Objective determination of the extratropical transition of tropical cyclones in the Northern Hemisphere

By JOSHUA STUDHOLME[1]*, KEVIN I. HODGES[2] and
CHRIS M. BRIERLEY[1], [1]*Department of Geography, University College London, London, UK*;
[2]*Department of Meteorology, National Centre for Earth Observation, University of Reading, Reading, UK*

ABSTRACT

Extratropical transition (ET) has eluded objective identification since the realisation of its existence in the 1970s. Recent advances in numerical, computational models have provided data of higher resolution than previously available. In conjunction with this, an objective characterisation of the structure of a storm has now become widely accepted in the literature. Here we present a method of combining these two advances to provide an objective method for defining ET. The approach involves applying K-means clustering to isolate different life-cycle stages of cyclones and then analysing the progression through these stages. This methodology is then tested by applying it to five recent years from the European Centre of Medium-Range Weather Forecasting operational analyses. It is found that this method is able to determine the general characteristics for ET in the Northern Hemisphere. Between 2008 and 2012, 54% ($\pm$7, 32 of 59) of Northern Hemisphere tropical storms are estimated to undergo ET. There is great variability across basins and time of year. To fully capture all the instances of ET is necessary to introduce and characterise multiple pathways through transition. Only one of the three transition types needed has been previously well-studied. A brief description of the alternate types of transitions is given, along with illustrative storms, to assist with further study.

Keywords: tropical cyclones, extratropical transition, operational analyses, clustering

1. Introduction

A tropical cyclone (TC) can undergo profound structural and behavioural changes as it moves into the midlatitudes to become an extratropical cyclone, a process referred to as extratropical transition (ET) (Elsberry et al., 2000; Jones et al., 2003). ET results in an accelerated system producing intense rainfall, strong winds and large surface water waves (Jones et al., 2003). These systems not only play a role in the planetary climate system (Emanuel, 2008), but also have a strong and lasting effect on civilisation (e.g. Jones et al., 2003; Emanuel et al., 2012). Wind damage (Evans and Hart, 2008), inland flooding (DiMego and Bosart, 1982) and perilous maritime conditions (Sinclair, 2002) are all associated with ET. Identifying the ET of individual storms has historically been based upon a combination of the subjective analysis of satellite images by specialised individuals and a few objectively defined parameters

(Jones et al., 2003). This unsatisfactory practice is largely the consequence of a lack of understanding of the processes underlying ET and a scarcity of high-resolution data. Over the last decade many advances have been made in both areas and we now have sufficient understanding and resources to classify ET objectively.

There are many reasons why it is desirable to develop and establish a method for objectively determining ET. Fundamentally, as the ability of models to produce realistic simulations of smaller atmospheric systems, for example TCs, improves and the volume of observed weather and climate data grows, analysing it by means of human interpretation becomes increasingly unfeasible. So, there is a clear requirement for the establishment of a robust and computationally implementable scheme. One particular realm of interest is weather prediction. The acute effect of ET on midlatitude weather and climate prevents it from being ignored. Not least since the midlatitudes are the most populated areas of the planet (Small and Cohen, 2004). The modern method of numerical weather prediction (NWP) accounts for the inherently chaotic nature of

---

*Corresponding author.
email: joshua.studholme.11@ucl.ac.uk

This article has been republished with minor changes. These changes do not impact the academic content of the article.

the atmosphere by making multiple perturbations to input data and thus producing multiple forecasts (Froude et al., 2014). Analysing these multiple model outcomes (the European Center for Medium-Range Weather Forecasting produces 50 ensemble members, Buizza et al., 1999) every 6 hours (the regularity of most operational forecasting centres) is simply unfeasible. It is hoped that the scheme presented here could be helpful in solving this problem. Computationally implementable schemes could both normalise classification across time and be complimentary to the forecaster analysis. It would also provide the ability to objectively examine ET in climate models and any future changes.

Specifically, ET details the transformation of an axisymmetric, warm-core TC into an asymmetric, cold-core, vertically tilted extratropical cyclone (Muramatsu, 1985; Jones et al., 2003). ET can be said to be complete when the system has developed frontal characteristics. It is current practice to describe ET to occur as a two-stage process (Klein et al., 2000; Elsberry et al., 2000). Initially a storm is said to be going through the transformation stage if the warm-core vortex becomes baroclinic. The system then undergoes reintensification depending on the ambient environment through which the storm is passing. This often occurs in association with an upper-level trough (e.g. Ritchie and Elsberry, 2003; Elsberry and Ritchie, 2007).

At the turn of the millennium, TCs and extratropical cyclones had been compartmentalised into different areas of study (Hart and Evans, 2001). However, as satellite data became increasingly available, it came to be understood that storm systems transform into different types as they pass through their life cycle and into different environments (Jones et al., 2003). This realisation was followed by attempts to both develop a theory of and formally characterise ET as well as the development of basin-specific climatologies. Despite initial advances, ET was described in 2000 as being 'poorly understood and incompletely researched' (Elsberry et al., 2000). Over the last decade and half, a diverse and insightful literature has developed.

In a seminal paper, Jones et al. (2003) summarised the literature's state of knowledge for ET and galvanised future work by outlining questions yet to be answered. They stressed the importance of the development of an objective description of storm characteristics and outlined various candidates. One of these, the Cyclone Phase Space (CPS) proposed by Hart (2003) has now become widely used and studied (e.g. Arnott et al., 2004; Hart et al., 2006; Veren et al., 2009; Song et al., 2011). It has been assessed against various other frameworks and has been deemed effective (Kofron et al., 2010; Wang et al., 2012). The CPS is defined by three parameters, one for storm asymmetry and two thermal proxies for the upper and lower troposphere. Evans and Hart (2003) define the onset of ET in the

CPS as when the asymmetry parameter exceeds a critical value and to be complete when the lower troposphere's thermal proxy decreases below a certain value. In parallel with the development of this methodology, the last decade and a half has seen multiple climatologies for ET in various basins (e.g. Elsberry et al., 2000; Hart and Evans, 2001; Song et al., 2011). These marked the important transition from the pre-2000 work which was largely based on case-studies of individual storms to using statistics for a collection of events to describe the phenomenon basin by basin. However, all attempts to describe and explain ET has been hampered by two limiting factors: firstly, the lack of an all-encompassing and objective definition of ET, and secondly, a severe lack of data of sufficient resolution to examine storm structure. Therefore, the aim of this work is to develop a methodology that can utilise newly available high-resolution data in which TC structure is adequately simulated to surpass these historical limitations.

Since their first conception, models of Earth's climate system have continued to evolve ever-finer resolution and to incorporate characterisation of previously unresolved processes. However, even in 2006 validation of cyclone structure in forecasts by operational numerical models had yet to be performed (Evans et al., 2006) despite the fact that tropical cyclone-like vortices had been known to be simulated in even quite low-resolution models (e.g. Bengtsson et al., 1995). The first attempt to numerically model TCs in 1969 acknowledged the formidable difficulties in this task (Ooyama, 1969). Many advances have been made since then but we are yet to achieve a semblance of true representation. Strachan et al. (2013) describe TC simulations as 'increasingly credible' but not complete. Indeed, as hard as it is to effectively model TCs, fundamentally ET remains even harder as it involves many distinctly nonlinear processes (Elsberry and Ritchie, 2007).

Despite a lack of abundant data, a number of studies have emerged developing a methodology to objectively partition a storm's life cycle into a number of distinct phases. Arnott et al. (2004) investigated the structural evolution of transitioning systems in the CPS using a method from computer science and machine learning called K-means clustering. Here the data are grouped into a certain number of 'clusters'. It was determined that seven clusters were required to characterise the distinct structural regimes observed in the range of TC evolution from tropical cyclogenesis to ET. Recent improvements in the resolutions of operational forecast and reanalysis datasets provide a great opportunity to develop this method further and develop results of improved robustness.

As such, the aim of the work presented here is to establish and illustrate a method for objectively determining ET in numerical atmospheric data. The structure of the rest of the paper is as follows: the method of objective ET

detection is described before a case study storm is detailed, the findings of ET in the European Center for Medium-Range Weather Forecasting (ECMWF) operational analysis from 2008 to 2012 is then presented followed by a discussion of the benefits and implications of the technique introduced herein.

2. Methodology

Here we give a brief description of the data used and then an overview of the steps involved in the methodology before describing each individually in detail. We produced storm tracks for all Northern Hemisphere cyclones between 2008 and 2012 from the ECMWF NWP operational analysis. This is a high-resolution dataset produced through the process of data assimilation. The TCs were then identified and selected by matching against the best track record (Hodges and Emerton, 2014). The point in the CPS occupied by the storm tracks at each 6 hourly (i.e. synoptic time) timestep was then determined. As such the path through phase space of each storm was determined. All the CPS points for all storms in the dataset were then clustered according to an unsupervised K-means clustering algorithm. The properties of these clusters, such as number of members and their location in the phase space, were then investigated to determine the cluster, or clusters, that represented ET. These were then used to classify the life-cycle progression of individual storms, thus objectively determining ET.

2.1. Data description

2.1.1. ECMWF operational analyses. The ECMWF operational analysis dataset, stored in their operational archive, is used. Whilst a number of datasets would be appropriate for the demonstration of this methodology, this operational analysis was used owing to its high resolution and as it complemented other work (including Hodges and Emerton, 2014). The data are temporally confined between the 1st of May and 30th of November, 2008 to 2012. These months are chosen as tropical storm activity in the Northern Hemisphere outside this time period is considered negligible. By restricting the study to these years, the number of changes to the operational system is minimised and the period of highest resolution of the system is used. The data are 6 hourly relative vorticity average between 850 and 600 hPa. For the sake of practical extraction and handling, the data are extracted at a T_L255 resolution (512×256) as in Hodges and Emerton (2014).

Changes in the operational analyses must be considered throughout the study as systems are continually under-

going improvement and development, these are summarised here. In 2009 the weak-constraint 4D-Var was implemented for the first time in the assimilation system. In 2010 the underlying model's horizontal resolution was increased from 25 to 16 km and a new cloud parameterisation scheme was introduced. In 2011, a modification of the entrainment and detrainment of convection was implemented and in 2012, the convective downdraft entrainment was modified and de-aliasing of the pressure gradient term was improved. For a fuller description of this dataset see Richardson et al. (2005). The documentation of the ECMWF operational analysis model and a fuller description of the changes is found online (ECMWF, 2014).

2.1.2. Cyclone detection and tracking. It is evident that the accuracy of any analysis of ET is reliant upon the storm tracks used. This work makes use of the now well established method by Hodges (1994, 1995, 1999). This has been used widely in recent studies (e.g. Froude, 2010, 2011; Grise et al., 2013; Strachan et al., 2013; Zappa et al., 2013). There is extensive description of this in the existing literature, for example Froude et al. (2007a, 2007b). The particular configuration of the methodology is described in Hodges and Emerton (2014). As such only the key points and alterations will be presented here.

2.1.2.1. Cyclone detection: The cyclone detection was performed on the relative vorticity field utilising the vertical vorticity average. Serra et al. (2010) demonstrated the use of the vertical vorticity average to capture more of the life cycle of storm, and Hodges and Emerton (2014) showed that this enabled determining tracks significantly longer than the best track records in the International Best Track Archive for Climate Stewardship (IBTrACS) database (Knapp et al., 2010). The data are averaged over the 850, 700, and 600 hPa levels.

To remove noise it is spectrally filtered to T63, spectral coefficients are tapered and large-scale background is removed for total wave numbers equal to five. As such, vorticity maxima are then determined in the resulting data. The method is documented to effectively capture more of storm life cycle than is traditionally tracked both operationally and by storm tracking algorithms (Hodges and Emerton, 2014). This includes early stages as easterly waves and life after transition (e.g. Serra et al., 2010; Hodges and Emerton, 2014).

2.1.2.2. Cyclone tracking: The sequence of vorticity maxima, i.e. the storm tracks, is determined with a nearest neighbour search algorithm (Hodges, 1994; Hoskins and Hodges, 2005). These are then refined using a cost function to determine the smoothest, and thus most probable, set of

tracks. All tracks lasting for less than two days are removed. The resulting storm tracks are then matched against the best track records in the IBTrACS dataset to identify the storms. Storms are declared to be matches if they temporally coincide and have a mean geodetic separation distance of less than 4 degrees. All storms in the best track data are found in the analysis. The benefit of this approach is obtaining more of the storm life cycle than IBTrACS tracks (Hodges and Emerton, 2014). The formative stage of TCs as easterly waves is captured as is the extratropical cyclone stage. Figure 1 presents all of the 2012 storm tracks. A summary of the storms found is detailed in Table 1. Each storm is assigned to a basin in reference to the basin where it achieves its maximum intensity. The basin definitions used are the same as Hodges and Emerton (2014).

2.1.3. Cyclone structure: phase parameters. The position in the CPS (Evans and Hart, 2003; Hart, 2003) is determined for each timestep. The CPS method has become widely used to characterise storm structure (e.g. Evans and Hart, 2003; Arnott et al., 2004; Evans et al., 2006; Hart et al., 2006; Veren et al., 2009). A full description is not repeated here: the interested reader is referred particularly to Evans and Hart (2003).

The CPS uses three thermal parameters to describe the storm system. These are all determined using the geopotential height fields in the analysis dataset. The parameters are a measure of the lower tropospheric thermal symmetry (B) and the lower- and upper-tropospheric thermal wind parameters ($-V_T^L$ and $-V_T^U$, respectively). The lower troposphere is defined as the 925–700 hPa layer and upper troposphere as the 500–300 hPa layer.

In other applications of the CPS the lower troposphere is defined as 900–600 hPa and the upper troposphere as

600–300 hPa. However, since one of the immediate aims of this project is to explore ET in data where these levels are not stored, we use the 925–700 hPa and 500–300 hPa for the lower and upper troposphere, respectively. This choice has been explored and it is believed that it should not substantially affect the results since these levels characterise the required atmospheric layers well enough and they represent a similar pressure difference (i.e. 200 hPa compared to 300 hPa). Lower tropospheric symmetry is an area-averaged difference of the 925–700 hPa layer within a 500-km radius of the storm centre:

$$B = h(\overline{Z_{700hPa} - Z_{925hPa}}|_R - \overline{Z_{700hPa} - Z_{925hPa}}|_L), \quad (1)$$

where Z is the geopotential height, and R and L denote right and left of the track, respectively. This parameter can be taken as a proxy indicating the frontal characteristics of the storm. In symmetrical systems, for example developed TCs, B is close to zero. As a system becomes increasingly asymmetric its value of B increases. A value of around 50 m would be expected of a classic baroclinic cyclone (Veren et al., 2009). The two thermal parameters are defined as vertical derivatives in the height change of isobaric surfaces similarly within a 500-km radius of the storm centre:

$$\frac{\partial(\Delta Z)}{\partial(lnp)}\Big|_{925hPa}^{700hPa} = -\big|V_T^L\big|, \quad (2)$$

$$\frac{\partial(\Delta Z)}{\partial(lnp)}\Big|_{500hPa}^{300hPa} = -\big|V_T^U\big|, \quad (3)$$

where ΔZ is the height change from the highest to the lowest point on a pressure surface within 500 km of the storm centre, and p represents pressure Veren et al. (2009). To compute these quantities the geopotential height field is sampled on a radial grid of radius 5 degrees and centred on the storm locations and rotated to the storm propagation
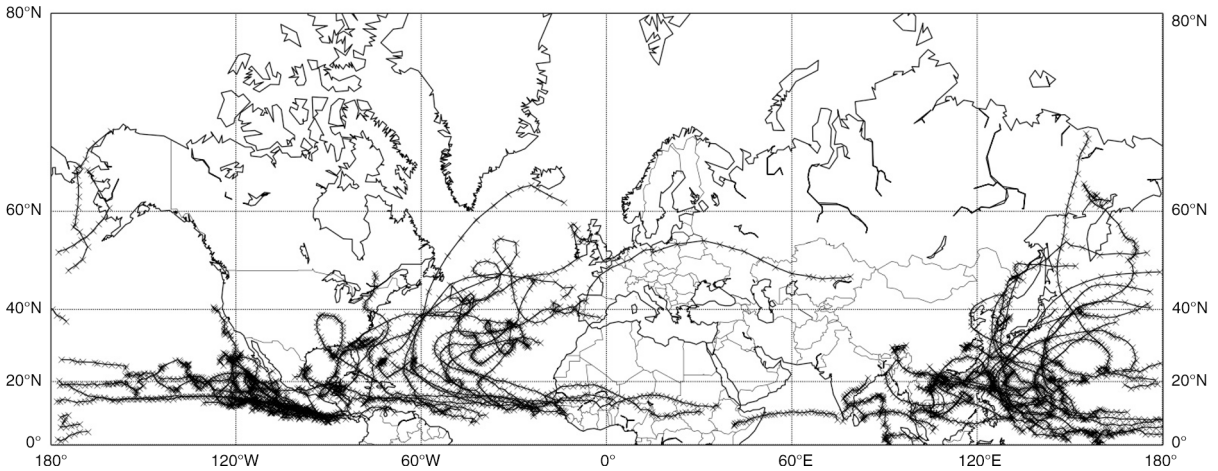

Fig. 1. Map of all storms in the dataset in 2012.

Table 1. All storms tracked in ECMWF operational analysis from 2008 to 2012. All these storms are used in the subsequent analysis

Basin/category	All years	2008	2009	2010	2011	2012
All basins	315	69	67	55	62	62
Atlantic Ocean	72	14	9	19	15	15
E. Pacific Ocean	81	19	20	12	14	16
W. Pacific Ocean	111	25	27	15	20	24
N. Indian Ocean	32	9	5	7	8	3
Tropical storm (<Cat. 1)	182	43	43	31	37	28
Category 1	47	10	8	8	8	13
Category 2	26	5	5	7	5	4
Category 3	27	5	6	3	5	8
Category 4 and 5	32	6	5	6	7	8

direction (see appendix of Bengtsson et al., 2007, for more details).

2.2. Classification and clusterings

By computing the CPS values for each of the timesteps in each of the storm tracks, a point cloud representing the realised occupancy of storms in the CPS is acquired. Doing this provides a map through the CPS with which to compare individual storms in relation to others. We desire to classify these points together to pick out life-cycle phases. To do this we use a cluster analysis (CA).

CA has been used widely in the atmospheric sciences. Recent examples of this include using CA to assess the performance of climate models (Yokoi et al., 2011) and precipitation patterns (Johnson et al., 2011). Arnott et al. (2004) applied this technique to the CPS. The process involves the grouping of similar objects together (Murphy, 2012). 'Similarity' is maximised within each of the groups (clusters) and minimised between the groups. This is judged from their position in phase space. Thus, two points are deemed similar if they are close together, and increasingly dissimilar as this distance increases.

Here a non-hierarchical approach is applied. This means that the data are divided into a number (usually denoted K) of clusters. This number is imposed on the dataset by the analyst. Using cluster validity functions, Arnott et al. (2004) find that seven clusters are required to accurately fit the life-cycle stages of storms and this result has been used in subsequent work (Evans et al., 2006; Veren et al., 2009). It is important however to stress that any meaning behind the clustering is a subjective imposition by the analyst. As such while seven is identified as being the *most adequate*, no number of clusters can be objectively chosen as being the *most realistic*.

The K-means clustering is done by grouping points together with the smallest distance in the CPS. This distance is defined by their Euclidean separation (Arnott et al., 2004):

$$d_{i,j} = [B_j - B_i]^2 + [(-V_T^L)_j - (-V_T^L)_i]^2 + [(-V_T^U)_j - (-V_T^U)_i]^2. \tag{4}$$

The implementation of K-means used was that of Pedregosa et al. (2011). The clustering procedure is started by picking out seven points randomly to initialise the clusters. The distance between each point in the dataset and the mean point of each cluster is then compared. The point is assigned to the cluster to which it is closest to.

This process (iterating over all the points in the dataset and comparing their distance to the mean cluster points, reassigning their cluster accordingly) is repeated until an equilibrium state is reached. This state is characterised by some 'inertia', the number of points that retain the same cluster classification upon each iteration. It is extremely rare for the inertial number to equal the number of points in the dataset since data rarely coheres to cleanly-divided groups. The random choice of the initial clusters effects the results and as such the algorithm is repeated 1000 times, each with different randomly chosen initial, singleton clusters. The final result is the clustering with the highest inertia between seed clusters. It is necessary to account for the differences in the range in variations in the CPS parameters so the following normalisation is applied to the data before eq. (4) is computed (Arnott et al., 2004):

$$X_{i,norm} = \frac{X_i - \bar{X}}{\sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2}} \tag{5}$$

Figure 2 shows all 17 754 points in the CPS coloured by their assigned cluster. This cloud of points, containing all storms in the dataset in the phase space, represents possible phase occupancy of storms over the course of their life cycles. Table 2 shows the corresponding statistics for each cluster. The light green points loosely correspond to the midpoint of traditionally understood ET as in this cluster the system has become asymmetric but is not yet a cold-core system (Jones et al., 2003). However, it is of utmost importance to note that these clusters are determined by geometry alone and any meaning assigned to them is imposed by the observer. As such, the boundaries of any such clustering should be read with this in mind. Figure 3 shows the mean statistics of each cluster. Here the size of the circle is proportional to the number of points in the cluster. This is a useful diagram through which to understand storm life cycles. TC, of course, spend most time as
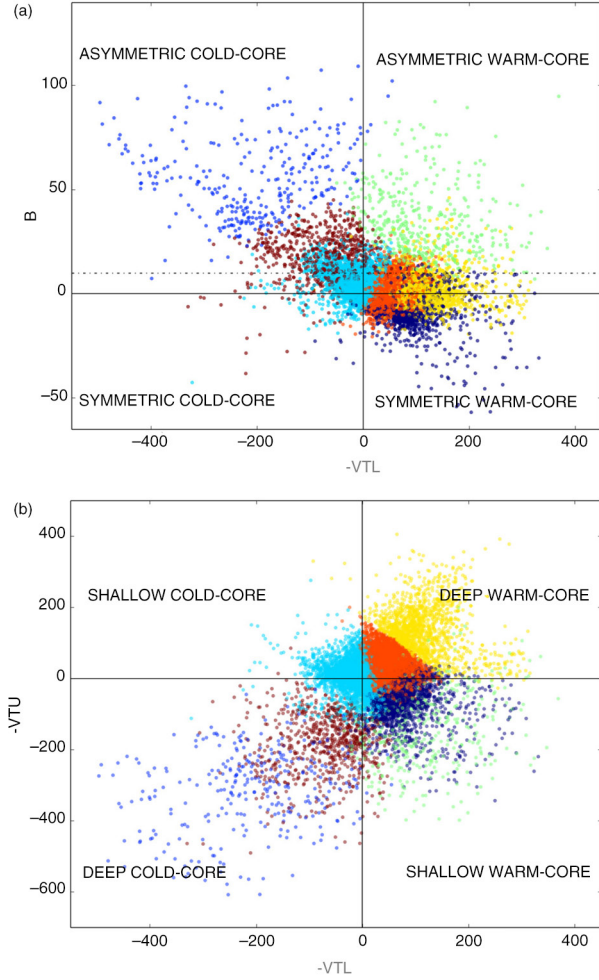
Fig. 2.   Scatter plots of the K-means clustering (seven clusters) for all data points in the CPS. (a) Symmetry parameter against lower tropospheric thermal parameter. The dash line at B = 10 m is the point marking the onset of ET defined by Hart (2003). (b) Upper against lower troposphere thermal CPS parameters.
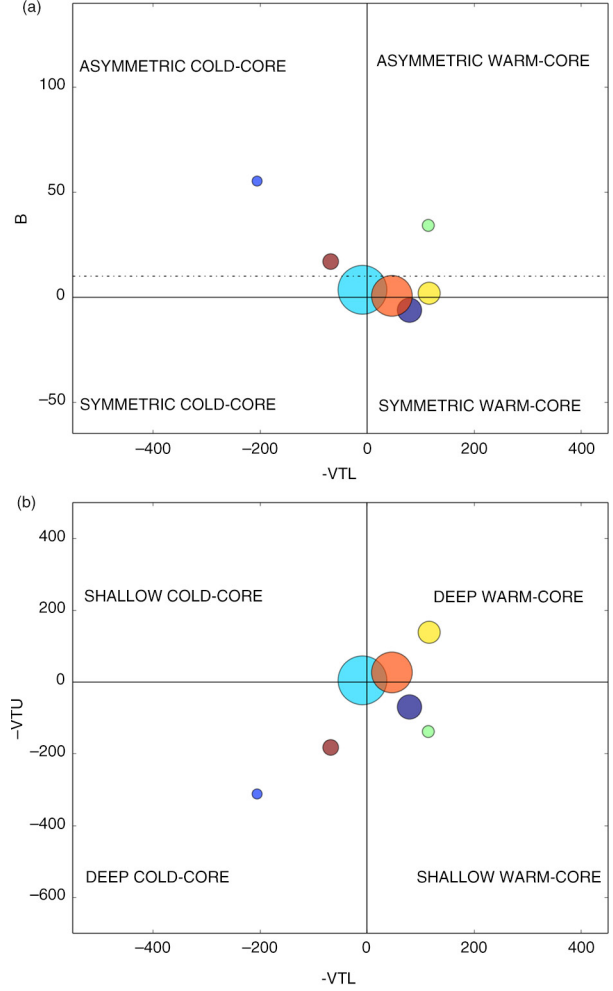

Fig. 3.   Mean statistics of the K-means clustering (seven clusters) for all data points in the CPS (Hart, 2003). (a) and (b) as above. Each point corresponds to a cluster and represented with its size being proportional to the number of points in the cluster. The statistics for these clusters is shown in Table 2.

symmetric warm-core systems. The range in cluster size is 8134 points. This variation demonstrates that all storms do not experience the same, or even similar, life cycles.

Table 2.   Cluster characteristics

Cluster number	Colour	Mean B	$Mean - V_T^L$	$Mean - V_T^U$	Number of points
1	Navy blue	−5.0	83.5	−52.5	2302
2	Red	15.8	−70.1	−191.5	863
3	Turquoise	0.6	19.0	7.6	8528
4	Yellow	1.9	97.9	117.7	2231
5	Blue	55.0	−204.9	−307.6	394
6	Light green	33.5	113.6	−136.6	480
7	Orange	8.0	−28.2	13.4	2956

2.3. Cluster characteristics

With this CPS clustering classification having been done, it is then possible to overlay a new storm onto this 'map' of grouped points to determine where it is in its life cycle at each timestep. In fact, separate clusterings are determined for each basin as storm characteristics vary considerably depending on their ocean basin. In this way, ET can be objectively classified according to recorded storm tracks in individual basins. In order to implement this method effectively various additional considerations and constraints need to be applied to the ET determination algorithm (shown in pseudocode in Algorithm 1). The first of these considerations distinguishes between the regional meteorological variations.
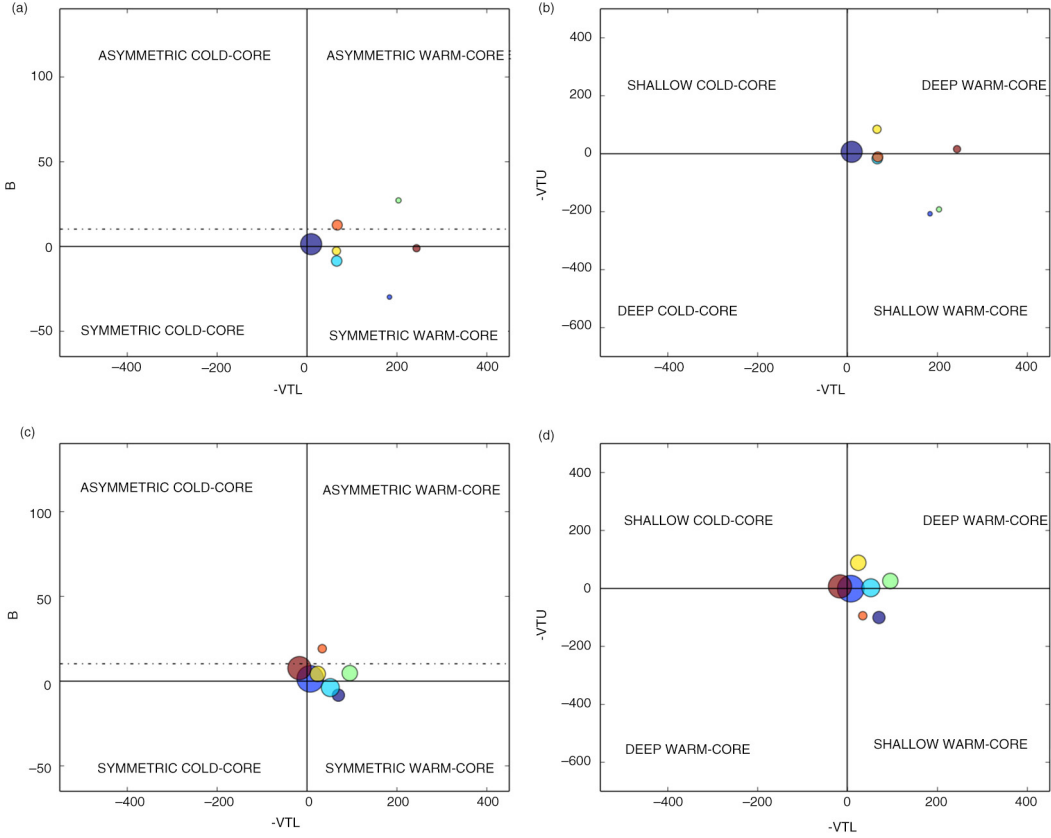
Fig. 4. Mean life-cycle stages for storm systems in (a) and (b) the northern Indian Ocean and (c) and (d) the eastern North Pacific. The size of the points is proportional to the number of points in the cluster. Note the same set of colours is used as in Fig. 4 but the same colour in different CPS diagrams do not correspond.

The set of clusters produced for each basin varies as a result of heterogeneous geographical and climatic conditions. These differences are represented clearly in the CPS. The western North Pacific and the North Atlantic basins' clusters closely resemble those of whole Northern Hemisphere (shown in Fig. 2). This is because the majority of Northern Hemisphere TCs occur in these basins and have the largest number of re-curving storms into higher latitudes. The characteristics for the eastern North Pacific and the northern Indian however are considerably different. The CPS diagrams for these basins are shown in Fig. 4.

In the northern Indian Ocean no cluster is assigned to an asymmetric cold-core system. This is because any system attempting transition will inevitably make landfall over the Indian continent and lose its energy source. While this lack of energy prevents any storm from surviving transition it does not preclude ET from beginning. Indeed storms that have undergone this process are identified in the analysis and assigned a cluster in the asymmetric warm-core quadrant in Fig. 4a. This cluster has a $-V_T^L$ around 200 implying that storms undergoing this structural transition do not survive very long although the $-V_T^U$ around -200

implies that they have completely lost their TC characteristics. At the end of their life cycles these systems are very shallow, asymmetric warm-core storms. In the CPS therefore they resemble monsoon depressions, a connection that could be explored in further work.

In the eastern North Pacific the clusters again are different from the mean Northern Hemisphere types dominated by the western North Pacific and the North Atlantic (Fig. 4). The general characteristics of the whole Northern Hemisphere are recognisable; however, they appear to be suppressed. Storms in this basin appear to not explore as much of the CPS as in other basins. There is a shallow, asymmetric warm-core cluster (orange) corresponding to a midpoint of ET. It however has fairly low mean B and $-V_T^L$ values. An interesting cluster is the one that approaches the deep, asymmetric cold-core (red). It appears as though there are not enough points in the second quadrant (i.e. corresponding to extratropical cyclones) for them to be assigned their own cluster. Eastern North Pacific storms tend to propagate zonally and in general do not recurve.

Despite these regional variations, one can assume there are clusters in each basin through which a storm must pass

during transition. It is these clusters that are desired to be identified by an algorithm to determine ET. As a result of these geographical variations, the corresponding basin's clustering is used when determining transition for individual storms rather than that of the whole hemisphere. The transition clusters must be subjected to a number of additional considerations before the time period of ET is finally identified.

2.4. Determination of ET

The canonical understanding of transition in the CPS is that a storm loses its symmetry faster than it becomes a cold-core system. Theoretically two alternate routes exist to transit from across CPS; namely developing a cold-core faster than losing symmetry and both aspects occurring simultaneously. In development of this methodology it became clear that not only are all three types theoretically possible, but all three occur within the ECMWF dataset. These three types of ET are listed in Table 3 along with their relative frequencies in the dataset. It is surprising to find that type two transitions, transitions via the symmetric cold-core route, are more frequent than the canonical type one. Since the annual range of the dataset is only 5 yr long, these frequencies may not be representative of long-term transition characteristics. Figure 5 shows composites of the transition types (identified using the approach described subsequently). It is not clear how these three types correspond with the three types described by Kitabatake (2011) and further analysis is required to reconcile these idealistic descriptions.

In light of this finding, any objective algorithm to detect ET must account for the fact that it can occur along multiple pathways. In the instance of the K-means clustering approach followed here, this requires assuming that three of the seven clusters represent transitory life-cycle stages. This is achieved by defining the three-dimensional space corresponding to each type of transition, and then determining which cluster has the most points in that space. All three are then considered in turn (see Algorithm 1 for pseudocode). All points of the storm's life cycle in these three clusters with asymmetry parameter B less than 8 and $-V_T^L$ greater than zero were removed from consideration

Table 3. Types of extratropical transition through the Cyclone Phase Space

Type	Characteristics	Relative frequency (absolute count)
One	Via asymmetric warm-core	35% (648)
Two	Via symmetric cold-core	45% (846)
Three	Directly to asymmetric cold-core	20% (370)

as potentially representing ET. This is because, despite these points being warm-core, symmetric systems, they are sometimes clustered into the potential ET clusters. Then all the points are grouped into chronological sequences. These sequences are potentially all periods of ET, the first point in the sequence represents the start of transition and the last point completion.

It is clear that a storm can experience multiple transitions during its life cycle as it moves around the phase space, and this is accounted for here. For each of these sequences the moment the asymmetry parameter (B) decreased or the lower troposphere thermal parameter ($-V_T^L$) increased, transition is recorded as having completed. This is because these changes represent a system moving back towards TC-like characteristics and thus imply ET is complete. Next, if the resulting set is a singleton group of points, the points immediately before and after in the life cycle are looked at. There is considerable 'noise' in the data where the storm's position in the CPS will flicker around point to point, varying considerably from the overall path. Thus, a number of points appear in the transition cluster that do not in all probability represent transition in any real meteorological sense. These, at this stage appearing as singleton potential ET sequences, are thus ignored. Finally, if a storm does not move at least 10 degrees in latitude over its whole life cycle, no transition is recorded even if one 'appears' in the CPS CA because this is not reconcilable with understandings of ET as the process of storms of tropical genesis moving into the midlatitudes. By implementing these constraints results were considerably improved, for example removing the higher latitude baroclinically induced TCs (McTaggart-Cowan et al., 2013) and monsoonal storms in the northern Indian Ocean that superficially resemble ET in the CPS.

3. Results

In applying this method for objectively determining ET, the transition characteristics of the storm tracks from 2008 to 2012 were produced. The results are presented in three parts. First, example storms for each type of ET are detailed. Secondly, the absolute number of storms and transitions is presented for both the Northern Hemisphere and basin-by-basin scale. Thirdly, the seasonal distribution of transition is explored. Although the data only span a period of 5 yr (a sixth of the normal timeframe understood to produce representative climatologies), the monthly averages are produced to obtain results that can be compared to existing climatologies. At the outset therefore it should be stated that these comparisons are done to first order at best.
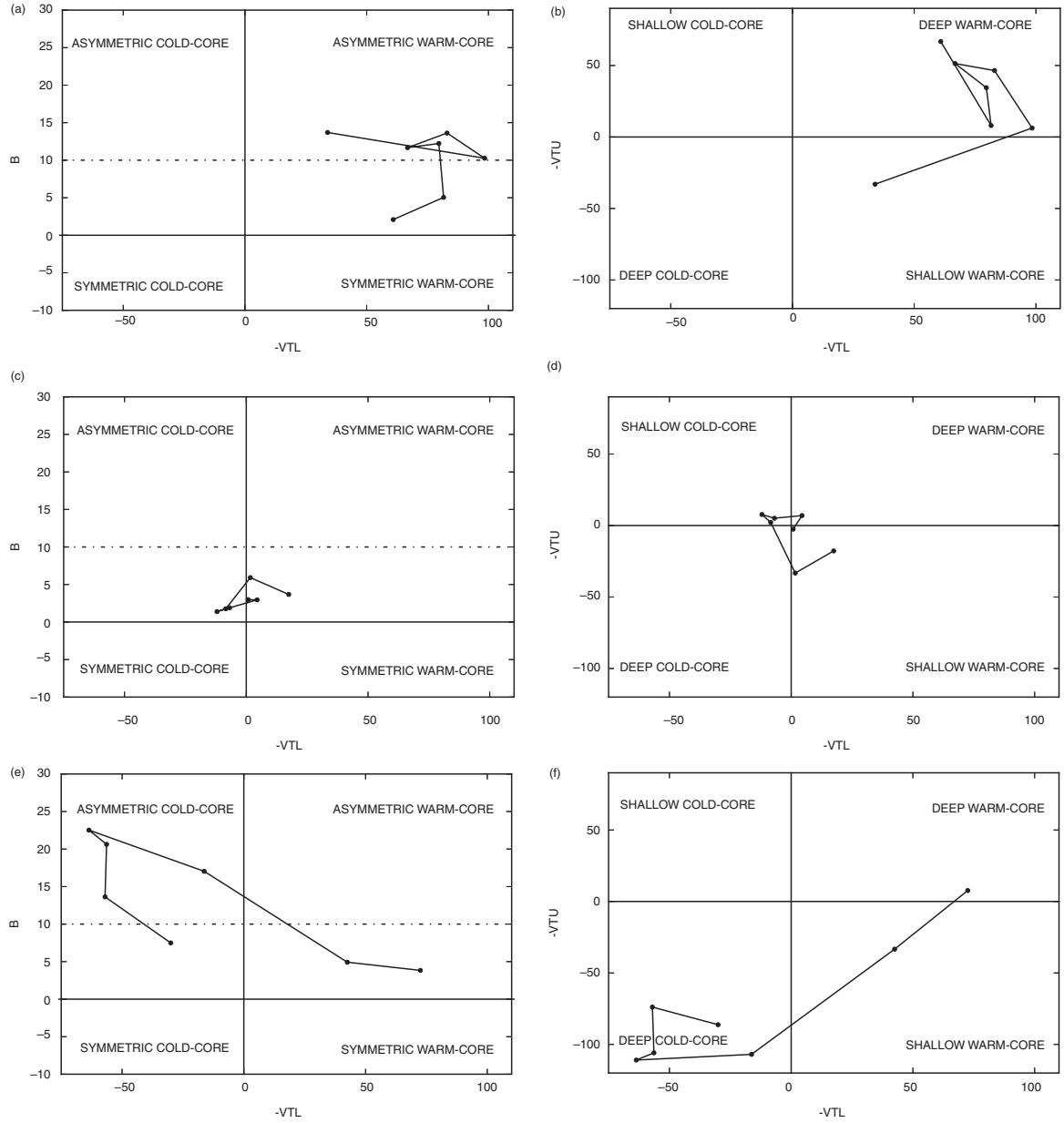
Fig. 5. Composite of extratropical transition for the Northern Hemisphere through the different pathways in the CPS. (a, b) Type one: asymmetric warm-core. (c, d) Type two: symmetric cold-core. (e, f) Type three: directly to asymmetric cold-core.

3.1. ET1: Hurricane Leslie, 2012

North Atlantic storm Leslie, which occurred in late August and September 2012, is presented as an example storm. It is selected as a typical result of ET determination by this method. The track is shown in Fig. 6. It was a long-lived TC and intensified to become a category 1 hurricane (by the Saffir-Simpson wind scale) for a short period. In the latter part of its life cycle it underwent ET to become a strong extratropical storm making landfall along the Burin Peninsula. A lower category storm is selected as a case study, rather than a category 4 or 5, because storms of this type represent the majority of TCs.

The American National Hurricane Centre (NHC) report Leslie as forming from a strong tropical wave moving off the west coast of Africa at the end of the 26th August (Stewart, 2013). The storm then crossed the Atlantic slowly intensifying. By 12:00 UTC 5th September it was 780 km south-east off the coast of Bermuda. It then expanded horizontally and strengthened to become a category 1 hurricane. In the
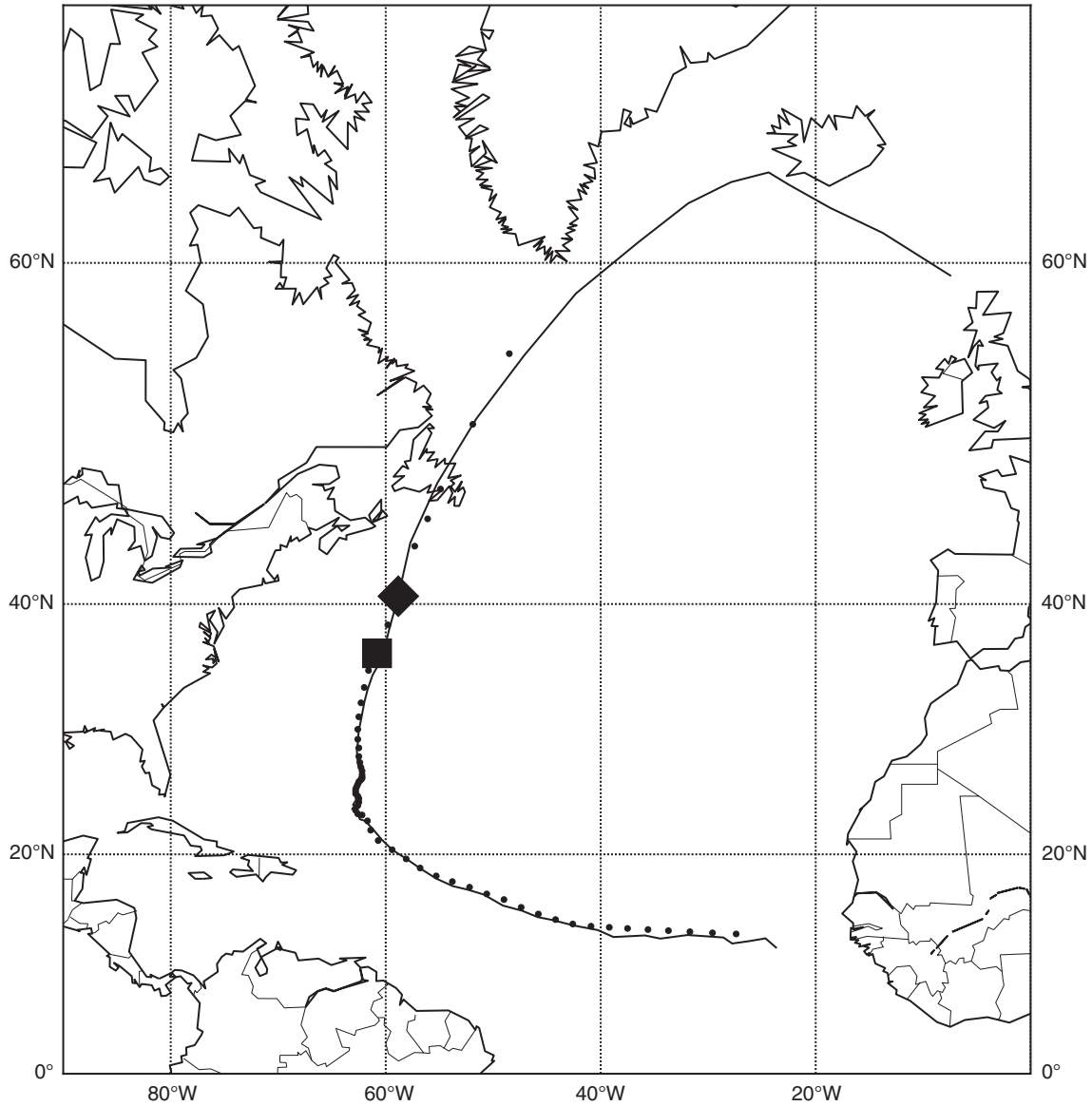
Fig. 6.   The track of North Atlantic Hurricane Leslie in 2012. Solid black line is this analysis's storms track. Dots are the best track path of the storm (Stewart, 2013). The square is the location of the start of ET as declared in the best track archive and the diamond is the start of ET as determined by this methodology

following days, it proceeded northwards very slowly but continued to have a very large radius around 280 km, over twice the average radius for a typical category 1 hurricane. The NHC report Leslie as merging with a cold front in the early morning of 11th September to form a powerful extratropical cyclone. At this point it was located 140 km southwest of St. Lawrence, Newfoundland, Canada.

The tracks shown in Fig. 6 demonstrate the benefit of using the tracking regime applied here. The solid black line representing the track used in the analysis identifies the storm earlier than the best track archive and follows it for much longer (for other storms this can be even more pronounced).

Hurricane Leslie continued to pass south of Greenland, it then skirted the south-west coast of Iceland and then continued down towards the north of the British Isles.

The start of transition of Hurricane Leslie is marked by the solid square and diamond. There is some discrepancy between the determination of ET start time. The solid black square shows the time at which ET was said to have started by the NHC report (Stewart, 2013). This is 09:00 UTC 10 September. This report records ET to be complete 24 hours later at 09:00 UTC in the morning of 11 September. We find ET to start at 00:00 UTC on the morning of the 11th and to be completed rapidly 6 hours later. The path of the
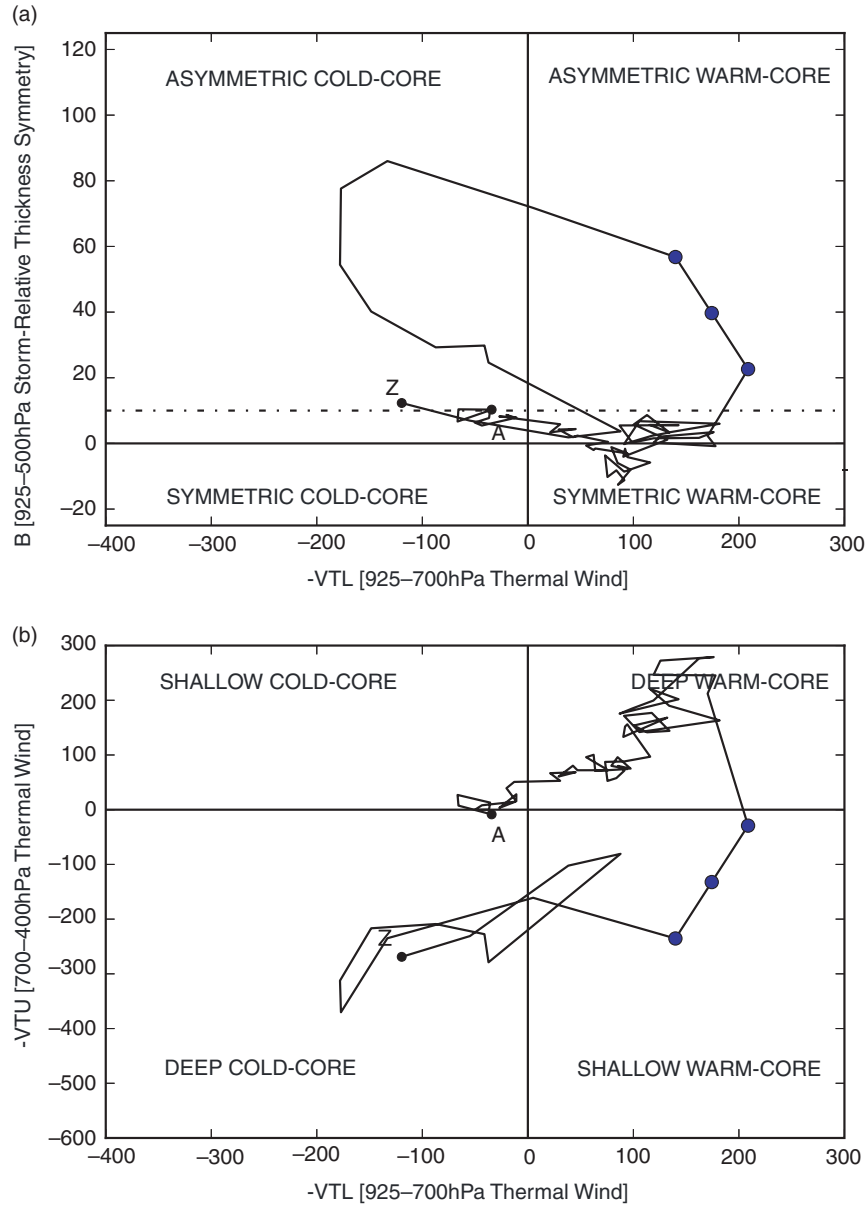
Fig. 7. The path of Hurricane Leslie through the CPS and ET determination. Points show period of extratropical transition. The green triangle marks the start of the track, and the red triangle the end point. (a) The CPS symmetry parameter against the lower tropospheric thermal parameter. (b) The upper and lower tropospheric thermal parameters of the CPS.

storm through the CPS is shown in Fig. 7. This transition represents a traditional type 1 ET as the storm becomes an asymmetric warm-core system before becoming a cold-core storm. The points show the steps through transition, the start (with a B value around 20), the mean point (B around 40) and completion ($B = 60$).

3.2. ET2: Tropical Storm Aere, 2011

Tropical storm Aere occurred in early May 2011 in the western North Pacific. It formed in the Philippine Sea on

the 5th and made landfall over Luzon before recurving and transitioning off the coast of Japan (Fig. 8). We find ET to have occurred at 18:00 UTC 11 May. It underwent ET on the 12th according to Angove and Falvey (2011). Yet the methodology presented here finds that transition occurred somewhat earlier.

Aere is illustrative of a 'type two' transition, which is the reverse of traditional ET. Instead of becoming an asymmetric warm-core system, it becomes cold-core faster than it becomes asymmetric (Fig. 9). This means that it does not really penetrate into the asymmetric
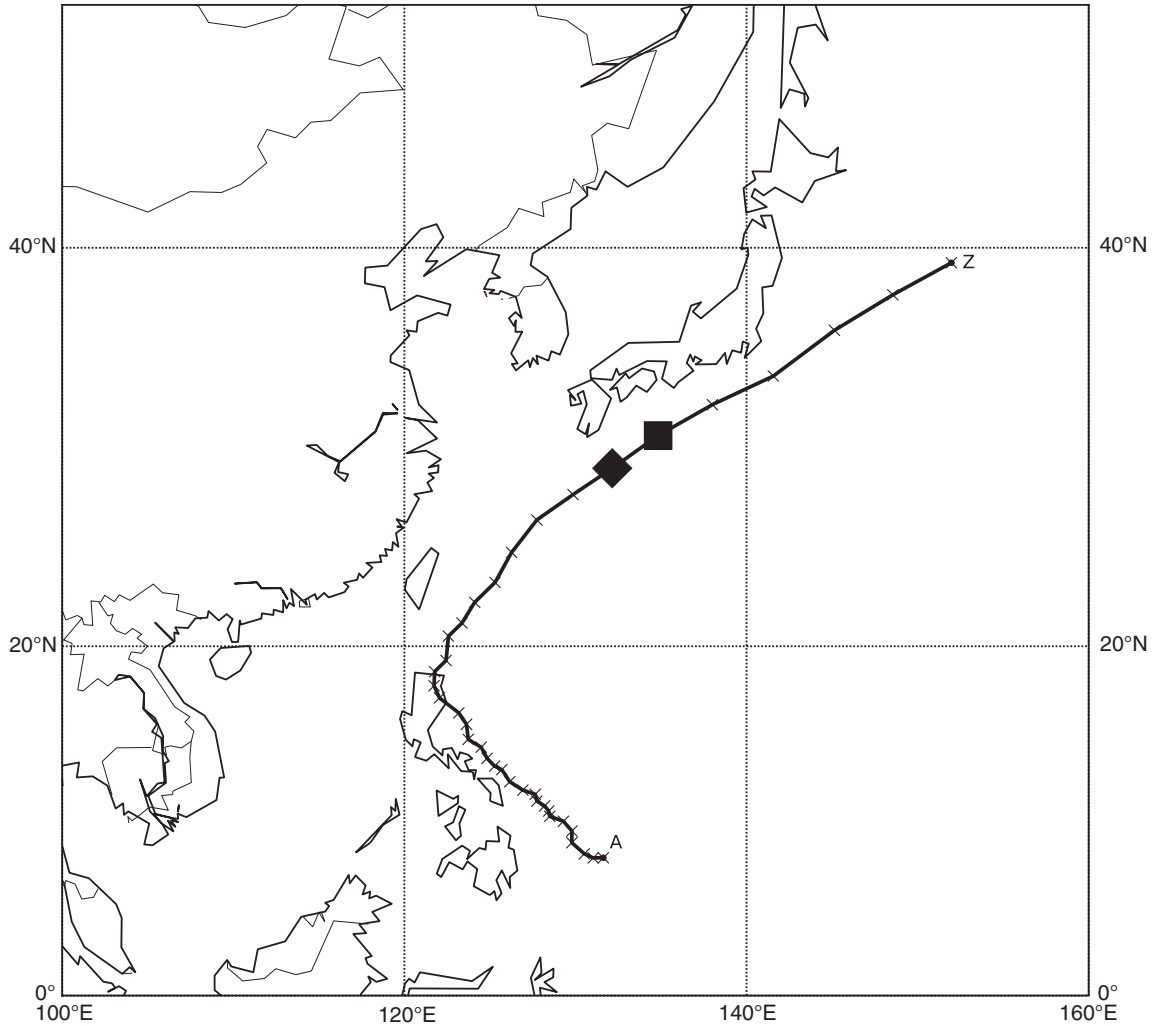
Fig. 8.   The track of the western North Pacific tropical storm Aere (May 2011). The annotations A and Z mark the start and end of the track, respectively. The crosses that divide the track are the timesteps (every 6 hours) of the operational analysis. The square is the location of the start of ET as declared in Angove and Falvey (2011) and the diamond is the start of ET as determined by this methodology.

warm-core quadrant of the CPS diagram (top-right in Fig. 9a).

3.3. ET3: Severe Tropical Storm Kirogi, 2012

Severe tropical storm Kirogi occurred in the western North Pacific in early August 2012. It developed in the Pacific, then rapidly headed North into the Sea of Okhotsk (Fig. 10). It formed on 3 August 2012 and dissipated on 10 August. We find ET to occur at 00:00 UTC 10 August. The JTWC report ET to occur on the 6 August (Evans and Falvey, 2013).

During transition this storm becomes asymmetric at a comparable rate as it becomes a cold-core system. This is evident in the storm's CPS diagram (Fig. 11). For this reason it is referred to as a type three transition.

3.4. Number of transitions

We find that in the whole Northern Hemisphere basin, out of an average number of storms per year between 2008 and 2012 of 59, 32 (54% $\pm$ 7) undergo transition (Fig. 12). The most storms occurred in 2008 (67), 30 (45%) of which are found to have undergone transition. This is also the lowest percentage transition of all 5 yr. There were the fewest storms in 2010 (53), 30 of which (57%) transitioned. The high percentage transition occurred in 2012 (64%), with 37 of 58 undergoing ET.

Figure 13 shows these annual totals by basin. In the North Atlantic, the year 2010 contained the most storms (19), of which 13 (56%) underwent ET. This was the lowest rate of transition here of all the 5 yr in the analysis in this basin. The highest rate of transition occurred in 2012 when
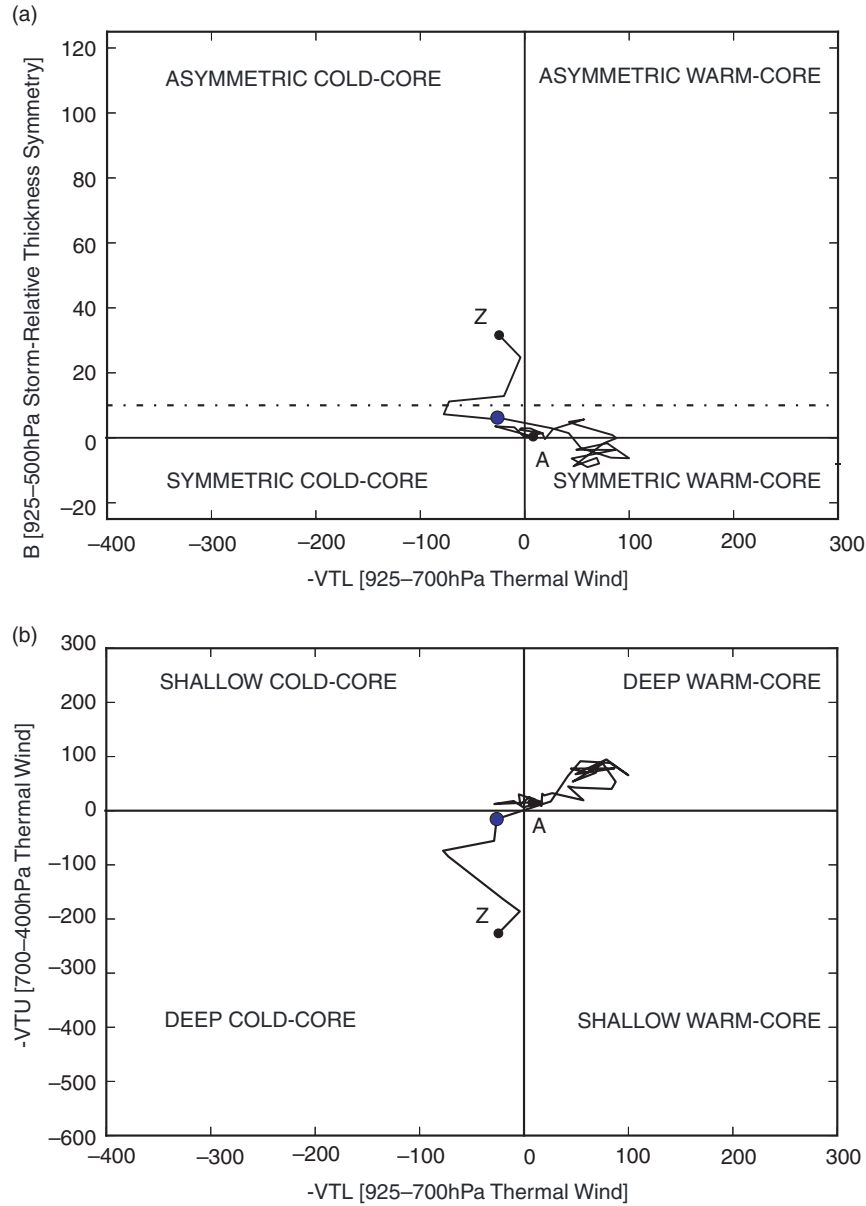
Fig. 9. The Cyclone Phase Space diagram for the tropical storm Aere, 2011. Point shows period of extratropical transition. The annotations A and Z mark the start and end of the track, respectively. (a) The CPS symmetry parameter against the lower tropospheric thermal parameter. (b) The upper and lower tropospheric thermal parameters of the CPS.

13 of 15 (87%) underwent ET. The lowest number of storms occurred in 2009 which was an El Niño year. It appears that this did not alter the proportion of storms undergoing ET.

In the western North Pacific most storms occurred in 2009 (27), of which 15 (56%) underwent ET. The year with the most transitions was 2012 when 71% (17 of 24 storms) are found to have undergone ET. In 2010, the year of fewest storms, 14 of 20 storms were found with ET events (67%). Over the time period, an average of 65% (±6) storms undergo transition every year.

The northern Indian Ocean is found to have less activity than the North Atlantic and the western Pacific. The eastern North Pacific is comparable to the North Atlantic in terms of number of storms but has few transitions. Most eastern North Pacific storms occurred in 2009 (20), of which nine (45%) underwent ET. This was also the year with the most transitions. In 2008 only three of 19 storms (15%) are found with an associated ET. This is the lowest rate of transition for the basin. The year of most storms in the northern Indian was 2008 (9) of which only two transitioned (22%). This is the lower percentage
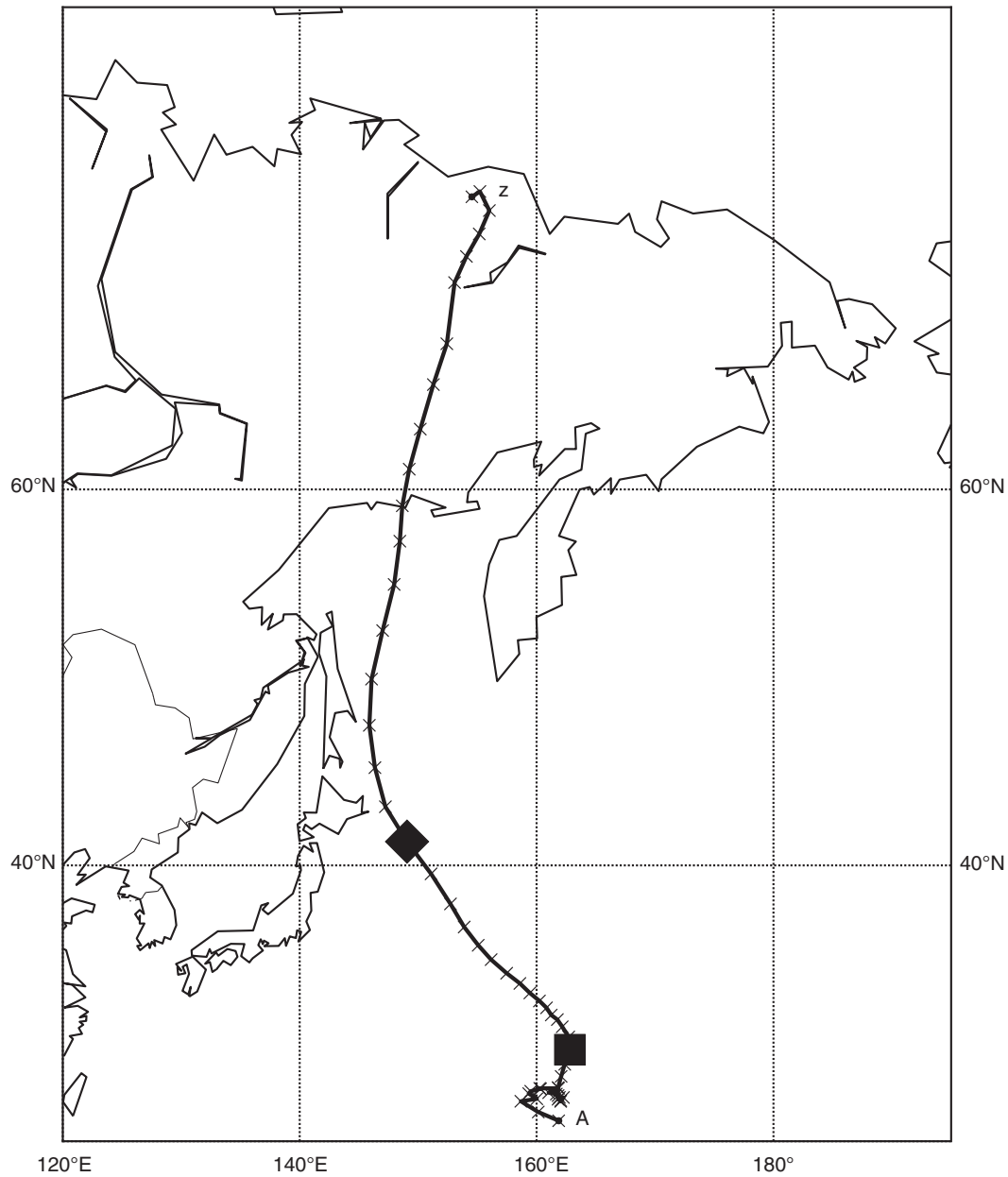
Fig. 10. The track of western North Pacific severe tropical storm Kirogi, August 2012. The square is the location of the start of ET as declared by the JTWC and the diamond is the start of ET as determined by this methodology.

transition over all years. The highest rate of transition was in 2009, three of five underwent ET (66.6%). In 2012, of three storms none attempted transition according to this methodology.

This analysis finds the basins with the smallest number of storm events have the most variability in transition rates, in the northern Indian Ocean this ranges between 0 and 60%. It is not clear how statistically significant this finding is. Average rates of transition over the period divide the four basins into two distinct groups, the North Atlantic and the western North Pacific have high average rates of

67% ($\pm$12) and 65% ($\pm$6), respectively whilst the eastern North Pacific and the northern Indian have 35% ($\pm$12) and 30% ($\pm$23), respectively.

3.5. Seasonal distribution of transition

Figure 14 shows monthly averages of all years. August is the month with most storm activity, with 14 storms on average occurring over the timespan. There is markedly different behaviour between the absolute number of storms and the rates of transition. The latter is characterised by

(a)

(b)
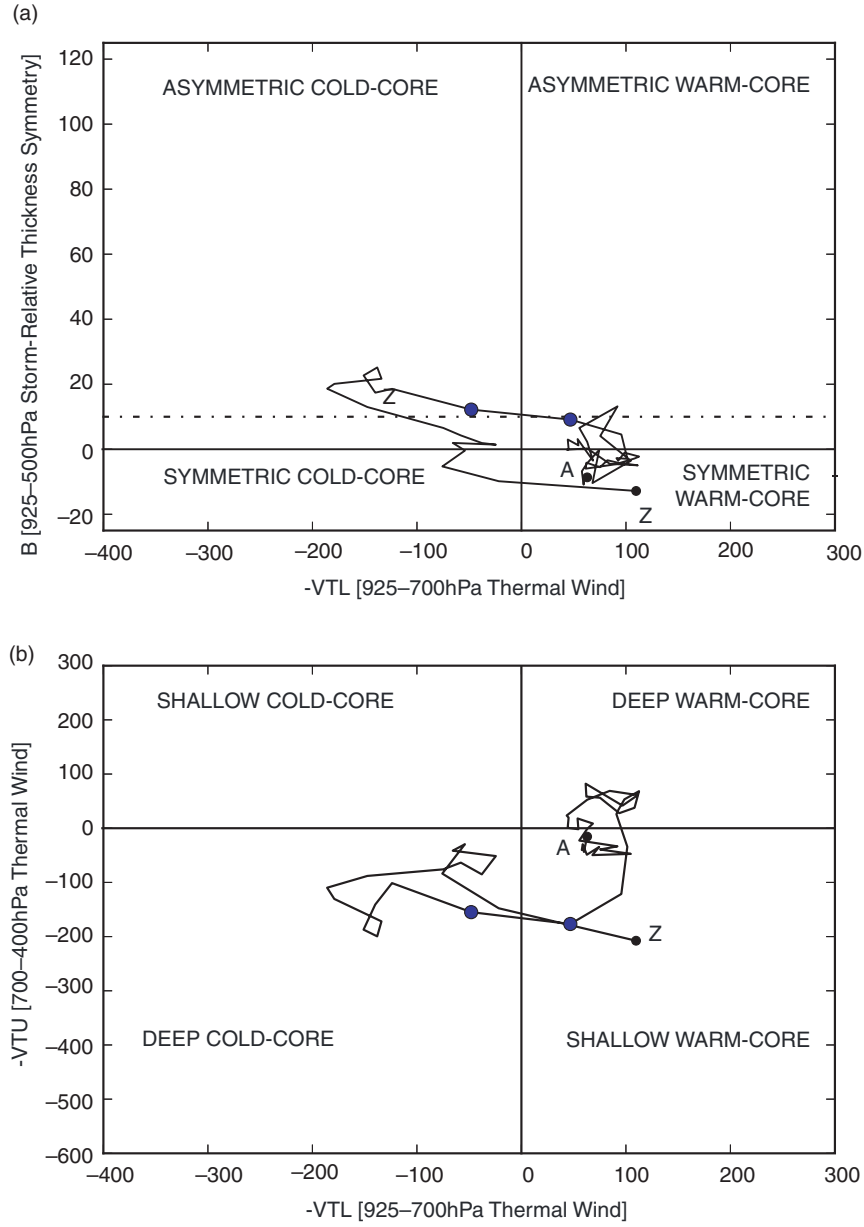
Fig. 11. The Cyclone Phase Space diagram for the severe tropical storm Kirogi, 2012. Points shows period of extratropical transition. The annotations A and Z mark the start and end of the track, respectively. (a) The CPS symmetry parameter against the lower tropospheric thermal parameter. (b) The upper and lower tropospheric thermal parameters of the CPS.

two peaks: one early on in the season in May and then again in September. The analysis finds no transitions occurring in November. The absolute number of transitions remains almost constant throughout August, September and October at eight.

Separating this information out by basin demonstrates distinctly different behaviour, this is shown in Fig. 15. In the North Atlantic most transitions occur in August and September, but highest percentage transition in July. Percentage transition rates seem to be approximately the

same shape as the distribution of the number of storms across the season. In the western North Pacific, similarly the most storms occur in August and September. However, the highest rate of transition is associated with the start of the season, May and June. In June, for example, 90% are determined to have undergone transition. There is then a drop before the rate of transition rises again in late September, early October.

In the northern Indian the highest transition rates occur during the time of peak storm activity (at the beginning and
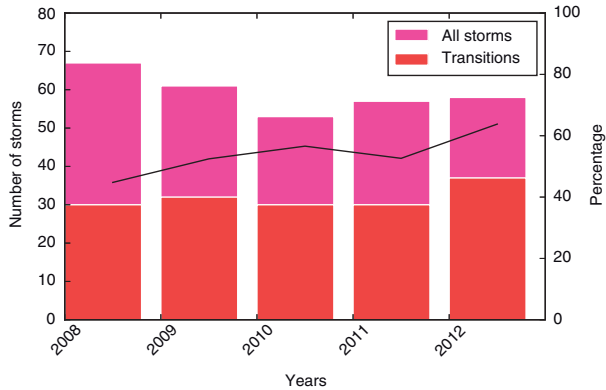
Fig. 12.   Annual number of storms and determined transitions between 2008 and 2012. The lighter colour shows all storms and the darker only those that attempted extratropical transition. Percentage transition is shown with the black line on the second y-axis.

end of the monsoon season). In contrast, the eastern North Pacific maintains its low rate of transition remarkably constant throughout the period. The absolute number of storms is distributed in unimodal fashion with the clear peak of four storms in August.

4. Discussion

The estimated rates and seasonal distributions of ET are discussed first. This is done predominantly with comparison to existing estimates and climatologies. The CPS conditions for the onset and completion time, referred to as the Hart parameters, of ET are then considered.

Subsequently, the associated errors in the determination of ET are considered and the mean transition pathways through the CPS are detailed.

4.1. Transition rates

This analysis finds that 68% (49 of 72) of North Atlantic hurricanes underwent transition between 2008 and 2012. This is considerably higher than the estimate reported by Hart and Evans (2001) who found that 46% (213 of 463) of storms of the same basin transitioned between 1950 and 1996. The monthly distribution of transition is comparable though, with Hart and Evans (2001) also finding transition rates to peak in August and September. Most transitions are found to occur at 23 degrees latitude in the North Atlantic and this seems to qualitatively correspond with previous findings (e.g. Jones et al., 2003). The methodology and dataset of Hart and Evans (2001) is fundamentally different from those used here. They use the best track data from the NHC based on reconnaissance measurements of pressure and wind, and satellite-based estimates of intensity. Typically when a storm moves above 40 degrees North or it is clear that it will not make landfall aircraft reconnaissance typically stops. Even so, this is not always the case: some individual storms are followed to higher latitudes; however, this is not particularly consistent between different ocean basins. Thus, best track data can be sparse for times when a system is most likely to undergo ET. To account for this, Hart and Evans (2001) use the European ECMWF (ERA-15) reanalysis to fill in the gaps. The resolution of those
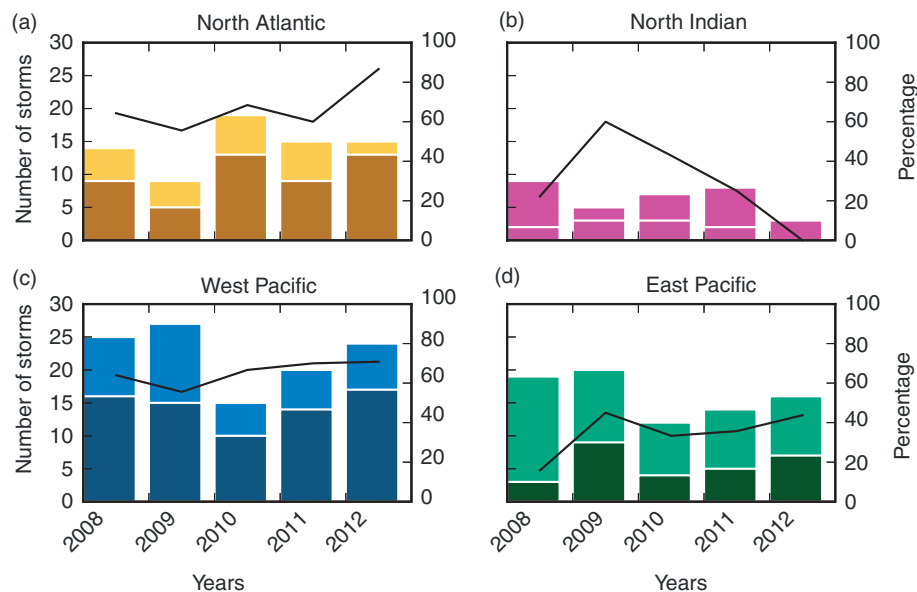

Fig. 13.   Annual storm and transition totals between 2008 and 2012 by basin. (a) The North Atlantic, (b) the northern Indian, (c) the western North Pacific and (d) the eastern North Pacific. Lighter colours show all storms whilst darker colours only transitions. Percentage transitions are shown on the second y-axis with the solid black lines.
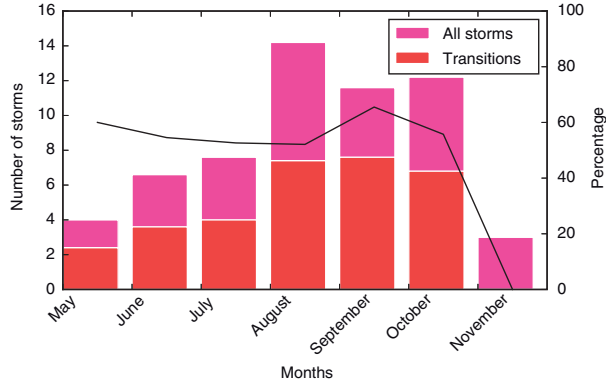
Fig. 14. Monthly distributions of storms and ET in the Northern Hemisphere. This is determined as the sum of the average of monthly totals across all basins in the years 2008–2012. The black line shows the percentage transition on the second y-axis.

analyses is 2.5 degrees and this has been shown to inadequately resolve the processes involved in ET.

We can therefore make an informed speculation that the primary limitation of Hart and Evans (2001) estimation is the shorter tracks. Across all ocean basins, the Best Tracks do have some indication of ET in some basins. However, since the best track dataset is a compilation from various sources under a range of schemes, this suffers subjectivity and differs from basin to basin. They are also designed specifically for tracking the tropical period of storm's life cycles. Using longer tracks, determined uniformly across basins, allows more time for ET to occur and objective determination.

We find 65% (72 of 111) of western North Pacific storms transition at a mean latitude of 25 degrees. Elsberry et al. (2000) find that 27% (30 of 112) of all storms in this basin transitioned between 1994 and 1998. There is agreement between Elsberry et al. (2000) and the results in the distribution of transition rates across the season. The peak rates of transition occur around September and October in both sets of results. The results here also agree with Song et al. (2011), in noting that the onset of ET in the western North Pacific is mostly characterised by $-|V_T^U|$ going from positive to negative. In other words, ET in the western North Pacific is predominately of type two as defined above. Kitabatake (2011) estimates that 16.8% of ET (45 of 268) in this basin takes the cold-core, type two transition route. Additionally, Kitabatake (2011) finds 40% of western North Pacific storms transitioned between 1979 and 2004, with a peak in September–October. Both the Elsberry et al. (2000) and Kitabatake (2011) methodology and data are similarly different from those applied here. The data used is from the Joint Typhoon Warning Center (JTMC) best track analyses. This is subject to similar limitations associated with the NHC best track data: namely the tracks are typically too short to adequately contain the transition period. The determination for ET is based on forecaster interpretation of satellite imagery. Thus, there are also issues associated with both the Elsberry et al. (2000) and Kitabatake (2011) climatologies.

The estimated average transition rate of 35% (28 of 81) in the eastern North Pacific is reassuring as it corresponds to experience. Storms in this basin generally propagate

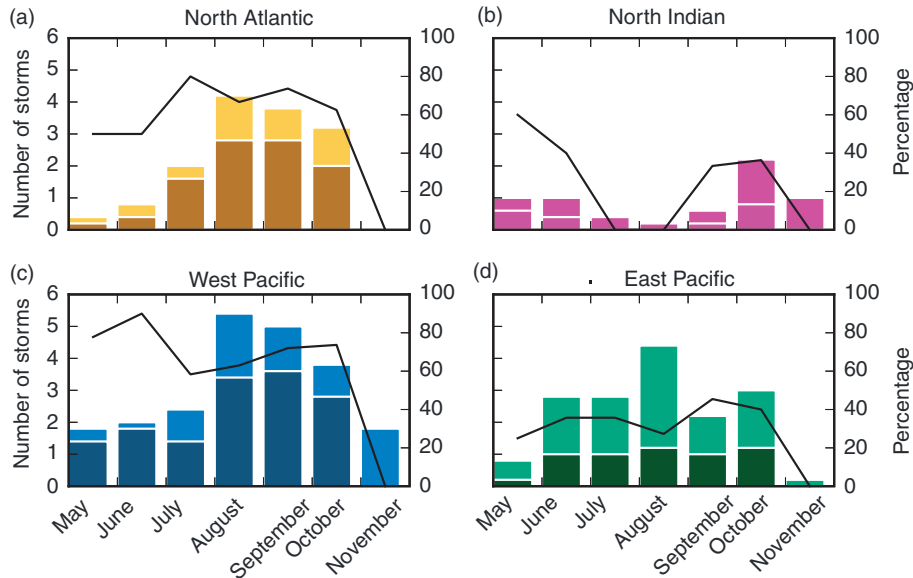*Fig. 15.* Storm and transition monthly distributions across each of the four basins. (a) The North Atlantic, (b) the northern Indian, (c) the western North Pacific and (d) the eastern North Pacific. Darker colours show the number of storms that underwent ET compared to all the storms in the basin (lighter colour). Percentage transitions rates are shown on the second y-axis as the solid black line.

zonally since the steering flow is predominantly from the East to the West and thus transition is less common than in the western side of the Pacific and in the Atlantic. Circulation in the tropical region here is controlled by the movements of the intertropical convergence zone (ITCZ) and in the midlatitudes by upper-level troughs (Ritchie et al., 2011). Early in the TC season, systems generally propagate westward before weakening over cooler sea surface temperatures (SSTs). These storms are therefore unlikely to undergo ET. Later in the season, the midlatitude trough extends further into the tropics and TCs are steered more to the north. These storms can then recurve back towards the North American continent, a process typically associated with the conditions of the subtropical highs (Song et al., 2011). The interaction with the midlatitude trough can bring about the onset of ET; however, it is rare for storms to complete transition as making landfall over Central and North America provides a poor energy source (Ritchie et al., 2011). Occasionally storms travel westwards across the Pacific in tropical latitudes before then moving polewards off the coast of the Asian continent and returning back to North America across the top of the Pacific in the region of Alaska and Canada. This latter movement is much rarer than the former, accounting in part for the lower rates of transition in the eastern North Pacific when compared to its western counterpart and the North Atlantic.

In the northern Indian Ocean the low rate of transition is supported by theoretical understanding. The position of the Asian continent at the North of this basin results in storms almost certainly making landfall, losing their supply of energy and therefore decaying before they can transition. Thus, the majority of tropical storms do not undergo ET. The finding that 31% (10 of 32) of storms do undergo some sort of transition is perhaps therefore surprising. However, it is clear that changes to storm systems in the CPS do resemble a transition. This is a phenomenon that is scarcely discussed and would warrant further investigation.

In all, a large part of the disparity between estimations of ET is believed to be the result of different track lengths. This analysis uses storm track invariably longer than other analyses. By charting a storm for a longer period of time, observing a structural transition will become increasingly more likely.

4.2. Hart parameters

Table 4 shows the CPS parameters at the onset and completion of ET and these values are compared with the definitions of ET presented elsewhere. Evans and Hart (2003) proposed a value of $B = 10$ m for the onset of ET and $-|V_T^L| < 0$ for completion. The values determined here

Table 4. Mean Hart parameters for extratropical transition determined from this dataset

Basin	Starting B	Ending $-V_T^L$
Atlantic Ocean	13.2	7.5
N. Indian Ocean	10.7	36.4
W. Pacific Ocean	12.8	9.5
E. Pacific Ocean	8.0	0.3
Northern Hemisphere	11.2	13.4

The distribution for the four individual basins are shown in Fig. 16.

for the eastern North Pacific broadly correspond to this, with the mean value of the B parameter at the start of transition in this basin being 8.0 m and at completion $-|V_T^L|$ being 0.3. However, there is considerable variability across the basins. In fact, the mean results for the whole hemisphere suggest that a starting symmetry value for ET of 10 m is appropriate (the mean starting B is found to be 11.2 m here), but the final value of $-|V_T^L|$ is rather different (13.4). There is only minor variation in values of starting B across the basins (variation of only 5.2 m). The variation in $-|V_T^L|$ at completion varies across an order of magnitude however, with the mean value in the eastern North Pacific being 0.3 and 36.4 in the northern Indian.

Within basins themselves there is also great variability in these values and this is demonstrated in Fig. 16. For example, B at ET onset in the western North Pacific is found to have an interquartile range of around 18 m. This is an important difference and the reason for this is currently unclear. It may depend on season, location of transition, the large-scale metrological condition, for example El Niño and other environmental factors. However, this does support the conclusion that routes through transition are much more heterogeneous than have been previously represented. There seems to be the most variation in the final $-|V_T^L|$ value. For example, in the North Atlantic this parameter ranges over 100.

4.3. Error estimation

The uncertainty associated with this analysis may be partitioned into error associated with the clustering and error associated with the determination of ET. In practical terms the error associated with the clustering is trivial. The K-means clustering algorithm is designed to prevent false cluster attribution by repeating itself a number times (1000) to achieve statistical robustness. There are alternate clustering methods, in particular one that allows a single point to exist simultaneously in more than one cluster (called 'fuzzy' clustering) (Bezdek, 1981), but it unclear whether these would improve the results.
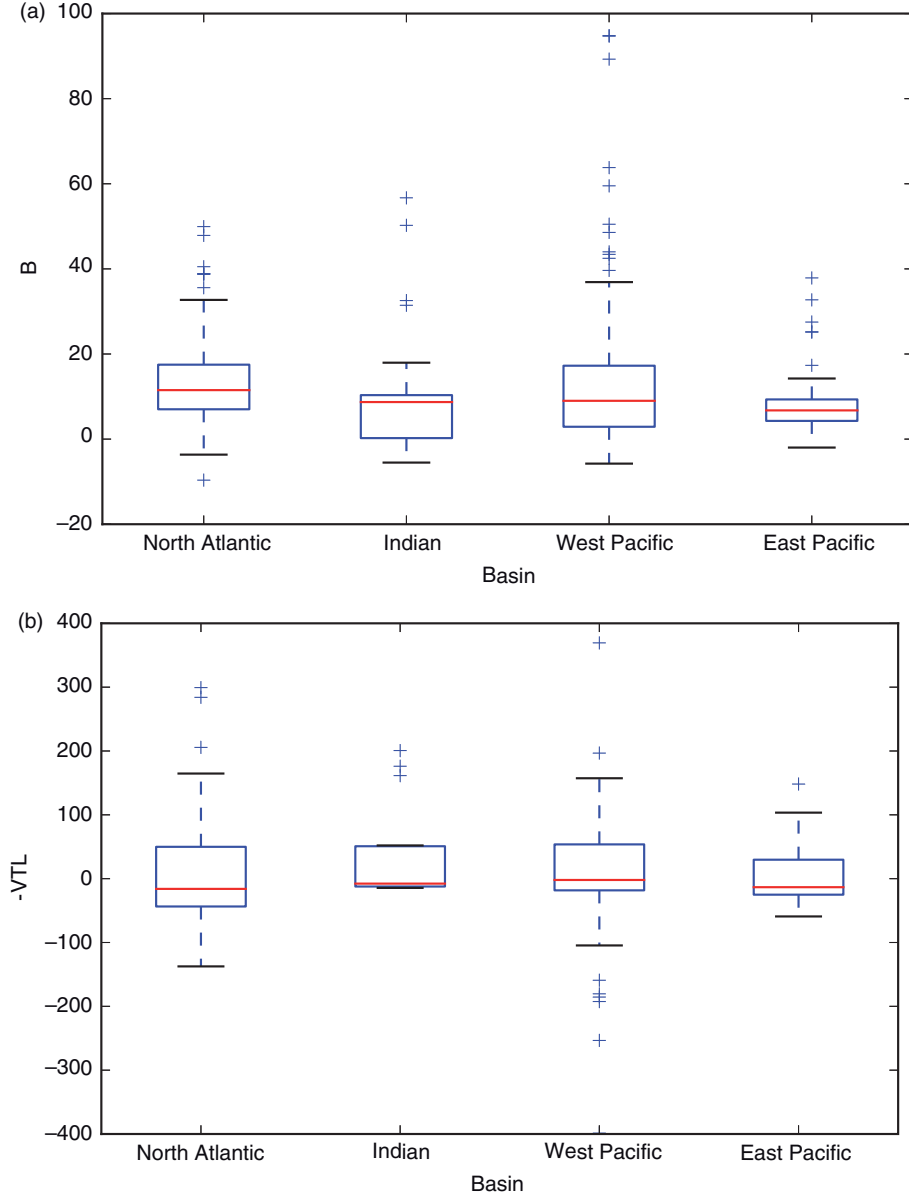
Fig. 16.   Box and Whisker plots of CPS parameter across the four basins. (a) *B* at ET onset and (b) $-|V_T^L|$ at ET completion.[10].

The second error source, determination of the start and end of transition, presents a difficult task. The most obvious scheme would be to compare ET onset and completion times between those estimated here and those recorded in the best track records. However, the best track archives do not record transition as a life-cycle stage in all basins, and those it does record, do it inconsistently. Rather a transitioning storm is still categorised as a TC. As such, error estimation remains the subject of continuing work. Comparison with existing climatologies indeed remains difficult as these are based on the best track records that

use significantly shorter track lengths and working definitions for ET diverge widely.

It remains important to consider the representation of ET, TCs and ECs in numerical models. Whilst models currently operate objectively better than ever before it is true that ET is still not truly represented (Evans et al., 2006). Quinting et al. (2014) find that the ECMWF operational analysis does not adequately capture deep convection. This is a clear limiting factor on the presented analysis if the method is to be used for the study of the phenomenon rather than for a forecasting purpose.

It suggests that in an operational context some implementation of 'calibration' might be required, which could introduce its own uncertainties. However, this is not an issue with the method used here.

4.4. Implications

One of the most important results of this work is the observation that there is huge variation in the pathways through ET. Figure 5 shows mean transitions for each of the three transition modes (one, two and three) for the Northern Hemisphere as estimated by the method presented here. Whilst the general behaviour is distinguishable it is clear that the paths do not move cleanly through the CPS. This is because they are the results of compressing a number of paths that despite operating in the same area of the CPS (i.e. being of the same type) diverge wildly. This supports other conclusions of this work in suggesting that a universal definition of transition is perhaps unfeasible.

This observation could spur further work since much remains unknown about the causes, nature and effects of ET. The presented method removes the requirement to manually identify transition and as such it could be applied to study the transition of tropical storms in a novel way. For example, the general synoptic conditions associated with ET could be explored to determine the causes and implications of transition. Whilst the ECMWF operational analyses dataset does not offer a long enough time period to do this there are others that do. For example, the Athena project produces high-resolution global climate simulations which extend back to 1960 (Jung et al., 2012). New, regularly updated high-resolution reanalysis are beginning to be produced such as the ERA-SAT and the NCEP-CFSR. These will be very fruitful data sources for further study. One particular area for further study is the variability demonstrated in the Hart parameters.

ET may be effected by the anthropogenic influences on climate. This method could be applied to high-resolution global climate models (GCMs) run under a number of future emission scenarios to explore how transition rates, times, types, and locations might operate in the near future. Future projections of TC activity indicate that greenhouse warming will result in a trend towards stronger tropical storms and there is much uncertainty about how the frequency of the globally averaged number of storms will be effected (Knutson et al., 2010). Investigating the role of ET in these projections is a fascinating avenue opened by the presented methodology.

Understanding the predictability of storms in numerical models is an active area of research (e.g. Froude et al., 2014). Determining a model's ability to predict a storm can be done by matching a modelled storm with best track data and then comparing its location and intensity (Froude, 2010).

A scheme for completing a similar analysis for ET is less clear since the underlying metrics (i.e. its onset and completion times, intensity, etc.) of a transition have historically been less clear. However, since this method no longer relies upon a subjective interpretation of ET it presents the opportunity to investigate the predictability of transition in forecasting models.

5. Summary

We find that 54% (± 7, 31 of 59) of the Northern Hemisphere's tropical storms undergo ET in a given year. Of the Northern Hemisphere ocean basins, the western North Pacific and the North Atlantic have a high rate of ET (approximately three in five) whilst the eastern North Pacific and northern Indian Oceans have much lower rates of transition (approximately 3 in 10). Storms undergoing transition in these latter basins rarely complete transition in part because the general steering flow results in them making landfall and thus forces them to abandon their source of energy.

It is difficult to compare the results of this analysis to other estimates as methodologies diverge widely as do the periods of analysis. Additionally there is no other study that uses a comparably high number of storms or uses storm tracks that capture as much of a storm's life cycle. Whilst there are a number of alterations that could be made to the method, it is not clear that they would improve the results. In fact, the error associated with the presented method in terms of transition's onset and completion time remains very difficult to estimate. Nonetheless, there is no reason to suggest this analysis's estimates of transition are less accurate than others.

The method is effective at determining ET transition. Indeed, it is also hoped that the results of this study will inform further theoretical work into the phenomenon of transition. One application in work currently underway is the analysis of the effect of lead times on forecasts of transition and to produce composites of ET, complementing the work of Evans et al. (2006) and others. More generally, it could be used to examine the environmental conditions and various storm structures associated with ET to better understand the phenomenon. Although the method presented is focused on ET, it is recognised that is could be applied to a range of other problems such as studying the behaviour of the El Niño Southern Oscillation.

6. Acknowledgement

Algorithm 1. Steps in the extratropical transition determination algorithm.

Cluster track against reference clustering of all points in database of the same basin

for *Clusters* **do**

    if *Of all clusters this one has the most points in the $B > 10$, $-V_T^L > 0$ space*

    then

        This cluster represents ET type 1

    else if *Of all clusters this one has the most points in the $0 < B < 10$, $-V_T^L < 0$ space* **then**

        This cluster represents ET type 2

    else if *Of all clusters this one has the most points in the $10 < B < 40$, $-100 < -V_T^L < 0$ space* **then**

        This cluster represents ET type 3

    end if

end for

for *ET type clusters* **do**

    Remove all points with $B < 8$ and $-V_T^L > 0$

    Group all points into chronological sequences

    for *Each of these sequences ('potential ETs')* **do**

        The moment B drops or $-V_T^L$ increases declare the transition to be over

        if *Sequence is a singleton set* **then**

            Look the points immediately before and after

        end if

        Remove sequence from list of potential transitions if latitude change of the whole storm life – cycle is less then 10 degrees

    end for

end for

References

Angove, M. D. and Falvey, R. J. 2011. *Annual Tropical Cyclone Report: 2011*. Technical Report, JTWC, Pearl Harbor, Hawaii.

Arnott, J. M., Evans, J. L. and Chiaromonte, F. 2004. Characterization of extratropical transition using cluster analysis. *Mon. Weather Rev.* **132**(12), 2916–2937.

Bengtsson, L., Botzet, M. and Esch, M. 1995. Hurricane-type vortices in a general circulation model. *Tellus A.* **47**, 175–196.

Bengtsson, L., Hodges, K., Esch, M., Keenlyside, N., Kornblueh, L. and co-authors. 2007. How may tropical cyclones change in a warmer climate? *Tellus.* **59**(4), 539–561.

Bezdek, J. C. 1981. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York.

Buizza, R., Miller, M. and Palmer, T. N. 1999. Stochastic Representation of Model Uncertainties in the ECMWF Ensemble Prediction System. *Q. J. R. Meteorol. Soc.* **125**(560), 2887–2908.

DiMego, G. J. and Bosart, L. F. 1982. The transformation of Tropical Storm Agnes into an extratropical cyclone. Part I: The observed fields and vertical motion computations. *Mon. Weather Rev.* **110**(5), 385–411.

ECMWF, 2014. Online IFS Documentation. Online at: http://www.ecmwf.int/research/ifsdocs/

Elsberry, R. L., Klein, P. M. and Harr, P. A. 2000. Extratropical transition of western North Pacific tropical cyclones: an overview and conceptual model of the transformation stage. *Weather Forecast.* **15**(4), 373–395.

Elsberry, R. L. and Ritchie, E. A. 2007. Simulations of the extratropical transition of tropical cyclones: phasing between the upper-level trough and tropical cyclones. *Mon. Weather Rev.* **135**(3), 862–876.

Emanuel, R. 2008. The hurricane-climate connection. *Bull. Amer. Meteor. Soc.* **89**, ES10–ES20.

Emanuel, K., Fondriest, F. and Kossin, J. 2012. Potential economic value of seasonal Hurricane forecasts. *Weather Clim. Soc.* **4**(2), 110–117.

Evans, A. D. and Falvey, R. J. 2013. *Annual Tropical Cyclone Report: 2012*. Joint Typhoon Warning Center, Pearl Harbor, Hawaii, pp. 1–117.

Evans, C. and Hart, R. 2008. Analysis of the wind field evolution associated with the extratropical transition of Bonnie (1998). *Mon. Weather Rev.* **136**(6), 2047–2065.

Evans, J. L., Arnott, J. M. and Chiaromonte, F. 2006. Evaluation of operational model cyclone structure forecasts during extratropical transition. *Mon. Weather Rev.* **134**(11), 3054–3072.

Evans, J. L. and Hart, R. 2003. Objective indicators of the life cycle evolution of extratropical transition for Atlantic tropical cyclones. *Mon. Weather Rev.* **131**, 909–925.

Froude, L. S. R. 2010. TIGGE: comparison of the prediction of Northern Hemisphere extratropical cyclones by different ensemble prediction systems. *Weather Forecast.* **25**(3), 819–836.

Froude, L. S. R. 2011. TIGGE: comparison of the prediction of Southern Hemisphere extratropical cyclones by different ensemble prediction systems. *Weather Forecast.* **26**(3), 388–398.

Froude, L. S. R., Bengtsson, L. and Hodges, K. 2007a. The predictability of extratropical storm tracks and the sensitivity of their prediction to the observing system. *Mon. Weather Rev.* **135**(2), 315–333.

Froude, L. S. R., Bengtsson, L. and Hodges, K. 2007b. The prediction of extratropical storm tracks by the ECMWF and NCEP ensemble prediction systems. *Mon. Weather Rev.* **135**(7), 2545–2567.

Froude, L. S. R., Bengtsson, L. and Hodges, K. 2014. Atmospheric predictability revisited. *Tellus A.* **65**, 19022, DOI: http://dx.doi.org/10.3402/tellusa.v65i0.19022

Grise, K. M., Son, S. W. and Gyakum, J. R. 2013. Intraseasonal and interannual variability in north American storm tracks and its relationship to equatorial pacific variability. *Mon. Weather Rev.* **141**(10), 3610–3625.

Hart, R. 2003. A Cyclone Phase Space derived from thermal wind and thermal asymmetry. *Mon. Weather Rev.* **131**(4), 585–616.

Hart, R. and Evans, J. L. 2001. A climatology of the extratropical transition of Atlantic tropical cyclones. *J. Clim.* **14**(4), 546–564.

Hart, R., Evans, J. L. and Evans, C. 2006. Synoptic composites of the extratropical transition life cycle of north Atlantic tropical cyclones: factors determining posttransition evolution. *Mon. Weather Rev.* **134**(2), 553–578.

Hodges, K. 1994. A general-method for tracking analysis and its application to meteorological data. *Mon. Weather Rev.* **122**(11), 2573–2586.

Hodges, K. 1995. Feature tracking on the unit-sphere. *Mon. Weather Rev.* **123**(12), 3458–3465.

Hodges, K. 1999. Adaptive constraints for feature tracking. *Mon. Weather Rev.* **127**(6 II), 1362–1373.

Hodges, K. and Emerton, R. 2014. Submitted paper: the prediction of Northern Hemisphere tropical cyclones by the ECMWF ensemble prediction system. *Mon. Weather Rev.* 1–71.

Hoskins, B. J. and Hodges, K. 2005. A new perspective on Southern Hemisphere storm tracks. *J. Clim.* **18**(20), 4108–4129.

Johnson, A., Wang, X., Kong, F. and Xue, M. 2011. Hierarchical cluster analysis of a convection-allowing ensemble during the Hazardous weather Testbed 2009 spring experiment. Part I: Development of the object-oriented cluster analysis method for precipitation fields. *Mon. Weather Rev.* **139**(12), 3673–3693.

Jones, S. C., Harr, P. H., Abraham, J., Bosart, L. F., Bowyer, P. J. and co-authors. 2003. The extratropical transition of tropical cyclones: forecast challenges, current understanding, and future directions. *Weather Forecast.* **18**(6), 1052–1092.

Jung, T., Miller, M. J., Palmer, T. N., Towers, P., Wedi, N. and co-authors. 2012. High-resolution global climate simulations with the ECMWF model in project Athena: experimental design, model climate, and seasonal forecast skill. *J. Clim.* **25**(9), 3155–3172.

Kitabatake, N. 2011. Climatology of extratropical transition of tropical cyclones in the Western North Pacific defined by using cyclone phase space. *J. Meteorol. Soc. Jpn.* **89**(4), 309–325.

Klein, P. M., Harr, P. A. and Elsberry, R. L. 2000. Extratropical Transition of Western North Pacific Tropical Cyclones: An Overview and Conceptual Model of the Transformation Stage. *Wea. Forecasting* **15**, 373–395.

Knapp, K. R., Kruk, M. C., Levinson, D. H., Diamond, H. J. and Neumann, C. J. 2010. The International Best Track Archive for Climate Stewardship (IBTrACS). *Bull. Am. Meteorol. Soc.* **91**(3), 363–376.

Knutson, T. R., McBride, J. L., Johnny Chan, J. C., Emanuel, K., Holland, G. and co-authors. 2010. Tropical cyclones and climate change. *Nat. Geosci.* **3**(3), 157–163.

Kofron, D. E., Ritchie, E. A. and Tyo, J. S. 2010. Determination of a consistent time for the extratropical transition of tropical cyclones. Part I: Examination of existing methods for finding "ET time." *Mon. Weather Rev.* **138**(12), 4328–4343.

McTaggart-Cowan, R., Galarneau, T. J. Jr., Bosart, L. F., Moore, R. W. and Martius, O. 2013. A global climatology of baroclinically influenced tropical cyclogenesis. *Mon. Weather Rev.* **141**(6), 1963–1989.

Muramatsu, T. 1985. *The Study on the Changes of the Three-Dimensional Structure and Movement Speed of the Typhoon through Its Life Time*. Technical Report 14, Meteorological Research Institute, Tokyo, Japan.

Murphy, K. P. 2012. *Machine Learning. A Probabilistic Perspective*. The MIT Press, Cambridge, MA.

Ooyama, K. 1969. Numerical simulation of the life cycle of tropical cyclones. *J. Atmos. Sci.* **26**, 3–40.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B. and co-authors. 2011. Scikit-learn: machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830.

Quinting, J. F., Bell, M. M. and Harr, P. A. 2014. Structural characteristics of T-PARC Typhoon Sinlaku During Its Extratropical Transition. *Mon. Weather Rev.* **142**, 1945–1961.

Richardson, D., Buizza, R. and Hagedorn, R. 2005. Final report of the 1st Workshop on the THORPEX Interactive Grand Global Ensemble (TIGGE). *WMO, WWRP-THORPEX No.5*. Reading, United Kingdom. Available at: http://www.wmo.int/pages/prog/arep/wwrp/new/documents/TIGGEFirstWorkshop Report.pdf

Ritchie, E. A. and Elsberry, R. L. 2003. Simulations of the extratropical transition of tropical cyclones: contributions by the midlatitude upper-level trough to reintensification. *Mon. Weather Rev.* **131**(9), 2112–2128.

Ritchie, E. A., Wood, K. M., Gutzler, D. S. and White, S. R. 2011. The influence of Eastern Pacific tropical cyclone remnants on the Southwestern United States. *Mon. Weather Rev.* **139**(1), 192–210.

Serra, Y. L., Kiladis, G. N. and Hodges, K. 2010. Tracking and mean structure of easterly waves over the Intra-Americas Sea. *J. Clim.* **23**(18), 4823–4840.

Sinclair, M. 2002. Extratropical transition of southwest Pacific tropical cyclones. Part I: Climatology and mean structure changes. *Mon. Weather Rev.* **130**(3), 590–609.

Small, C. and Cohen, J. E. 2004. Continental physiography, climate, and the global distribution of human population. *Curr. Anthropol.* **45**(2), 269–277.

Song, J., Han, J. and Wang, Y. 2011. Cyclone phase space characteristics of the extratropical transitioning tropical cyclones over the western North Pacific. *Acta Meteorologica Sinica.* **25**(1), 78–90.

Stewart, S. 2013. *NHC Tropical Cyclone Report: Hurricane Leslie (AL122012)*. Technical Report, National Hurricane Center, Miami, Florida, USA.

Strachan, J., Vidale, P. L., Hodges, K., Roberts, M. and Demory, M. E. 2013. Investigating global tropical cyclone activity with a hierarchy of AGCMs: the role of model resolution. *J. Clim.* **26**(1), 133–152.

Veren, D., Evans, J. L., Jones, S. C. and Chiaromonte, F. 2009. Novel metrics for evaluation of ensemble forecasts of tropical cyclone structure. *Mon. Weather Rev.* **137**(9), 2830–2850.

Wang, Q., Li, Q. and Fu, G. 2012. Determining the extratropical transition onset and completion times of Typhoons Mindulle (2004) and Yagi (2006) using four methods. *Weather Forecast.* **27**(6), 1394–1412.

Yokoi, S., Takayabu, Y. N., Nishii, K., Nakamura, H., Endo, H. and co-authors. 2011. Application of cluster analysis to climate model performance metrics. *J. Appl. Meteorol. Climatol.* **50**(8), 1666–1675.

Zappa, G., Stephenson, D. B., Sansom, P. G., Ferro, C. A. T. and Shaffrey, L. 2013. Simple uncertainty frameworks for selecting weighting schemes and interpreting multimodel ensemble climate change experiments. *J. Clim.* **26**(12), 4017–4037.