The correlation at 200 hPa suggests that some of the observed motion of MSCs in this class during January–March is due to the advection of cirrus near 200 hPa. Figure 2 shows that the long class (class 2) typically propagates more slowly than the Int-West class, and the weaker negative correlation of midtropospheric winds from the ERA-40 with the daily number of MCSs is consistent with this. Figure 2 also shows that the long class contains some eastward propagating systems (<25%). The positive correlation between the number of MCSs each day and the 200-hPa zonal winds from the ERA-40 during February and March suggests that advection of cirrus by the upper westerlies may contribute to the observed eastward propagating long class MCSs.

The geographical distribution of the four classes of MCSs can be explained by the December–February average zonal wind at 700 and 500 hPa, which is plotted in Fig. 7. Instead of an average over the whole season,

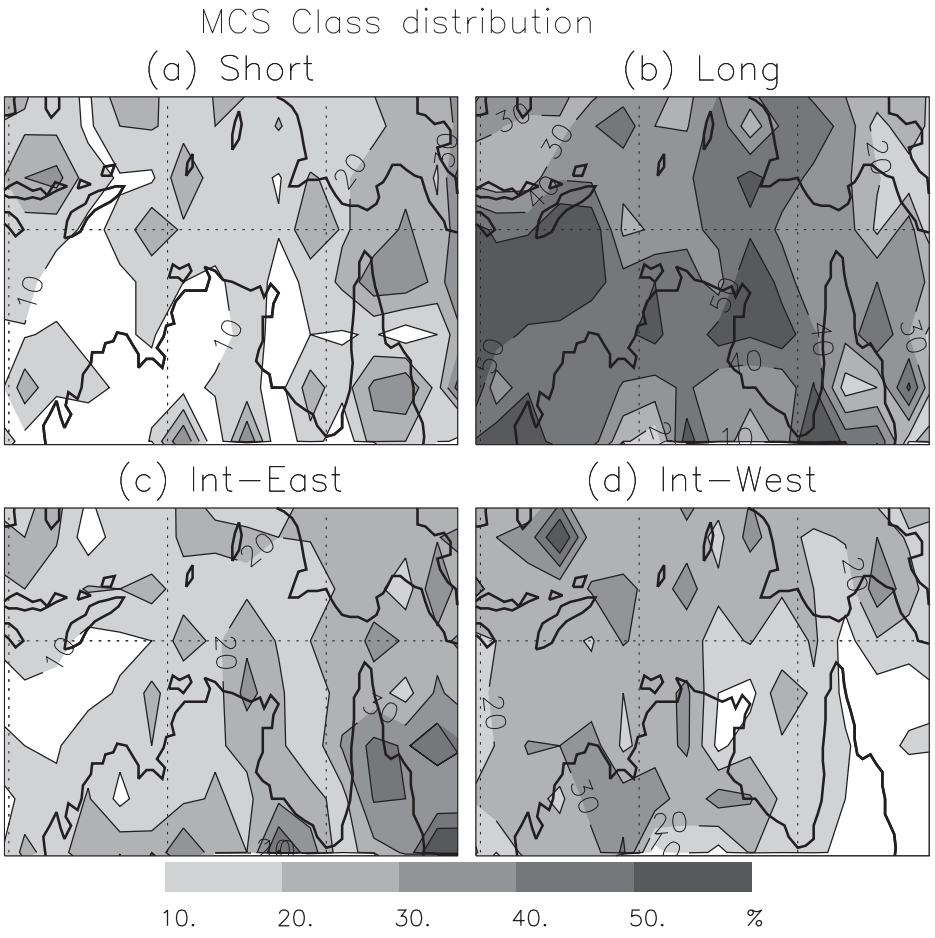

FIG. 5. As in Fig 4, but of the geographical distribution of the convective system classes.

the average for December–February is used since the majority of MCSs are observed during these months. Comparing the geographical distribution of the zonal winds to the most common regions for the genesis of MCSs in each class (Fig. 4) shows that the Int-East regime makes a large contribution to the genesis of MCSs in an environment with mean easterly winds at 700 hPa and westerly winds at 500 hPa. This result is consistent with a mean trade wind regime with mid- to upper-level westerlies. The Int-East class MCSs located equatorward of 10°S are more likely to be associated with monsoon conditions than those poleward of 10°S, because the zonal winds are westerly through the depth of the troposphere near the equator in the mean during December–February.

The Int-West class makes its largest contribution to the genesis of MCSs in the western half of the domain (Fig. 3), where the December–February mean zonal winds are easterly (Fig. 7). This area includes Darwin, which alternates between deep easterly and deep westerly winds during these months due to the monsoon (Drosdowsky 1996). The occurrence of these westerly wind bursts over Darwin is highly variable. Drosdowsky (1996) found that deep westerly winds occurred on average for 39 days from December to March, which is approximately one-third of all days. This is consistent with Fig. 7, where mean easterlies prevail over the top end of Australia (Fig. 1) and hence the Int-West class dominates on this time scale (Fig. 3).

The short class makes its largest contribution to the genesis of MCSs in the eastern half of the domain and equatorward of 10°S, mostly over the oceans. These are regions of average deep westerly zonal winds (Fig. 7), consistent with the correlations with zonal wind (Fig. 6). The mean properties of the short class show them to be mainly westward propagating (Fig. 2). However, the correlations in Fig. 6 are small, implying large variability in the wind environment in which they form.

5. Intraseasonal variation of MCS classes

In section 4 it was shown that the most favored geographical regions for the genesis of MCSs in the four

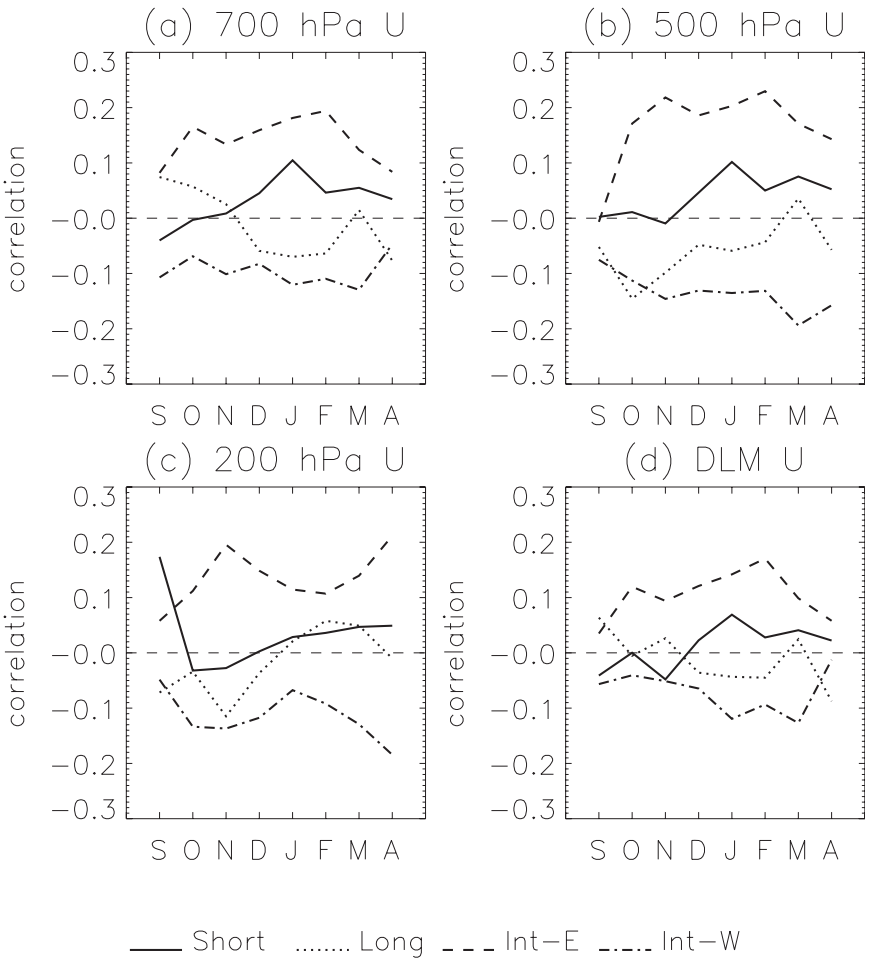

FIG. 6. The mean monthly daily correlation between the number of MCSs for each of the four MCS classes and the ERA-40 zonal winds at (a) 700 hPa, (b) 500 hPa, and (c) 200 hPa and (d) a deep-layer, pressured-weighted mean zonal wind using the levels at 925, 850, 700, and 500 hPa.

classes could be explained in part by the December–February midtropospheric zonal wind field. The wind over northern Australia varies intraseasonally, with alternating deep tropospheric westerly and easterly winds associated with the active and break phases of the monsoon (Drosdowsky 1996). How these changes affect the relative contribution of the four MCS classes is examined now.

Figures 8 and 9 show the relative contribution from each of the four classes to the total number of MCSs observed in a 5° × 5° box centered on Darwin and Weipa, respectively, with the results stratified into three regimes according to the 500-hPa zonal wind U_{500} from the 2300 UTC radiosonde ascent. The three wind regimes are (a) an easterly environment defined by $U_{500} < -5$ m s^{-1}, (b) a light wind environment defined by -5 m s$^{-1} < U_{500} < 5$ m s^{-1}, and (c) a westerly wind environment defined by $U_{500} > 5$ m s^{-1}. The long class is frequently observed in all conditions at both locations. Being the longest lived of the four classes, MCSs that belong to the long class are the most likely to propagate into an area such as Darwin or Weipa from outside. Plots of the four classes stratified by the zonal wind that consider only those systems generated within a 5° × 5° box surrounding the location (not shown) show that the long class makes a smaller contribution to MCS activity compared to Figs. 7 and 8.

In an easterly environment, the relative contribution to the total number of MCSs from those in the Int-West class is about two times greater than that from those in the Int-East class. Conversely, in a westerly environment, MCSs in the Int-East class make a relative contribution more than twice that from Int-West MCSs. In the weak wind environment the relative contribution of the intermediate regimes depends on the location, which is consistent with Fig. 4. The environment in which the short

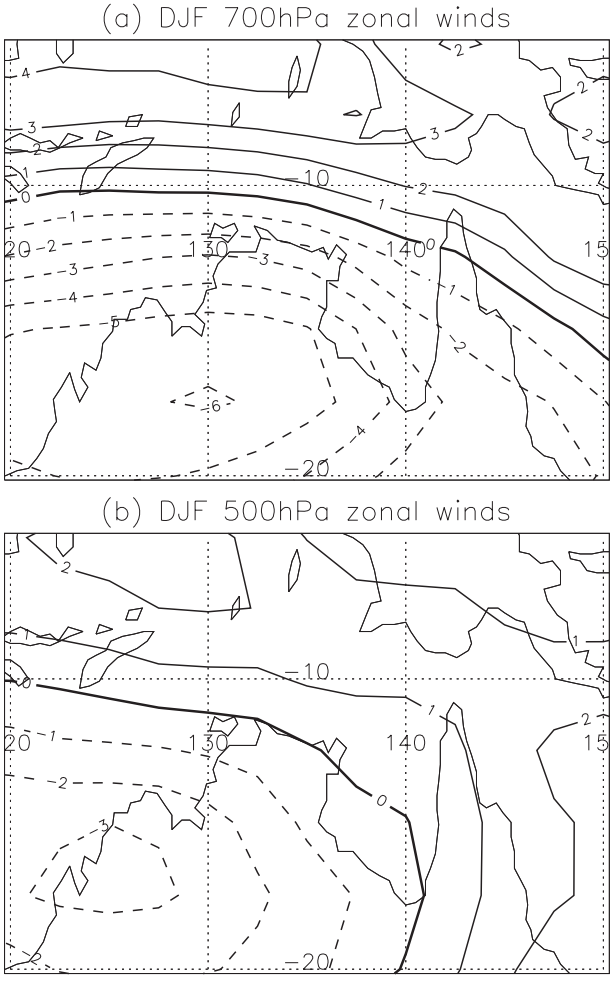

FIG. 7. Maps of the mean December–February zonal winds for (a) 700 hPa and (b) 500 hPa.

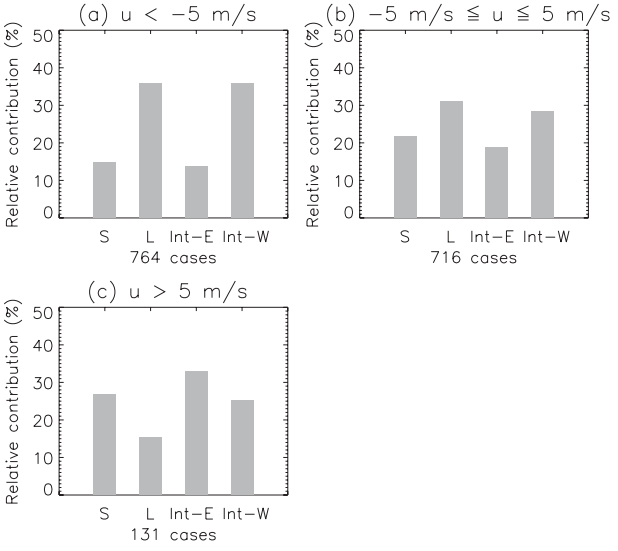

FIG. 8. The relative contribution of the four MCS classes passing within a 5° × 5° box around Darwin stratified by the 500-hPa zonal wind for wind classes (a) $U_{500} < -5$ m s^{-1}, (b) -5 m s$^{-1} \leq U_{500} \leq 5$ m s^{-1}, and (c) $U_{500} > 5$ m s^{-1}.

class makes its maximum contribution to total number of MCSs varies geographically, consistent with the weak correlation between zonal wind and the number of MCSs in the short class (Fig. 6).

One of the major influences on the intraseasonal variability of the north Australian monsoon is the Madden–Julian oscillation (Hendon and Liebmann 1990; Wheeler and Hendon 2004). Wheeler and Hendon developed a real-time MJO index based on a principle component analysis of the 850- and 200-hPa wind fields and outgoing longwave radiation (OLR). This real-time MJO index consists of eight phases describing the progression of the MJO from the Indian Ocean to the eastern Pacific. During the active phases of the MJO, negative OLR anomalies associated with enhanced convective activity are accompanied by 850-hPa westerly winds and 200-hPa easterly winds. Wheeler and Hendon found that, when the index had sufficiently large amplitude (a value of at least one), the onset of the first active phase of the MJO for the monsoon was associated with the onset of the first day of deep westerly winds associated with the monsoon at Darwin, which occurred during an active MJO phase. Figure 10 shows the relative contribution to the total number of MCSs in each class at each phase of the MJO, including all events where the amplitude of the real-time MJO index is less than one (since their removal does not significantly change results) for Darwin and Weipa. The total number of MCSs is modulated by the state of the MJO, reaching a maximum during phase 5 for Darwin and phase 6 at Weipa, consistent with the eastward propagation of the MJO. The relative contribution to each of the four MCS classes is affected also by the state of the MJO. The short class makes its maximum contribution during the suppressed early and late phases of the MJO, with peaks during phases 2 and 8 at Darwin and phase 3 at Weipa. This result is consistent with Fig. 2, which shows that MCSs in the short class have a weak tendency to propagate toward the west, suggesting that short class MCSs occur in a break monsoon period environment. The long class also makes its largest contribution to the total number of MCSs at Darwin during the early and late phases of the monsoon, consistent with a tendency to propagate westward (Fig. 2). At Weipa, the long class makes its maximum contribution during suppressed monsoon conditions (phase 2) and near the peak of the monsoon at Weipa (phase 5). Figure 2 shows that about 25% of the long class propagate eastward and, therefore, are associated with the active phase of the MJO.

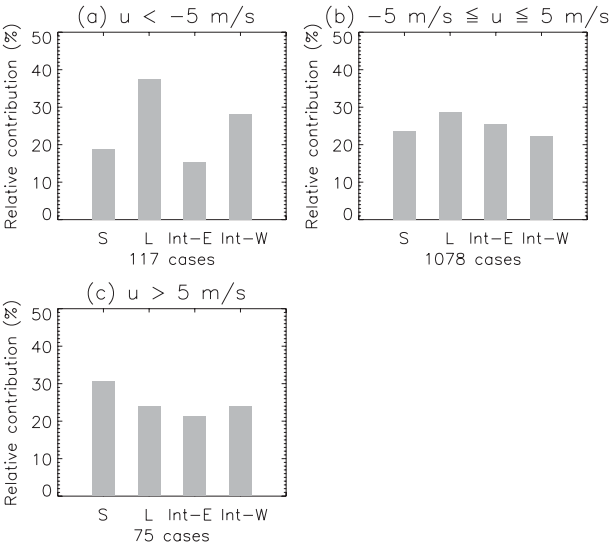

FIG. 9. As in Fig. 8, but around Weipa.

The phase of the MJO also affects the relative occurrence of the intermediate classes. At Darwin, the Int-West class (Int-East class) makes its largest (smallest) contribution during the suppressed phases of the MJO (3–4). The Int-East class makes its largest contribution to the number of MCSs at Darwin during phases 6–7, when the Int-West class makes its minimum contribution. Hence, the relative contribution of the two MCS classes changes with the state of the MJO. At Weipa the Int-West class makes its largest contribution to the number of MCSs during the suppressed phases 1–2 and 8. The Int-West class makes its smallest contribution to the number of MCSs during phase 6 at which time the Int-East class makes its largest contribution. As was the case for Darwin, this change in importance for the intermediate classes during phase 6 represents a change in synoptic conditions associated with the MJO. The peak in the number of Int-East class MCSs during phases 2–3 at Weipa is not associated with the active phase of the MJO (see section 4).

6. Discussion

The application of cluster analysis to a large dataset of mesoscale convective systems identified and tracked using IR data yields four physically meaningful classes of MCSs. MCSs are clustered by the algorithm according to their lifetime, direction of propagation, minimum lifetime cloud top temperature, maximum lifetime cloudy area, and cloudy area expansion rate over the first hour. The resulting classes can be described according to their lifetime and direction of propagation and are referred to as short, long, Int-East, and Int-West. Although all four classes can occur at any given location at any given time, the geographical distribution of the genesis locations of the MCS classes show that the results of the clustering are not random but reflect the synoptic environment in which the MCSs in those classes form. The occurrence of the Int-West class is related to the deep easterly steering that occurs on average during the monsoon season over the western half of the study area. The relative contribution of the Int-West class tends to increase during the season, even during months where deep westerly winds are expected. This is partly because deep westerly winds occur over northern Australia only on about one-third of all days during December–March (Drosdowsky 1996).

The Int-East class is not always associated the active monsoon, as shown by the geographical distribution

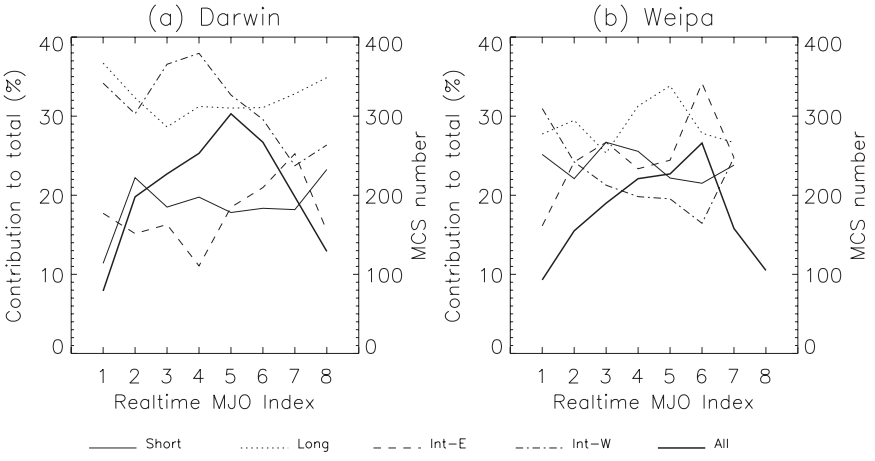

FIG. 10. The relative contribution to the total number of MCSs passing through a 5° × 5° box centered on (a) Darwin and (b) Weipa for the four MCS classes as well as the total number of MCSs for each of the eight phases of the Wheeler and Hendon (2004) real-time MJO index.

(Fig. 4). Instead, the Int-East class is often associated with midlatitude westerly winds in the midtroposphere (Fig. 6), consistent with the annual cycle (Fig. 3), which shows that the Int-East class occurs more often outside of the monsoon months of December–February. MCSs in the Int-East class that are located equatorward of 10°S are more likely to be associated with deep westerly flow associated with the active monsoon. The relative occurrence of the Int-East class is also modulated by changes in the midtropospheric flow as shown by the distributions at Darwin and Weipa, with more Int-East class MCSs being observed than Int-West class MCSs during a westerly wind environment (Figs. 7 and 8). A significant factor in the modulation of the synoptic wind and hence the relative occurrence of the intermediate classes is the MJO (Fig. 10).

The long class appears to be associated mostly with break period continental convection, particularly in the western half of the study region (Fig. 4) where the mean midtropospheric zonal winds during December–February are easterly (Fig. 7). With 50% of the long class having $R_{MAX} > 130$ km (corresponding to an area of ≥ 5300 km^2) and lifetimes of 6 or more hours and being mostly westward propagating, approximately half of this class may be identified with the MCCs of Miller and Fritsch (1991). Over continental Australia (Figs. 3b,c), the relative contribution of the long class to the total number of MCSs follows the cycle of total number of MCSs, peaking during December–February. The increase in the number of the long class over other classes may be linked to an increase in convective instability over the continent during the season. The long class has the most rapidly expanding initial cloudy area. Machado and Laurent (2004) showed that the initial expansion of the cloudy area was not related to the divergence of the upper wind field but, instead, suggested that it must be related to the transport of condensate to the upper levels of the atmosphere. The transport of condensate is related to updraft strength, suggesting a link between the state of the boundary layer and the occurrence of the long class. Miller and Fritsch (1991) noted that clusters of MCCs tended to form in regions where low-level jets of high equivalent potential temperature were found. McBride and Frank (1999) note that the highest values of equivalent potential temperature occurred during break monsoon periods when the steering winds were easterly, consistent with the westward propagation of most MCSs in the long class. Compared to that over continental Australia, the seasonal cycle of the long class over the oceans is comparatively weak.

The short class appears to slightly prefer a westerly wind environment (Fig. 6) over the oceans (Fig. 4), although short class MCSs are mostly westward propagating. Over continental Australia, the short class is more common early and late in the monsoon season. As short class MCSs have the smallest initial cloudy areas expansion rates, it is possible that they are associated with a more stable boundary layer with less convective available potential energy than the long class.

The technique developed here could readily be applied to other regions around the world. It is likely that the division of MCSs into short, intermediate, and long lifetimes would be observed; however, the direction of propagation of each class will depend on the large-scale steering. For example, Miller and Fritsch (1991) identified that most MCCs over the China region were eastward propagating compared to the Australia and New Guinea regions. This would likely be reflected in the direction of propagation for the long class over that region.

Future work will examine the differences between the MCS classes using independent data such as the rainfall and microphysical data from the Tropical Rainfall Measurement Mission (TRMM) satellite (Kummerow et al. 1998).

REFERENCES

Afifi, A., V. A. Clark, and S. May, 2004: *Computer-Aided Multivariate Analysis*. Chapman and Hall/CRC, 489 pp.

Arnott, J. M., J. L. Evans, and F. Chiaromonte, 2004: Characterization of extratropical transition using cluster analysis. *Mon. Wea. Rev.,* **132,** 2916–2937.

Arnup, S. J., and M. J. Reeder, 2007: The diurnal and seasonal variation of airmass boundaries in the Australian region. *Mon. Wea. Rev.,* **135,** 2995–3008.

Drosdowsky, W., 1996: Variability of the Australian summer monsoon at Darwin: 1957–1992. *J. Climate,* **9,** 85–96.

Gong, X., and M. B. Richman, 1995: On the application of cluster analysis to growing season precipitation data in North America east of the Rockies. *J. Climate,* **8,** 897–931.

Hendon, H. H., and B. Liebmann, 1990: A composite study of onset of the Australian summer monsoon. *J. Atmos. Sci.,* **47,** 2227–2240.

Jakob, C., and G. Tselioudis, 2003: Objective identification of cloud regimes in the tropical western Pacific. *Geophys. Res. Lett.,* **30,** 2082, doi:10.1029/2003GL018367.

Kondo, Y., A. Higuchi, and K. Nakamura, 2006: Small-scale cloud activity over the Maritime Continent and the western Pacific as revealed by satellite data. *Mon. Wea. Rev.,* **134,** 1581–1599.

Kummerow, C., W. Barnes, T. Kozu, J. Shiue, and J. Simpson, 1998: The Tropical Rainfall Measuring Mission (TRMM) sensor package. *J. Atmos. Oceanic Technol.,* **15,** 809–817.

Liu, C., E. J. Zipser, and S. W. Nesbitt, 2007: Global distribution of tropical deep convection: Different perspectives from TRMM infrared and radar data. *J. Climate,* **20,** 489–503.

Machado, L. A. T., and H. Laurent, 2004: The convective system area expansion over Amazonia and its relationships with convective system life duration and high-level wind divergence. *Mon. Wea. Rev.,* **132,** 714–725.

Maddox, R. A., 1980: Meoscale convective complexes. *Bull. Amer. Meteor. Soc.,* **61,** 1374–1387.

Mapes, B. E., and R. A. Houze Jr., 1993: Cloud clusters and superclusters over the oceanic warm pool. *Mon. Wea. Rev.,* **121,** 1398–1416.

Mathon, V., H. Laurent, and T. Lebel, 2002: Mesoscale convective system rainfall in the Sahel. *J. Appl. Meteor.,* **41,** 1081–1092.

McBride, J. L., and W. M. Frank, 1999: Relationships between stability and monsoon convection. *J. Atmos. Sci.,* **56,** 24–36.

Miller, D., and J. M. Fritsch, 1991: Mesoscale convective complexes in the western Pacific region. *Mon. Wea. Rev.,* **119,** 2978–2992.

Nicholls, N., J. L. McBride, and R. J. Ormerod, 1982: On predicting the onset of the Australian wet season at Darwin. *Mon. Wea. Rev.,* **110,** 14–17.

Pope, M., C. Jakob, and M. J. Reeder, 2008: Convective systems of the north Australian monsoon. *J. Climate,* **21,** 5091–5112.

Rossow, W. B., G. Tselioudis, A. Polak, and C. Jakob, 2005: Tropical climate described as a distribution of weather states indicated by distinct mesoscale cloud property mixtures. *Geophys. Res. Lett.,* **32,** L21812, doi:10.1029/2005GL024584.

Uppala, S. M., and Coauthors, 2005: The ERA-40 re-analysis. *Quart. J. Roy. Meteor. Soc.,* **131,** 2961–3012.

Wheeler, M., and H. H. Hendon, 2004: An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Mon. Wea. Rev.,* **132,** 1917–1932.

Williams, M., and R. A. Houze Jr., 1987: Satellite-observed characteristics of winter monsoon cloud clusters. *Mon. Wea. Rev.,* **115,** 505–519.

Yuter, S. E., and R. A. Houze Jr., 1998: The natural variability of precipitating clouds over the western Pacific warm pool. *Quart. J. Roy. Meteor. Soc.,* **124,** 53–99.

Zuidema, P., 2003: Convective clouds over the Bay of Bengal. *Mon. Wea. Rev.,* **131,** 780–798.