

# *The evolution, seasonality, and impacts of western disturbances*

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Hunt, K. M. R. ORCID: <https://orcid.org/0000-0003-1480-3755>,  
Turner, A. G. ORCID: <https://orcid.org/0000-0002-0642-6876>  
and Shaffrey, L. C. ORCID: <https://orcid.org/0000-0003-2696-752X> (2018) The evolution, seasonality, and impacts of western disturbances. Quarterly Journal of the Royal Meteorological Society, 144 (710). pp. 278-290. ISSN 1477-870X doi: <https://doi.org/10.1002/qj.3200> Available at <https://centaur.reading.ac.uk/73137/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1002/qj.3200>

Publisher: Royal Meteorological Society

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

[www.reading.ac.uk/centaur](http://www.reading.ac.uk/centaur)

**CentAUR**

Central Archive at the University of Reading

Reading's research outputs online



# The evolution, seasonality and impacts of western disturbances

Kieran M. R. Hunt,<sup>a\*</sup> Andrew G. Turner<sup>a,b</sup> and Len C. Shaffrey<sup>b</sup>

<sup>a</sup>Department of Meteorology, University of Reading, UK

<sup>b</sup>NCAS–Climate, University of Reading, UK

\*Correspondence to: K. M. R. Hunt, Department of Meteorology, University of Reading, Earley Gate, PO Box 243, Reading RG6 6BB, UK. E-mail: k.m.r.hunt@reading.ac.uk

Western disturbances (WDs) are upper-level synoptic-scale systems embedded in the subtropical westerly jet stream (STWJ), often associated with extreme rainfall events in north India and Pakistan during boreal winter. Here, a tracking algorithm is applied to the upper-tropospheric vorticity field for 37 years of ERA-Interim reanalysis data, giving a catalogue of over 3000 events. These events are analysed in a composite framework: the vertical structure is explored across a large number of dynamic and thermodynamic fields, revealing a significant northwestward tilt with height, strong ascent ahead of the centre, which sits above the maximum surface precipitation, and a warm-over-cold, dry-over-moist structure, among other signatures of strong baroclinicity. Evolution of the structures of cloud cover and vertical wind speed are investigated as the composite WD passes across northern India. Cloud cover in particular is found to be particularly sensitive to the presence of the Himalayan foothills, with a significant maximum at 300 hPa approximately 1 day after the WD reaches peak intensity. *k*-means clustering is used to classify WDs according to both dynamical structure and precipitation footprint and the relationship between the two sets of WDs is explored. Finally, the statistical relationship between the STWJ position and WDs on interannual time-scales is explored, showing that WD frequency in north India is highly sensitive to the jet location over Eurasia. Years with a greater number of WDs feature a STWJ shifted to the south, a pattern that is substantially more coherent and reaches as far west as North America during boreal winter. This suggests that it may be possible to predict the statistics of western disturbance events on seasonal time-scales if suitable indicators of jet position can also be predicted.

**Key Words:** western disturbances; subtropical jet stream; India, baroclinic; orography; composite; tracking

Received 30 May 2017; Revised 3 October 2017; Accepted 12 October 2017; Published online in Wiley Online Library

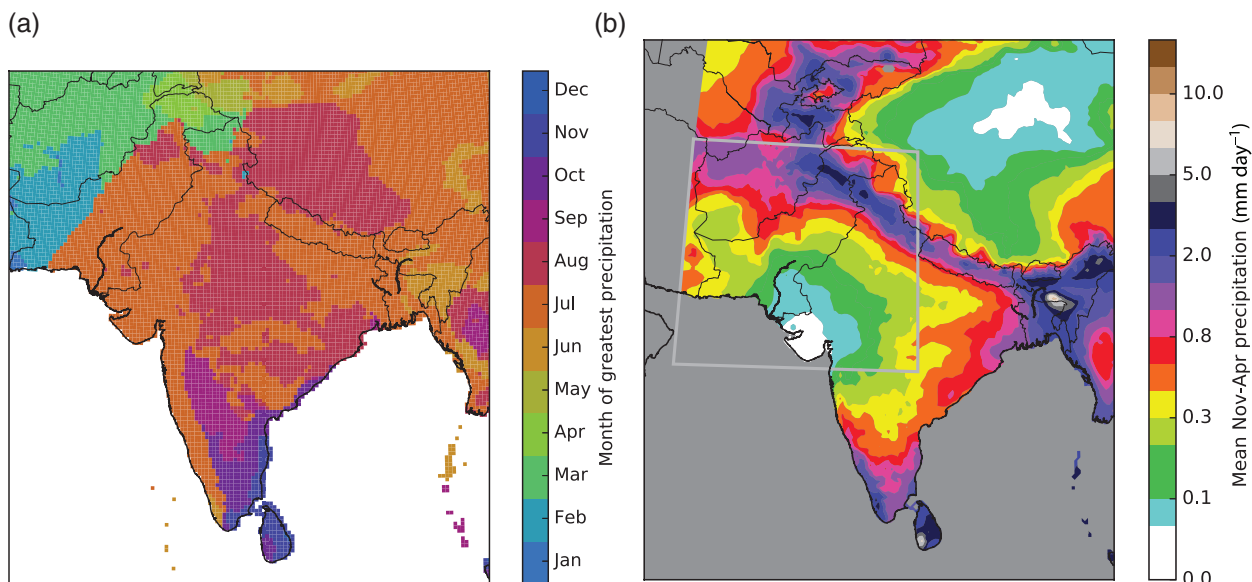
## 1. Introduction

Western disturbances (WDs) are formally defined by the India Meteorological Department (IMD) as: ‘[A] cyclonic circulation/trough in the mid and lower tropospheric levels or as a low pressure area on the surface, which occurs in middle latitude westerlies and originates over the Mediterranean Sea, Caspian Sea and Black Sea and moves eastwards across north India’.\* WDs are, at the most fundamental level, synoptic-scale vortical perturbations embedded in the subtropical westerly jet stream (Dimri *et al.*, 2016). They are often associated with extreme rainfall events in the Karakoram and Hindu Kush regions of Pakistan and north India (e.g. Chakravorty and Basu, 1957; Mooley, 1957; Sharma and Subramaniam, 1983; Rangachary and Bandyopadhyay, 1987; Lang and Barros, 2004; Roy and Roy Bhowmik, 2005; Kotal *et al.*, 2014; Dimri *et al.*, 2015) and have been the subject of a number

of modelling case studies (e.g. Ramanathan and Saha, 1972; Chitlangia, 1976; Gupta *et al.*, 1999; Azadi *et al.*, 2002; Dimri, 2004; Semwal and Dimri, 2012; Dimri and Chevuturi, 2014; Patil and Kumar, 2016). Despite the abundant interest in these systems, little headway has been made in developing a systematic understanding of them. Cannon *et al.* (2016) recently developed a wave-tracking algorithm based on 500 hPa geopotential height anomaly using daily NCEP–CFRSR reanalysis data (Saha *et al.*, 2010). However, their computed frequency of 5–6 events passing through north India per year is roughly an order of magnitude less than most anecdotal climatologies suggest it should be, around 4–7 per month during boreal winter (Chattopadhyay, 1970; Mohanty *et al.*, 1998; Dimri, 2006). Either way, it is still not clear what the seasonal cycle of western disturbances looks like, despite some studies that have looked at variability in the mid-tropospheric geopotential height (e.g. Madhura *et al.*, 2015).

Knowing that the westerly jet responsible for bringing these disturbances to the Indian subcontinent is both stronger and further south over the winter (Schiemann *et al.*, 2009), we can reasonably

\*<http://imd.gov.in/section/nhac/wxfaq.pdf>



**Figure 1.** Some precipitation statistics computed for north India, using APHRODITE data (1951–2007): (a) the month with the highest climatological precipitation and (b) mean precipitation during November–April, with dark grey indicating areas out of the data domain. The track filtering domain is marked in (b) by a light grey box, bounded by the coordinates 20–36.5° N, 60–80° W (see section 2.2 for explanation). The dark grey in (b) indicates missing data.

assume that winter WDs will be correspondingly more frequent and responsible for more precipitation. Figure 1(a) shows the climatologically rainiest month over South Asia, computed from Asian Precipitation – Highly Resolved Observational Data Integration towards Evaluation of water sources (APHRODITE) data (see section 2.1.1). As expected, the southwest monsoon dominates the picture, with July and August the wettest months across the monsoon trough, whereas in the south the rain shadow of the Western Ghats is wettest during the northeast monsoon in the late autumn. In the vicinity of the northern mountain ranges, however, there is a significant area where late winter and spring bring the wettest weather (Palazzi *et al.*, 2013). This, we might hypothesize, is the area most susceptible to heavy rainfall from WDs.

Without a large dataset of WDs, it is impossible to make any assertions about their gross structure or interaction with the large-scale, other than by theoretical or case-study analysis. It has been shown that, despite the complicated orography and fairly small spatial scale of WDs, they are typically well represented in reanalyses (Mohanty *et al.*, 1998, 1999).<sup>†</sup> Therefore we must ask whether we can use a reanalysis to understand better the evolution, spatial distribution and seasonality of WDs.

We will start with an outline of the data sources and methods used in this study in section 2, before discussing our results in section 3. These results include exploration of the seasonal cycle (section 3.1) and structure (section 3.2) of western disturbances, as well as an examination of their evolution (section 3.3), an automated classification through dynamic structure and precipitation (section 3.4) and a statistical view of interaction with the jet (section 3.5), which, as we shall show, offers the prospect for predicting the statistics of western disturbances on seasonal time-scales. We conclude the discussion in section 4.

## 2. Data and methodology

### 2.1. Data sources

#### 2.1.1. APHRODITE

APHRODITE (Yatagai *et al.*, 2009, 2012) is a daily, 0.25° resolution, gridded, gauge-based precipitation product. Covering

<sup>†</sup> Global reanalyses at this time were fairly coarse: for example ERA-15, released in 1997, was at T106 (roughly 190 km at the Equator).

the period from 1951–2007 inclusive, it is one of the longest precipitation datasets over continental Asia and one of the few that uses surface observations. APHRODITE performs comparatively well against satellite-only or blended products, particularly over orography (Guo *et al.*, 2015).

#### 2.1.2. TRMM

The *Tropical Rainfall Measuring Mission* (TRMM: Simpson *et al.*, 1996; Kummerow *et al.*, 1998) was a joint National Aeronautics and Space Administration (NASA)/Japan Aerospace Exploration Agency (JAXA) satellite with onboard 13.8 GHz precipitation radar and a five-channel (spanning 10.7–85.5 GHz) microwave imager that was operational between late 1998 and early 2015. Among the products released by the project was the multi-satellite precipitation analysis (Huffman *et al.*, 2007, 2010), most commonly known by its algorithm name, 3B42 (version 7). This combines data collected by the TRMM satellite with infrared (IR) images from a selection of geostationary satellites to produce a continuous, three-hourly, 0.25° resolution product between 50° N and 50° S. The advantages of 3B42 over a gauge-based product are temporal resolution and coverage over the oceans. TRMM 3B42 has been shown to perform well over India, despite struggling over orography (Nair *et al.*, 2009; Dinku *et al.*, 2010). Performance over orography was, however, greatly improved in version 7 (Zulkafli *et al.*, 2014). Note that when we come to look at composites of precipitation using TRMM 3B42, the northern limit of the data used will be 46.5° N, well within the boundaries of the dataset.

#### 2.1.3. ERA-Interim

ERA-Interim (Dee *et al.*, 2011) is a global, 0.7° reanalysis product developed by the European Centre for Medium-Range Weather Forecasting (ECMWF). Covering 1979 to the present day, it has 37 output vertical levels from the surface to 0.1 hPa, 27 of which are below 100 hPa. The fixed choice of numerical weather prediction model helps to mitigate spurious trends. Data from aircraft, radiosondes, buoys and satellites are assimilated.

## 2.2. Tracking algorithm

Given the difficulty of working with the formal definition provided by the IMD, we started by locating previous case studies in the literature that fell within the ERA-Interim period (1979–present)

and mentioned WDs by specific date. After removing duplicates, 19 candidate WD case studies were found, ranging from March 1979–June 2013. Examination of the vertical structure of vorticity in each of these indicated that the strongest signal was found between 450 and 300 hPa, with a spatial scale of the order of several hundred kilometres.

In order to expand this case study selection into a complete catalogue of western disturbances, we developed a tracking algorithm to be run on ERA-I reanalysis data. The prescription is as follows.

- (1) Compute the mean relative vorticity in the 450–300 hPa layer, then truncate the spectral resolution at T63 (~ 200 km at the Equator). We shall call this quantity  $\xi$ .
- (2) Locate all local maxima in  $\xi$  subject to some radius  $\delta$ , such that a point is considered a local maximum if no points with a distance  $\delta$  have a greater value of  $\xi$ . We shall call this set of local maxima  $\chi_i$ .
- (3) For each  $\chi_i$ , associate local positive non-zero values of  $\xi$  and integrate to find the centroid of  $\xi$  for each.<sup>‡</sup> We shall call this set of points  $X_i$ .
- (4) (a) To group the candidate points into tracks: for each  $X_i$  at time point  $j$ , seek and attach the nearest neighbour from time point  $j + 1$ , so long as it is within some distance  $\Delta$ , using the *kd*-tree nearest-neighbour algorithm (e.g. Yianilos, 1993).  
(b) The efficacy of this step can be increased by introducing the concept of a background velocity, important when considering the high-frequency, high-velocity nature of WDs. Here this is done by biasing the search radius using the contemporaneous wind field: for example, in a wind field  $\mathbf{u}$ , the central location from which the nearest neighbour is sought is not  $X_i$  but  $X_i + \mathbf{u}(X_i) \cdot (t^{j+1} - t^j)$ . Simply put, rather than starting the nearest neighbour search at the location of the candidate point at the previous time point, we assume it is advected by the background winds and start the search from the location where it would have ended up after such advection.
- (5) We also hold the tracks in memory for one time point, looking for a candidate in time point  $j + 2$  within  $2\Delta$  of  $X_i$ . This prevents breaking a track into two pieces unnecessarily in the event of a candidate apparently disappearing for a single time point.
- (6) These resulting tracks are then filtered three times. Firstly, ‘stubs’ of length shorter than 2 days are rejected. Secondly, tracks that do not pass through Pakistan or north India, here defined as 20–36.5° N, 60–80° E (see Figure 1(b)), are rejected as not of interest to this study. Finally tracks whose geneses are east of their lysens and thus do not propagate eastwards are rejected.
- (7) The values of  $\delta$  and  $\Delta$  are determined empirically by running the algorithm over the 19 case studies identified from previous literature and choosing the combination giving the closest match. These were chosen to be 850 km and 1000 km (6 h)<sup>-1</sup>, respectively.

We applied this algorithm directly to ERA-Interim data for the entire available period (1979–2016), recovering 3090 WDs in total at a frequency of approximately 6–7 per month; a spatial summary of these tracks is given in Figure 2. Figure 2(b) shows all 3090 tracks, revealing a large diversity in behaviour, with a substantial subset originating from as far west as the North American continent. Since genesis regions vary from the Arctic Circle to as far south as the Caribbean, the governing dynamics of these vortices will also vary. Figure 2(b) shows that such extremes are relatively uncommon, however.

<sup>‡</sup>The reader may like to think of this as a process whereby ‘blobs’ of positive vorticity surround the local maxima and a two-dimensional integration is

### 2.3. *k*-means clustering algorithm

One aspect of this study will be to determine whether rainfall patterns associated with different WD events can be grouped together in some way. To that end, we must choose a suitable clustering algorithm. The *k*-means clustering algorithm (Hartigan and Wong, 1979) is an unsupervised feature-learning algorithm that separates a given set of *n*-dimensional (vector) points into a pre-ordained number (*k*) of clusters that are as distinct as possible. In our case, the components of the vector are given by the individual voxel (or pixel) values of the field of interest. Formally, it aims to select clusters such that the sum of the square of the Euclidean distances of each point in the cluster to its mean is minimized, i.e. the objective function is given by

$$F = \sum_{j=1}^k \sum_{i=1}^{n_j} \|\mathbf{x}_{i,j} - \mu_j\|^2, \quad (1)$$

where  $\mathbf{x}_{i,j}$  is a point vector in cluster *j* and  $\mu_j$  is the centre of cluster *j*, which has *n<sub>j</sub>* points. To achieve this, *k* cluster centres are initialized randomly, then each  $\mathbf{x}$  is assigned to a cluster *j* according to the closest centre. The centres are then recalculated based on the mean value of all assigned points and the step is repeated until convergence. This method can be susceptible to local minima, so it is repeated with different initializations (we use 10) to ensure that the computed minimum of *F* is indeed global. *k*-means clustering has been used in meteorological applications before, by e.g. Kulkarni *et al.* (1992), who classified different rainfall paradigms over peninsular India. For a discussion on the application to our data, see section 3.4.

### 2.4. Cloud colour maps

This method for representing the three-dimensional distribution of cloud cover is presented in full detail in Figure 17 of Hunt and Turner (2017), but we shall summarize briefly it here. As in that study, we use the ERA-Interim definitions for low, medium and high cloud height, i.e. LC < 1.9 km ≤ MC < 6.3 km ≤ HC.<sup>§</sup> Consider the typical RGB colour space, where a coordinate has the form (0 ≤ *R* ≤ 255, 0 ≤ *G* ≤ 255, 0 ≤ *B* ≤ 255). Then, to compute the colour for a given pixel, we scale the values of LC, MC and HC, using them as direct coordinates, scaled by a maximum of 50% observed cloud cover. Thus, for an area of 0% LC, 50% MC and 0% HC, the resulting pixel colour would be green (0, 255, 0).<sup>¶</sup>

## 3. Results

### 3.1. Seasonal cycle

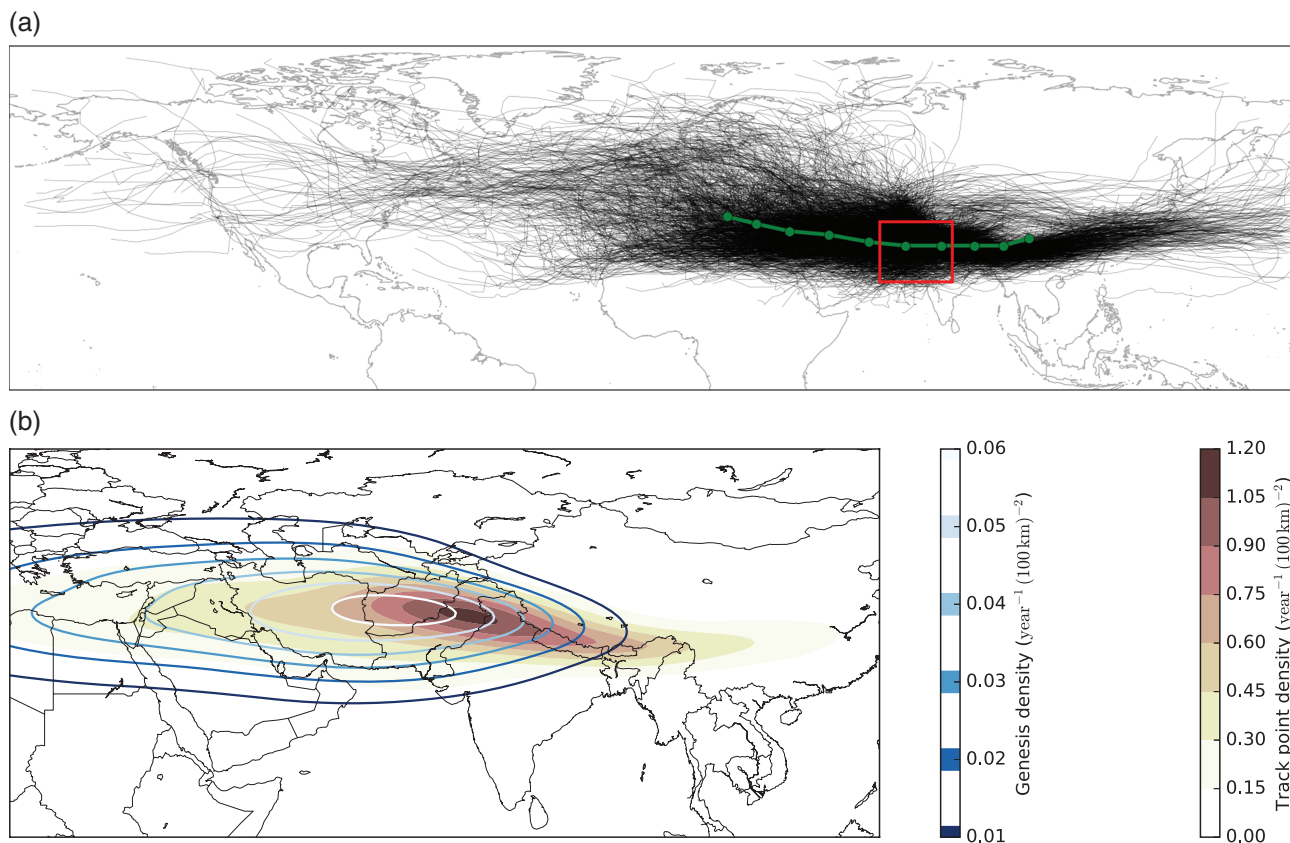
It is well known that the western disturbance frequency peaks during winter (e.g. Singh and Agnihotri, 1977; Dimri *et al.*, 2015), when the subtropical westerly jet (STWJ) is at its most southerly position; we can use this as a consistency check for our database by ensuring the tracked WDs follow the expected seasonal cycle. To do this robustly, we employ the resolution-invariant<sup>||</sup> improvement to the conventional histogram: the kernel density estimation of the probability density function (PDF). Here, the datum of interest for a given WD is the day of the year

computed to find their centroids – this is analogous to finding their centres of mass using vorticity as the density function.

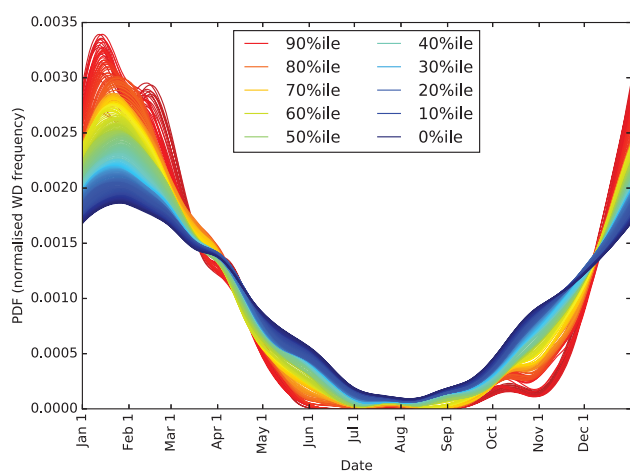
<sup>§</sup>To within 50%, 1.9 and 6.3 km correspond to pressures of 800 and 450 hPa, respectively.

<sup>¶</sup>This would be (0, 128, 0) usually, but we have halved the thresholds here to increase contrast, on account of the low overall cloud cover.

<sup>||</sup>This is not strictly true, as one must select a bandwidth for the convolving function, but the result here is virtually independent of that, except for very large or very small values.



**Figure 2.** Representation of some bulk spatial statistics of WD tracks. (a) All 3090 tracks found in ERA-Interim (1979–2016); (b) contour plots of genesis (lines) and track point (solid) densities, with units of  $\text{year}^{-1} (100 \text{ km})^{-2}$ . In (a), the red box, as in Figure 1(b), indicates the box through which tracks must pass to be considered; the segmented green track indicates the mean position for 5 days either side of maximum intensity (the central point is that on the inside left of the red box).



**Figure 3.** Gaussian (periodic) kernel density estimates of the probability density functions for the Julian day of WD incidence on northern India, for a range of WD strengths. The day used is that during which the WD enters the box given in Figure 1(b) for the first time. Here, for example, the 90th percentile line is a PDF for all WDs with peak T63 450–300 hPa relative vorticity in the top 10%. Resolution is a quarter-percentile. The zeroth percentile blue line is the PDF for the whole dataset.

in which it enters the ‘north India’ box given in Figure 1(b). A ribbon of these PDFs, filtered and coloured by relative intensity, is given in Figure 3; the PDF for the set of all WDs is given by the bluemoest line (the zeroth percentile), whereas, for example, WDs with peak T63 450–300 hPa relative vorticity (in the north India box) falling in the top half of all systems comprise the 50th percentile set. The expected annual cycle is recovered with peak and trough locations independent of WD strength, approximately at the beginning of February and August, respectively. In August, the reduced count is due in part to the more northerly location of the STWJ: the westerlies tend to head north of the Tibetan Plateau and are replaced over the subcontinent by the tropical easterly jet.

As we filter out weak – but progressively raising the minimum threshold – events, i.e. moving towards the redder lines, we see an accentuation of the existing pattern: the strongest systems are relatively more likely than the weaker ones to be found during the winter. In fact, the strongest third of WDs almost never occur during the monsoon months of July and August, whereas they are almost twice as likely to be found during late winter. The crossover dates, roughly the start of December and April, could be considered appropriate boundaries for a ‘western disturbance’ season.

### 3.2. Composite structure

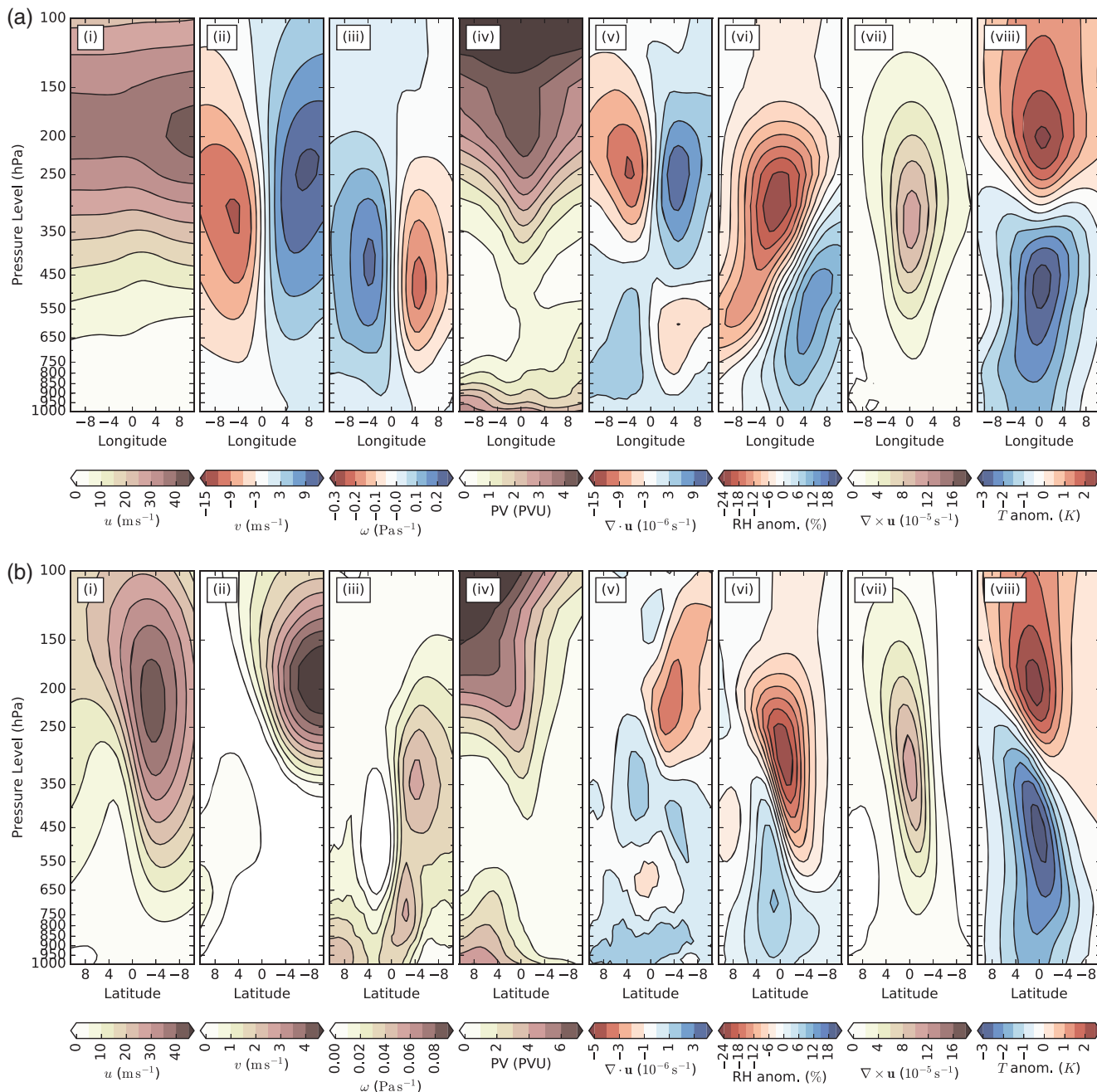
One key advantage of collecting a large catalogue of events is that we can investigate the mean system-centred structure of them to gain an insight into underlying internal and external mechanisms. Here, we will look at the vertical structure of winds, vorticity, potential vorticity, divergence, moisture, and temperature, as well as the footprints of precipitation and cloud cover. Note that, where reanalysis data would be below the surface, they have not been included in the composite.

#### 3.2.1. Vertical structure

We start by exploring the composite vertical structure of the tracked western disturbances. To ensure this discussion is robust in the context of the high variability of track geneses and lyses (see Figure 2(a)), we only composite WDs with centres inside the box shown in Figure 1(b). Figures 4(a) and 4(b) show west–east and north–south slices, respectively, through the centre of the WD composite for a selection of fields, which we shall discuss from left to right for several fields of interest.

##### 3.2.1.1. (i) Zonal wind.

Shown in the leftmost panel, the composite structure of  $u$  clearly displays the westerly jet with which WDs are associated. The jet



**Figure 4.** Vertical composites of western disturbance structure over northern India and Pakistan, taken from (a) west to east and (b) north to south through the centre, which is at  $0^\circ$ . From left to right, the composite fields are as follows: (i) zonal wind speed [ $\text{m s}^{-1}$ ], (ii) meridional wind speed [ $\text{m s}^{-1}$ ], (iii) vertical wind speed [ $\text{Pa s}^{-1}$ ], (iv) potential vorticity [PVU;  $10^{-6} \text{K m}^2 \text{kg}^{-1} \text{s}^{-1}$ ], (v) divergence [ $10^{-6} \text{s}^{-1}$ ], (vi) relative humidity [%], (vii) relative vorticity [ $10^{-5} \text{s}^{-1}$ ] and (viii) temperature anomaly [K]. The anomalies are computed against a 20 day mean. Note that the colour scales between the two sets of cross-sections may differ for a given field. Longitude and latitude are both defined relative to the centre of the composite WD.

is strongest at a height of around 200 hPa, weakening behind the WD and strengthening ahead of it, where it reaches a mean speed of over  $40 \text{ m s}^{-1}$ . Further, as should be expected from a jet–cyclone coupling in the Northern Hemisphere, the jet core sits several hundred kilometres to the south of the WD centre. Where one may expect the zonal wind to have both signs, reflecting circulation, in the north–south cross-section, the signal from the jet is so strong that this is essentially not found.

3.2.1.2. (ii) Meridional wind.

In the second panel, the cross-sections of composite  $v$  give an insight into WD circulation without the dominant westerly contribution from the STWJ. A strong cyclonic circulation is noted in the upper troposphere, albeit slightly asymmetric; the northerly lobe at 250 hPa contrasts the significantly lower southerly lobe at 350 hPa, i.e. the rotation is below the core of the jet. The peak composite wind speed is around  $10 \text{ m s}^{-1}$ , which is comparable to other synoptic systems found in South Asia, e.g.

monsoon depressions (Hunt *et al.*, 2016). We also note that in the north–south cross-section there is evidence of some contribution of converging, weak northerly wind from the jet.

3.2.1.3. (iii) Vertical wind.

In the third panel from the left, the cross-sections of composite  $\omega$  are the first indication here that the WD–jet coupling is nonlinear. One would typically expect some forced uplift at the centre of a non-solenoidal vortex in the absence of interaction with background flow or shear; here, we are far removed from that scenario: strong ascent ahead of the WD with a maximum at 450 hPa, matched by a similar descent behind it. This can be understood in terms of upper-level convergence on inspection of the vertical structure of  $u$ : streamlines of the jet bunch up on approaching the vortex as they are diverted southward, before spreading again on passing. Mass continuity requires that divergence aloft must sit above ascending air and vice versa. The north–south cross-section reveals little in the way of coherent

structure, other than to say that descent is generally ubiquitous and strongest underneath the jet.

#### 3.2.1.4. (iv) Potential vorticity.

In the fourth panel from the left, potential vorticity (PV) has a particularly distinct zonal structure. Vertically, there are maxima at the tropopause and at the surface, with a minimum in the mid-troposphere at 550 hPa. Evidence of a central maximum is present as depression of the PV isolines above 450 hPa. This structure, at least in the mid-troposphere and above, is similar to those of baroclinic waves (e.g. Molinari *et al.*, 1995), nascent extratropical cyclones (Davis, 1992) or even the general case of baroclinic instabilities (Robinson, 1989). The meridional structure bears a strong similarity, differing in having a more prominent lower-tropospheric maximum and a pronounced northwestward slant of the upper-tropospheric maximum. The former is a likely consequence of the interaction between the lower-level circulation and the Himalayan foothills, a coupling that has previously been shown to generate areas of high PV for synoptic-scale tropical systems (Hunt *et al.*, 2016); the poleward component of the latter, as we will discuss shortly, is a common feature of baroclinic instabilities.

#### 3.2.1.5. (v) Divergence.

The zonal structure of divergence (central panel) confirms the strongly baroclinic structure that was inferred earlier. Strong upper-level ( $\sim 250$  hPa) divergence sits ahead of the WD, mirrored by equivalent convergence behind it. Consistent with the ascent–descent pattern we saw in  $\omega$ , a quadrupole emerges, with a weaker, anti-symmetric pattern in the lower troposphere. It is very much plausible that it is this combined convergence and ascent ahead of the WD that interacts constructively with the orography in northern India to produce the intense rainfall that is commonly associated with these systems. We will go into further detail later (in our discussion of Figure 8). Conversely, the meridional structure through the centre is weak and complex. There is convergence associated with the southern edge of the jet, downwards from which a weak wavelike structure with wavelength  $\sim 7$  km appears to emanate. We note that this wavelike structure is consistent with the weak descent seen in the third panel.

#### 3.2.1.6. (vi) Relative humidity.

In the fourth column, the zonal and meridional structures of the relative humidity as anomalies to the 20 day mean are shown. In each instance, there is an axial tilt with height: the nil isohume leans towards the northeast with increasing altitude. Zonally speaking, there is a substantial lower-tropospheric wet anomaly co-located with the ascent east of the centre balanced with an upper-tropospheric ( $\sim 300$  hPa) dry anomaly, the maximum of which is only about 200 km west of the centre. This eastward axial tilt is most likely explained by the vertical wind shear (cf.  $u$ ), but this is not definitely the case. Conversely, the northward tilt in the meridional cross-section, which is structurally very similar, lies parallel to isotachs in  $u$ .

#### 3.2.1.7. (vii) Relative vorticity.

Third from the right, the structure of relative vorticity in both zonal and meridional cross-sections supports our choice of it as a tracking variable and confirms that the heights used (450–300 hPa) were suitable. The central maximum is at an altitude of around 325 hPa, higher even than the extent of an extratropical cyclone (e.g. Dacre *et al.*, 2012), but still underneath the 200 hPa jet. We can rule out orography as a cause by examining the evolution of this field through the WD lifetime (not shown): whilst the magnitude of the relative vorticity waxes and wanes, the spatial scale and location have very low variability. Also notable is the lack of rotational symmetry in this field: this is highly uncommon among synoptic-scale circulations (Dacre *et al.*, 2012; Hurley and Boos, 2015), but not waves of similar

scale (Boettcher and Wernli, 2011). The zonal extent for a given relative vorticity isosurface is much greater than its meridional extent, giving the appearance of a ‘squashed’ vortex. Looking at the meridional cross-section, there is a significant northward tilt of the vortex with height. This northwestward tilt is classic behaviour for a baroclinic disturbance, of which WDs are largely believed to be examples (Brahmananda Rao and Trivikrama Rao, 1971), despite some recent studies claiming the contrary (Hara *et al.*, 2004). Given this and the structure of potential vorticity discussed recently, we can conclude with some certainty that WDs are primarily baroclinic disturbances. The structure of the geopotential anomaly (not shown) has a broad, deep minimum that strongly resembles the case study structure for both observations and a mesoscale model given in Dimri *et al.* (2015); its Laplacian is also proportional to relative vorticity above 450 hPa, indicating geostrophy.

#### 3.2.1.8. (viii) Temperature.

Second from the right, temperature is given as an anomaly to the 20 day mean, in order to disregard anticipated variations from the seasonal cycle. The structure is almost cylindrically symmetric, with a  $\sim 2$  K warm anomaly centred at 200 hPa above a  $\sim 3$  K cold anomaly at around 500 hPa. This differs from the mature extratropical cyclone structure, which typically has a warm anomaly throughout the troposphere, but bears a good similarity to its nascent, immature counterparts (Dacre *et al.*, 2012). Indeed, the cold and moist anomaly in the mid/lower troposphere is also a signature of so-called ‘medicane’ (‘Mediterranean hurricanes’: Emanuel, 2005), some of which continue eastward before interacting with the STWJ and becoming WDs (Madhura *et al.*, 2015). As with relative humidity and potential vorticity, the meridional cross-section demonstrates some northward tilt with height, with the gradient perpendicular to the isotachs of  $u$  – this is a signature of baroclinicity, although it could be explained alternatively by thermal wind balance.

### 3.2.2. Horizontal structure

Western disturbances are commonly referred to in the context of the impact they have at the surface – usually in terms of precipitation. Here, we will look at the composite footprints of WDs as they enter north India. Figure 5 shows the horizontal structure of precipitation, cloud cover, surface pressure and winds at the surface.

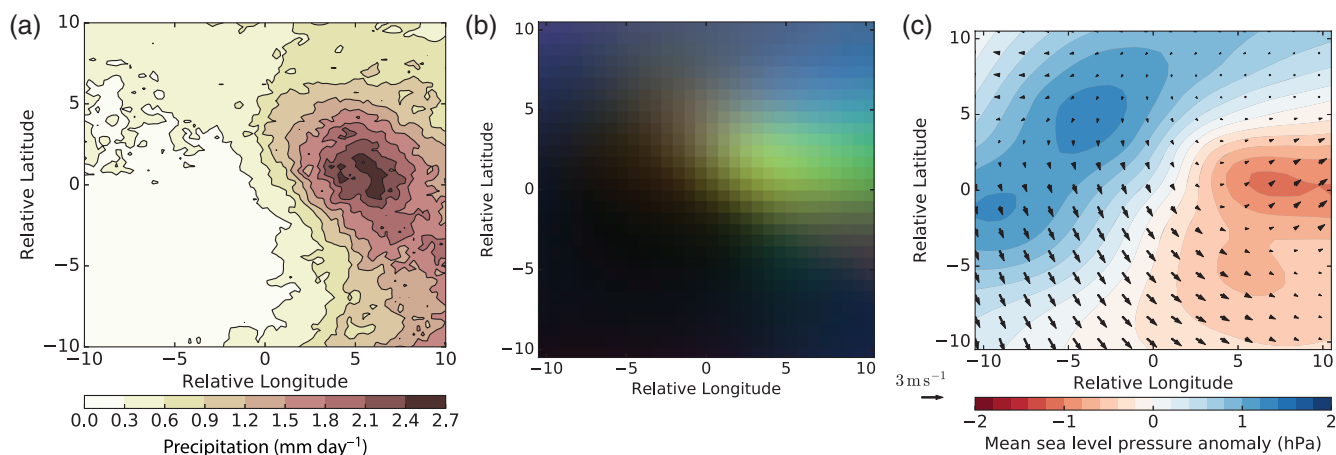
#### 3.2.2.1. Precipitation.

Figure 5(a) shows the composite surface precipitation ( $\text{mm day}^{-1}$ ) from TRMM 3B42 for WDs in north India. The maximum is spatially coherent and located about 500 km due east of the centre of the WD, and is thus co-located with the ascent and low-level convergence shown in Figure 4(a). We also note the fairly low mean rainfall rate of the composite: less than  $3 \text{ mm day}^{-1}$  at the peak. This is because a large number of these WDs ( $\sim 95\%$ ) do not really precipitate heavily; of those that do – i.e. the remaining 5%, which have a frequency of about  $4 \text{ year}^{-1}$  – the peak mean precipitation rate is around  $40 \text{ mm day}^{-1}$ .

#### 3.2.2.2. Cloud cover.

Figure 5(b) shows the amount of low-, mid- and high-level cloud (LC, MC, HC respectively) in ERA-Interim associated with the composite WD. A description of the methodology used in this figure is given in section 2.4. Examination reveals that much of the area under the composite WD is almost black, indicating no substantial cloud cover at any level; this is most obviously the case to the west and southwest of the WD centre, where there is organized large-scale descent (see Figure 4(a),  $\omega$ ). Towards the north and east, however, a complex pattern emerges: the far north ( $\sim 600$  km from the centre) is dominated by high cloud, whereas the west and centre have a stronger presence of low and mid cloud.





**Figure 5.** A selection of surface impacts from the composite WD: (a) precipitation [ $\text{mm day}^{-1}$ ] from TRMM 3B42; (b) colour-wheel cloud cover [red: low; blue: medium; green: high; black: no cloud; white: full cloud] from ERA-Interim, see figure 17 of Hunt and Turner (2017) for a full description; (c) mean sea-level pressure anomaly [hPa] with overlaid (absolute) 10 m winds from ERA-Interim. Longitude and latitude are relative: the origin is the tracked centre of the composite WD.

This large coverage of cirrus in the north fits existing case-study satellite observations (Rakesh *et al.*, 2009) and has previously been attributed to frontal (Bhaskara Rao and Morey, 1971) and extratropical cyclone-like dynamics (Agnihotri and Singh, 1982). Low- and particularly mid-level clouds feature strongly west of the centre. These are co-located with the region of strong ascent shown in Figure 4(a) and it is likely that they are enhanced upon WD interaction with the north Indian orography (Bhaskara Rao and Morey, 1971).

### 3.2.2.3. Sea-level pressure and winds.

Given in Figure 5(c) are the composite sea-level pressure anomaly and 10 m winds. Consistent with the vertical structure we saw in Figure 4, there is a depression several hundred kilometres east of the centre reaching a little over 1 hPa in magnitude, balanced by an equal-magnitude high-pressure area several hundred kilometres to the west/northwest. This apparently weak surface low is the result of variability in the vertical extent of the WD: if one filters events by vortical intensity (as in Figure 3), so as to remove the weakest 90%, then the magnitude of the depression rises to 4 hPa. 10 m winds in the composite WD are fairly light, reaching an average of  $2 \text{ m s}^{-1}$ ; these are mostly comprised of northerlies flowing from the high-pressure area in the west, although there is some cyclonic component, and indeed weak northerlies associated with the low-pressure anomaly in the east. There are competing drivers for the near-surface winds here: extension of the mid-level vorticity into the boundary layer, forcing centralized, cyclonic winds; or a cyclone–anticyclone pair driven by the surface pressure anomalies.

### 3.3. Evolution

As we have already discussed, the structure of a western disturbance will change as it interacts with northern India, which we can conclude on two counts: the drastically changing and impinging orography and the sudden propensity for extreme precipitation in the Karakoram/Hindu Kush region. Therefore, as part of a complete picture of the climatology of these systems, we should discuss their evolution.

For the meridional structure of cloud cover, this is given in Figure 6; the timeline is defined around the zero-day mark, the time point at which, in the north India box, the WD has maximum 450–300 hPa relative vorticity. Figure 6 shows the meridional structure of the evolution of cloud cover in ERA-Interim over 10 days. Before we discuss this further, we must issue the obvious caveat that cloud in ERA-I is a purely modelled product and thus subject to some uncertainty and model biases. There is a distinct and significant change in cloud cover over the period presented here: initially ( $-5$  days), there is a lower-tropospheric maximum

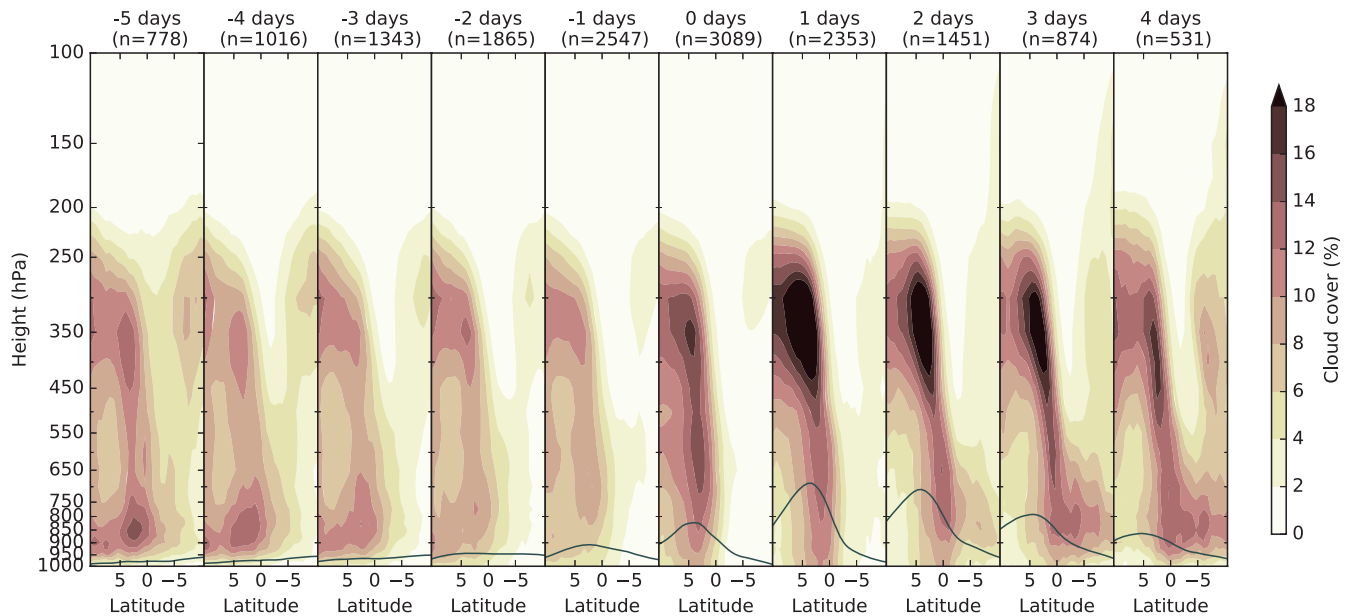
at 850 hPa, as well as a weak maximum aloft at 350 hPa. Over the subsequent 4 days, as the WD approaches peak strength, the cloud generally thins, particularly in the lower levels, although the upper-level maximum is roughly maintained. At peak vortical strength, the cloud pattern starts to shift in accordance with interaction with the Himalayas: the 350 hPa maximum starts to develop while extending to lower altitudes. One day after peak vorticity, the cloud cover reaches a maximum with a now clear vertical extent to the surface approximately 100 km to the north of the centre. This is consistent with the time at which the interaction with the north Indian orography is most crucial, when the maximum rainfall would be likely to occur. The magnitude and extent of the upper-level maximum are maintained for several days before decreasing, during which time the spatial extent of the low-level cloud starts to increase.

For comparison, the evolution of vertical wind speed, given as zonal cross-sections, is shown in Figure 7. Here, the structure is markedly more symmetrical than we saw for cloud cover and appears to remain so for the duration of the event. As the WD intensifies at day 0, the peak ascent drops from 400 to 500 hPa but doubles in magnitude, before decreasing several days thereafter. This sudden intensification over India is indicative of either a strong coupling with the orography or some other source or region of high baroclinicity.

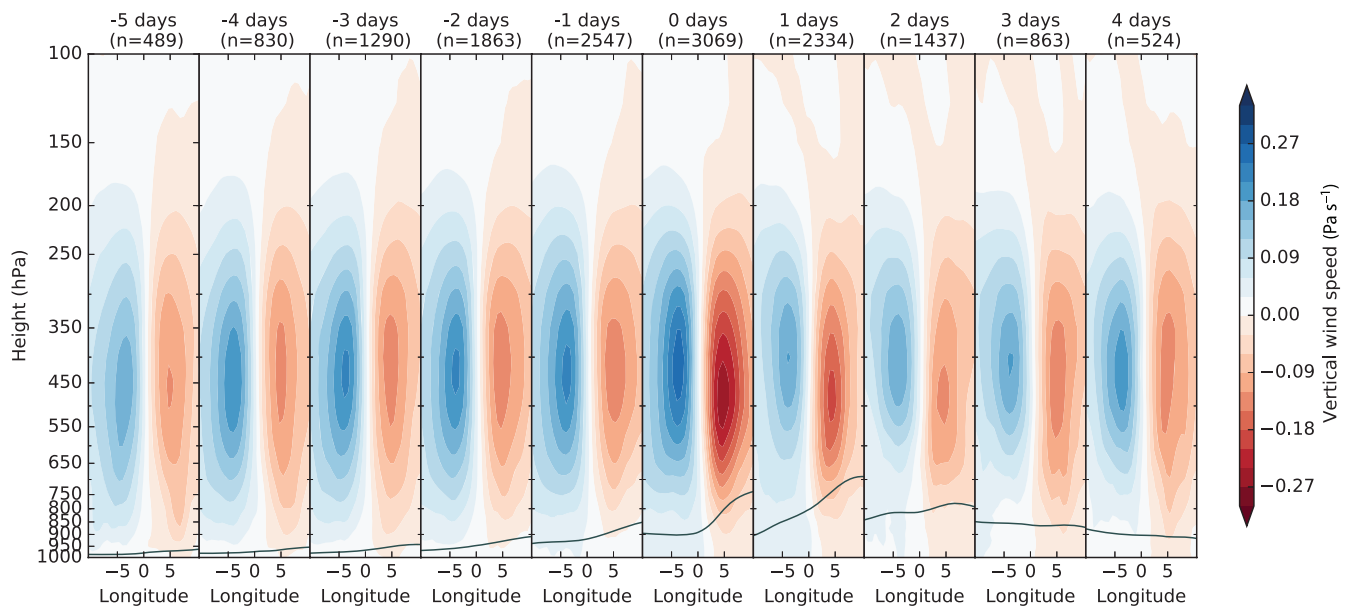
We complete this discussion with a look at the evolution of the divergence field (Figure 8). As before, the general pattern increases in magnitude as it reaches north India, before slowly relaxing again thereafter; here, however, the role of the orography is much clearer. The lower-tropospheric maxima in  $|\nabla \cdot \mathbf{u}|$  are initially below 900 hPa, but are elevated to around 650 hPa and strengthened by topographic interaction as the WD moves over north India and remain elevated until lysis.

### 3.4. Classification and variability

One corollary of the previous absence of a catalogue of western disturbances is that they remain without any form of robust classification. The distinction has been made in earlier studies between precipitating and non-precipitating disturbances, but there has been little work beyond case-study selections into any further classification (Srinivasan, 1971; Veeraraghavan and Nath, 1989). Such work permits the important exploration of the domains of governing dynamics. Here we use the  $k$ -means clustering algorithm outlined in section 2.3 to attempt to classify WDs based on their dynamical features. To do this, we select quasi-orthogonal dynamic fields of importance: relative vorticity, divergence and vertical wind speed; as these each have quite different statistical distributions, we normalize each field by subtracting the mean and dividing by the standard deviation. Then a suitable domain around the WD centre is chosen (here



**Figure 6.** Composite evolution of the meridional cross-section of cloud cover, from ERA-Interim. ‘0 days’ is defined as the time in the north India box (Figure 1(b)) where the 450–300 hPa vorticity is at a maximum. Then, for example, ‘+3 days’ is the composite of all systems (that still exist) 96 h after this point. The size of the population for each time point is given above the relevant subfigure. The dark teal line towards the bottom of each panel indicates the mean height of the orography (as indicated by surface pressure) for that particular composite subset. Figure 2(a) shows the mean location of the WDs comprising each panel.



**Figure 7.** Composite evolution of the zonal cross-section of vertical wind speed, from ERA-Interim. Definitions as in Figure 6.

1000 × 1000 km<sup>2</sup> from the surface to 100 hPa), before regridding each field such that the voxels are of equal volume (thus avoiding overweighting lower altitudes, where the level density is greater). For each time point\*\* the fields are collapsed into one dimension and the three variables (relative vorticity, divergence,  $\omega$ ) are concatenated into a single vector. These vectors then comprise the vector set  $\mathbf{x}$  given in Eq. (1). The algorithm is then run on this set of vectors to obtain the cluster mean vectors, from which the original three-dimensional fields can be reconstructed. As already mentioned, the number of clusters must be decided before running the algorithm; here we employ the heuristic ‘elbow method’ (e.g. Hardy, 1994), determining that four clusters explain a sufficient amount of the variance.

Figure 9 shows the structure of the mean for all three fields in each cluster. The rows give the normalized values of relative vorticity, vertical wind speed and divergence, respectively; the

columns represent the different clusters, in increasing order of maximum relative vorticity. Types 1, 2 and 3 are separated by intensity: type 1, the most common, is the weakest, with the lowest magnitudes of vorticity, divergence and ascent/descent; conversely, type 3 is the least common, but strongest, type of WD. The algorithm has then split type 3 further into long-wavelength (3a) and short-wavelength (3b) categories. Each cluster is well-populated, indicating the absence of any serious outliers, for example cyclonic features that are not WD-like. Data on the extrema that can be computed from each panel of Figure 9 are given in Table 1, showing that the central vorticity increases by about 30% with strengthening type and that the ascent/descent becomes increasingly asymmetrical – from being almost equal in type 1 to the ascent being around three times stronger than the descent in type 3 (and around eight times stronger than the ascent in type 1). Both the peak convergence and divergence increase with strengthening type as well. Lower level convergence is the more dominant of the two for the two weakest types, but is superseded by divergence aloft in type 3 WDs. As we have already mentioned, the key feature differentiating types 3a and

\*\*We are considering all time points where a tracked WD is inside the box given in Figure 1(b).

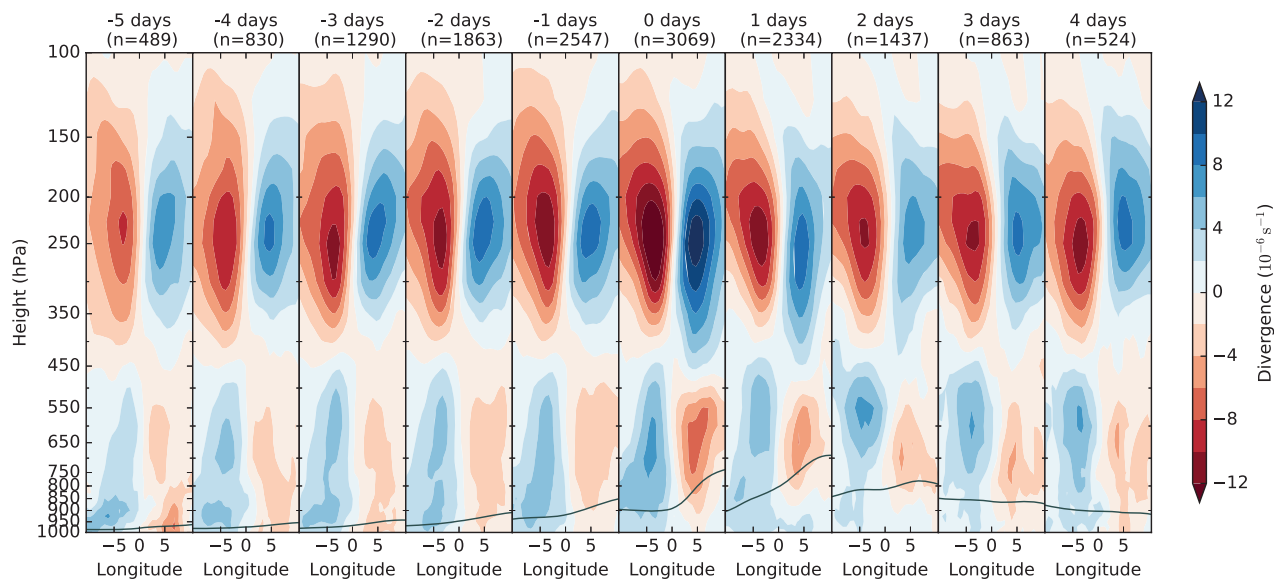


Figure 8. Composite evolution of the zonal cross-section of divergence [ $10^{-6} \text{ s}^{-1}$ ], from ERA-Interim. Definitions as in Figure 6.

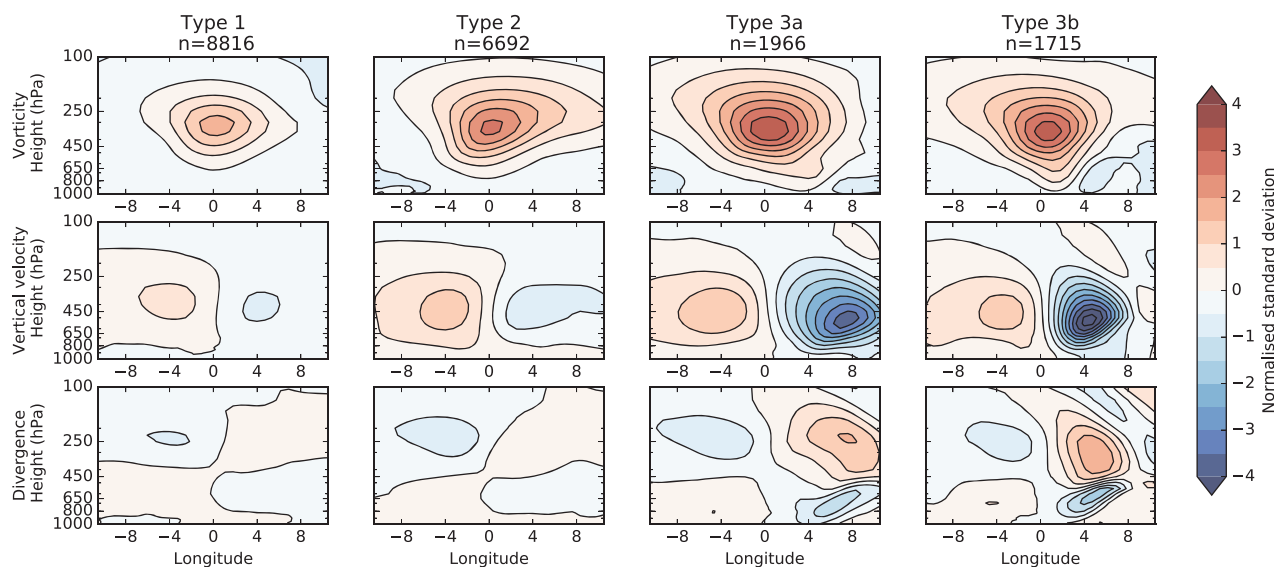


Figure 9. Classification of WDs based on a four-partition  $k$ -means clustering algorithm using normalized relative vorticity, vertical wind speed and divergence. Populations of each cluster are given at the top of each column.

3b is horizontal spatial scale. We can estimate the wavelength of these by doubling the distance between the peak ascent and descent: type 3a has wavelength  $\sim 2300$  km, compared with the more axially confined type 3b (and 1, 2) at  $\sim 1600$  km. In keeping with the method of classifying storms by their wind speed, we have included the maximum magnitude of meridional wind speed for the cluster mean. As expected from the general trends we have already seen across the types, the wind speed increases about  $3 \text{ m s}^{-1}$  across each type. Note that, because of the similar vorticity but shorter wavelength of type 3b WDs, their maximum wind speed is significantly lower than that of a type 3a WD and indeed comparable to type 2. This indicates that it would be difficult, though not impossible, to use wind speed alone to classify WDs accurately. Further, despite the correlation between surface pressure anomaly and upper-level vorticity we mentioned earlier, there is little such relationship with the dynamic types presented here: the mean magnitudes of the surface depression for type 1 and type 3 WDs are 1.0 and 1.4 hPa, respectively.

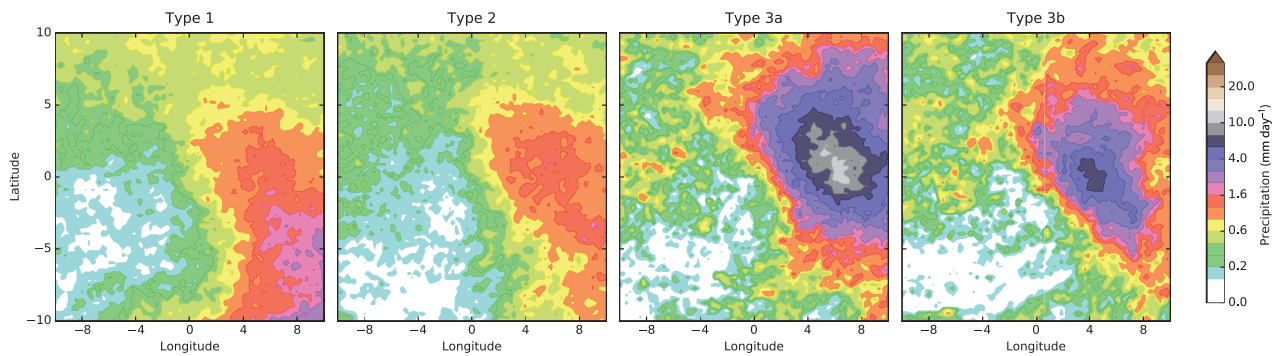
So, how do these dynamical classifications relate to precipitation? Figure 10 shows the mean rainfall for each cluster described above. For type 1 WDs, the precipitation maximum is far to the southwest of the centre, indicating likely orographic interference. For the remaining three types, the composite precipitation maximum is co-located with the

Table 1. Selected data derived from Figure 9, showing extrema of the composite mean for dynamic fields in each cluster type. Populations for each cluster are given above the respective panel.

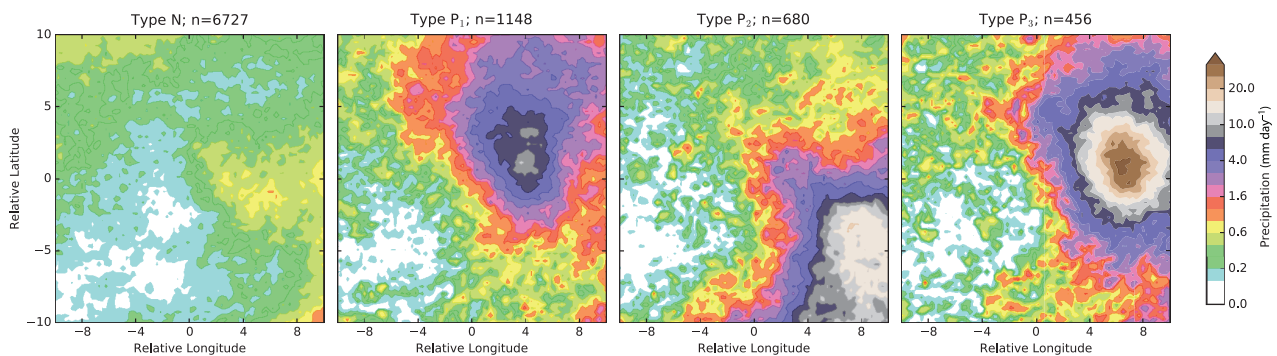
| Type | $\nabla \times \mathbf{u}$ max<br>( $10^{-5} \text{ s}^{-2}$ ) | $\omega$ max min<br>( $\text{Pa s}^{-1}$ ) | $\nabla \cdot \mathbf{u}$ max min<br>( $10^{-5} \text{ s}^{-2}$ ) | $ \mathbf{v} $ max<br>( $\text{m s}^{-1}$ ) |
|------|--|--|---|---|
| 1    | 8.70   | 0.15  -0.10                                | 0.68  -1.01   | 12.1  |
| 2    | 12.1   | 0.27  -0.16                                | 0.92  -1.43   | 15.5  |
| 3a   | 15.4   | 0.29  -0.73                                | 2.98  -2.52   | 20.9  |
| 3b   | 14.8   | 0.26  -0.92                                | 3.70  -3.31   | 15.9  |

maximum composite ascent, ranging from  $4\text{--}8^\circ$  to the west of the centre. The greatest mean rainfall is associated with type 3a WDs, which is perhaps surprising, given that their ascent is typically weaker than in type 3b. Despite this, it is clear that, on average, increased dynamical intensity tends to lead to an increased mean rainfall.

We can extend the idea of automated classification to rainfall. It has been alluded to both previously in this study and in previous literature (e.g. Srinivasan, 1971; Veeraraghavan and Nath, 1989) that WDs can be classified by precipitation. This is an inherently difficult problem, given the high spatial variance and long-tailed distribution of surface precipitation; indeed, running



**Figure 10.** Composite precipitation from *TRMM* 3B42 [ $\text{mm day}^{-1}$ ] for each of the dynamical classifications given in Figure 9.



**Figure 11.** Four-partition  $k$ -means clustering of WD surface precipitation ( $\text{mm day}^{-1}$ ) from *TRMM* 3B42 for the north India box given in Figure 1(b). Resulting clusters are labelled type N (non-precipitating) and type P (precipitating), with type P being subdivided further (nomenclature arbitrary) based on the location/intensity of precipitation.

**Table 2.** Count (with ratio of the total number of events for the type) for events with maximum rainfall, from *TRMM* 3B42, exceeding a given threshold, binned by event precipitating type (see Figure 11).

| Type           | $n(\%) > 65$<br>$\text{mm day}^{-1}$ | $n(\%) > 125$<br>$\text{mm day}^{-1}$ | $n(\%) > 200$<br>$\text{mm day}^{-1}$ | $n(\%) > 500$<br>$\text{mm day}^{-1}$ |
|----------------|--------------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|
| N              | 3461(51.4)                           | 1837(27.3)                            | 815(12.1)                             | 47(0.7)                               |
| P <sub>1</sub> | 1062(92.5)                           | 826(72.0)                             | 515(44.9)                             | 47(4.1)                               |
| P <sub>2</sub> | 671(98.7)                            | 595(87.5)                             | 480(70.6)                             | 107(15.7)                             |
| P <sub>3</sub> | 445(97.6)                            | 405(88.8)                             | 317(69.5)                             | 60(13.2)                              |

The thresholds are 65, 125, 200 and  $500 \text{ mm day}^{-1}$ ; of these, the first three correspond to official designations by the India Meteorological Department: ‘heavy’, ‘very heavy’ and ‘extremely heavy’.

a  $k$ -means clustering algorithm straight on *TRMM* 3B42 data for WDs tended to produce small groups of outliers, regardless of the number of clusters prescribed. To mitigate this and allow larger groups, we replaced precipitation  $P$  with the expression  $\log(1 + P)$ . As with Figure 9, the optimum number of clusters was four and the average for each cluster is given in Figure 11. These four clusters can be split into two types: those with strong evidence of a large-scale precipitation footprint and significant rainfall rate therein (type P) and those that do not (type N). The three type P clusters are separated by intensity and peak position, assigned a further number by descending order of population. The maxima of types P<sub>1</sub> and P<sub>3</sub> are co-located and the spatial structures are very similar, though the magnitude of the latter is some four times greater than the former. These precipitation footprints are very similar to that one would expect on consideration of the surface pressure (Figure 5(c)) and ascent (Figure 4(a)). The remaining type, P<sub>2</sub>, sits between the other two precipitating types in terms of magnitude, but the maximum is located far to the southwest rather than being to the near-west of the centre.

Unfortunately, it would be naive to assume that extreme rainfall events (especially those of smaller spatial scale) are necessarily confined to type P western disturbances, since the  $k$ -means clustering favours broad spatial similarities, which such events lack. To explore whether there is any meaningful correlation

between individual extreme events and these broader spatially derived categories, we can simply bin and count the maxima for each event in each type. Table 2 summarizes these data at various thresholds: 65, 125, 200 and  $500 \text{ mm day}^{-1}$ , the first three of which correspond to heavy, very heavy and extremely heavy rainfall, respectively, in IMD reports. We can see that the two metrics are strongly related: type P WDs are significantly more likely to have heavy rainfall events at any threshold than type N and rainfall events at the two highest thresholds are altogether more likely to be associated with type P WDs, despite the greater number of type N events. Of particular note are the rainfall events exceeding  $500 \text{ mm day}^{-1}$ ,<sup>††</sup> which tend to be associated with type P<sub>2</sub> WDs that have their precipitation maximum some 1000 km to the southwest of the WD centre. Rainfall events of this magnitude are almost exclusively associated with orography (Rajeevan *et al.*, 2008), implying that type P<sub>2</sub> is most likely to be interacting constructively with the north Indian topography to generate precipitation. Despite this, we note that it has been shown (Sharma and Subramaniam, 1983) that an upper-level trough over the peninsula can enhance an incident WD, resulting in a greater southward extent of the precipitation.

So, what is the relationship between these precipitation-based categories and the dynamical structure aloft? We have noted that two of the precipitating categories, type P<sub>1</sub> and type P<sub>3</sub>, have footprints that appear as though they are linked directly to the WD structure, whereas the other type P<sub>2</sub> does not. A simple Bayesian analysis comparing the dynamical categories and precipitation categories<sup>‡‡</sup> reveals that there are indeed strong relationships. Relatively speaking, type N is more likely to be associated with types 1 and 2 (by about 10%); type P<sub>1</sub> with type 3 (80%); type P<sub>2</sub> with type 1 (31%); and type P<sub>3</sub> with type 3 (315% for 3a, 31% for 3b). We can thus reasonably confirm our earlier hypothesis that

<sup>††</sup> It is probably more appropriate to refer to this quantity as  $20 \text{ mm h}^{-1}$ , as it is an instantaneous rate rather than a daily accumulation.

<sup>‡‡</sup> Consider that the two sets of categories were completely unrelated: we would then not expect the normalized frequency distributions of  $P(\{1,2,3a,3b\}|N)$  and  $P(\{1,2,3a,3b\}|P_1)$  to be significantly different, for example.

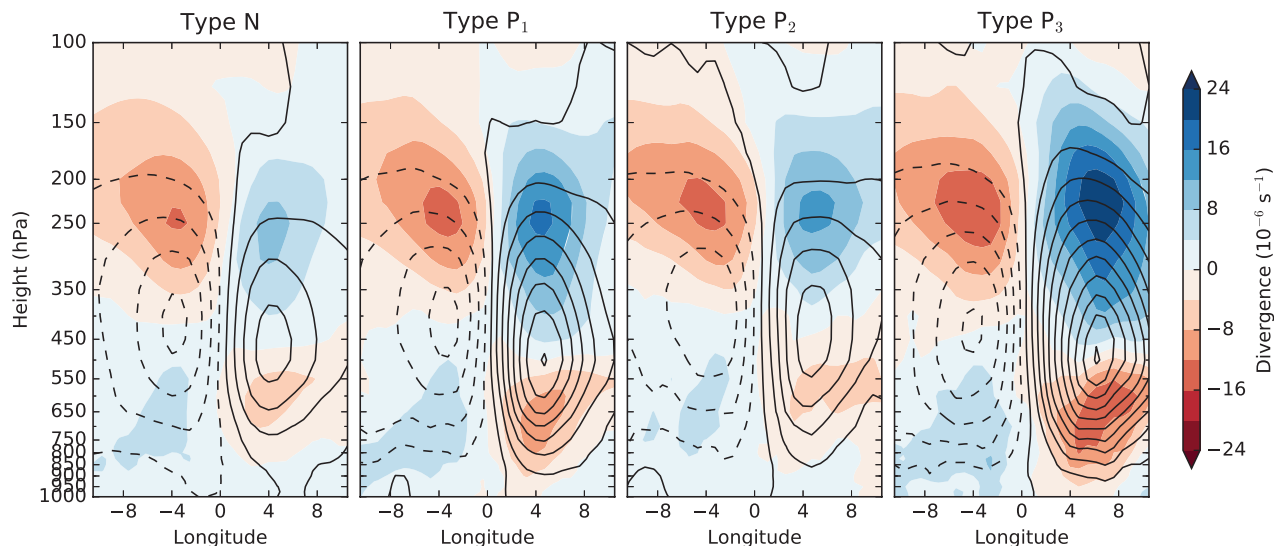


Figure 12. Divergence [shaded,  $10^{-6} \text{ s}^{-1}$ ] and vertical wind speed [contours, interval:  $0.05 \text{ Pa s}^{-1}$ , ascent solid] for the precipitation categories outlined in Figure 11.

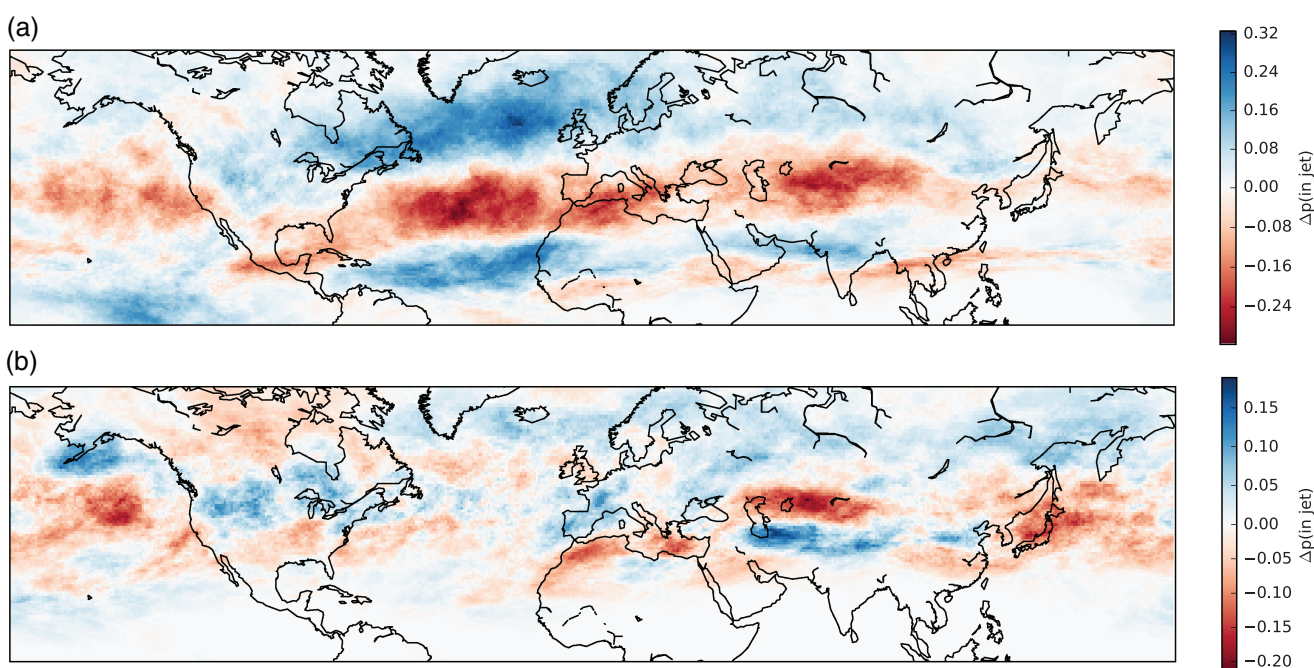


Figure 13. Difference in westerly jet location between years of high and low WD frequency for each of two seasons: (a) January–March and (b) June–September. A given point is considered to be in a westerly jet if, at 200 hPa,  $u > 0$  and  $u > 30 \text{ m s}^{-1}$ , computed from six-hourly ERA-Interim data. The mean location probability function is computed for each season and the change in probability is calculated between the selected years (given in the text). Note that the colour scales differ between the two subfigures.

types  $P_1$  and  $P_3$  are strongly coupled with the dynamical structure of WDs and that type  $P_2$  is not. To confirm this, we show the vertical structure of  $\omega$  and divergence for WDs across the four precipitation categories in Figure 12. As expected, the strongest ascent and convergence are found in type  $P_3$  WDs, then in type  $P_1$ , with the other two categories being the weakest. Note that these changes are only notable for dynamic variables; repeating this analysis with temperature and humidity does not reveal significantly different structures across the precipitation categories.

### 3.5. Interaction with the jet

Given the coupled nature of western disturbances and the subtropical westerly jet, we might expect there to be some correlation between the jet location and the frequency of WDs incident on north India. Following other authors (e.g. Schiemann *et al.*, 2009), we define a point as being in a westerly jet if, at 200 hPa, the following two conditions are met: the wind is westerly (i.e.  $u > 0$ ) and the magnitude of its speed is greater than  $30 \text{ m s}^{-1}$  (i.e.  $u > 30 \text{ m s}^{-1}$ ). This permits us to compute,

over a finite set of times, the probability that a given point will be in a westerly jet. We shall call this variable  $p(\text{in jet})$ . Noting the intra-annual variability in the latitude of the STWJ, it is sensible to consider this relationship in the context of different seasons; thus, we shall look at both boreal summer (June–September) and boreal winter (January–March). For each season, the four years with the highest (winter: 1979, 1995, 2012, 2013; summer: 1983, 1989, 2000, 2010) and lowest (winter: 1985, 1987, 2004, 2010; summer: 1990, 1992, 1995, 1997) WD frequency were considered. In Figure 13, the difference in  $p(\text{in jet})$  between the given high and low years for each season is shown. For the winter (Figure 13(a)), the spatial extent of the difference is considerable: both the probability increase over central India and the corresponding decrease over central Asia have a significant, coherent signal as far west as continental America. Conversely, in the summer case, the difference is much more meridionally confined and extends only as far west as the Caspian Sea. This implies that externally forced changes in the jet position, whether by modes of climate variability such as ENSO or planetary warming, may result in markedly different distributions of WDs.

#### 4. Conclusion

Starting from the description given by the India Meteorological Department and a selection of case studies, we developed and applied a tracking algorithm to 36 years of global reanalysis data; applying a location filter over northern India and Pakistan, we recovered over 3000 upper-level disturbances in the subtropical westerly jet stream (STWJ) that fit previous definitions of western disturbances (WDs). This frequency of 6–7 per month through northern India is close to previous estimates and substantially closer than previous efforts to track WDs (Cannon *et al.*, 2016). These disturbances have substantial variance in their genesis locations, ranging from the Arctic Circle to the Caribbean, but most commonly originate over Arabia, north Africa and the Mediterranean. We have shown that these disturbances impact India most frequently at the beginning of February and least frequently at the beginning of August, with the relative frequency of stronger WDs increasing during the winter months.

The vertical structure of these WDs was then explored using a composite of events over northern India. Many fields clearly revealed the highly baroclinic nature of these disturbances, since many features lean northwestward with height, including the PV maximum embedded in the upper troposphere. Aside from the evident nonlinear coupling with the STWJ, WDs typically look similar to immature extratropical cyclones and include a warm, dry core aloft over a cold, moist core in the middle and lower troposphere. Another key feature differentiating WDs from other synoptic features of the region is the markedly off-centre ascent, the maximum of which is elevated, at 450 hPa and some 600 km to the west, i.e. ahead of, the disturbance. Most precipitation associated with the composite WD is found to be associated with this large-scale ascent.

*k*-means clustering was used to classify WDs over north India according to both a combined dynamic field (using normalized vorticity, divergence and vertical wind speed) and precipitation. In the first instance, four distinct categories were found: three based on intensity, with the third splitting into two based on spatial scale – one short-wavelength (~1600 km) and one long (~2300 km). In the second instance, WDs were broadly split into two categories: those that have large-scale precipitation and those that do not. The former category can then be split into three new categories that are functions of rainfall intensity and location. Here, we showed that particularly extreme rainfall events are more likely to be part of a WD in which large-scale precipitation is found, rather than as a small isolated storm in a weakly precipitating WD. Further, we showed that there is a strong relationship between the dynamical categories and the precipitation categories, indicating that rainfall associated with WDs is related not just to the orography but also to the intensity of the system itself. The mechanism behind such extreme events remains unclear. Remaining questions include the following: what fraction of extreme rainfall events in these vulnerable areas of north India and Pakistan are western disturbances actually responsible for? What is the synoptic structure of the atmosphere during events for which WDs are not responsible?

Finally, we showed that the location/presence of the STWJ plays a crucial influence in the frequency of WDs in north India and Pakistan. We showed that this was true regardless of the season, but that the spatial coherence of the pattern is much higher during boreal winter, with the signal reaching as far west as America. Bearing in mind the relationship between western disturbances and flooding in Pakistan and north India, it is important to ask how the behaviour of the jet will vary in a changing climate; further, given that WDs are largely baroclinic phenomena, how will the baroclinicity of the upper atmosphere change over Eurasia in such scenarios?

Further work will first assess the validity of WD representation in a range of climate models before examining how jet structure, WD structure and count might vary in idealized climate scenarios.

#### Acknowledgement

KMRH, AGT and LCS are funded by the JPI-Climate and Belmont Forum Climate Predictability and Inter-Regional Linkages Collaborative Research Action via NERC grant NE/P006795/1.

#### References

- Agnihotri CL, Singh MS. 1982. Satellite study of western disturbances. *Mausam* **33**: 249–254.
- Azadi M, Mohanty UC, Madan OP, Padmanabhamurty B. 2002. Prediction of precipitation associated with a western disturbance using a high-resolution regional model: Role of parameterisation of physical processes. *Meteorol. Appl.* **9**: 317–326.
- Bhaskara Rao NS, Morey PE. 1971. Cloud systems associated with western disturbances: A preliminary study. *Indian J. Meteorol. Geophys.* **22**: 413.
- Boettcher M, Wernli H. 2011. Life cycle study of a diabatic Rossby wave as a precursor to rapid cyclogenesis in the North Atlantic – dynamics and forecast performance. *Mon. Weather Rev.* **139**: 1861–1878. <https://doi.org/10.1175/2011MWR3504.1>.
- Brahmananda Rao V, Trivikrama Rao S. 1971. A theoretical and synoptic study of western disturbances. *Pure Appl. Geophys.* **90**: 193–208.
- Cannon F, Carvalho LMV, Jones C, Norris J. 2016. Winter westerly disturbance dynamics and precipitation in the western Himalaya and Karakoram: A wave-tracking approach. *Theor. Appl. Clim.* **125**: 27–44.
- Chakravorty KC, Basu SC. 1957. The influence of western disturbances on the weather over northeast India in monsoon months. *Indian J. Meteorol. Geophys.* **8**: 261–272.
- Chattopadhyay J. 1970. Power spectrum analysis of atmospheric ozone content over north India. *Pure Appl. Geophys.* **83**: 111–119.
- Chitlangia PR. 1976. Mean model of western depression. *Indian J. Meteorol. Geophys.* **87**: 157–162.
- Dacre HF, Hawcroft MK, Stringer MA, Hodges KI. 2012. An extratropical cyclone atlas: A tool for illustrating cyclone structure and evolution characteristics. *Bull. Am. Meteorol. Soc.* **93**: 1497–1502. <https://doi.org/10.1175/BAMS-D-11-00164.1>.
- Davis CA. 1992. A potential-vorticity diagnosis of the importance of initial structure and condensational heating in observed extratropical cyclogenesis. *Mon. Weather Rev.* **120**: 2409–2428.
- Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen I, Kållberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavolato C, Thépaut JN, Vitart F. 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **137**: 553–597. <https://doi.org/10.1002/qj.828>.
- Dimri AP. 2004. Impact of horizontal model resolution and orography on the simulation of a western disturbance and its associated precipitation. *Meteorol. Appl.* **11**: 115–127.
- Dimri AP. 2006. Surface and upper air fields during extreme winter precipitation over the western Himalayas. *Pure Appl. Geophys.* **163**: 1679–1698.
- Dimri AP, Chevuturi A. 2014. Model sensitivity analysis study for western disturbances over the Himalayas. *Meteorol. Atmos. Phys.* **123**: 155–180.
- Dimri AP, Niyogi D, Barros AP, Ridley J, Mohanty UC, Yasunari T, Sikka DR. 2015. Western disturbances: A review. *Rev. Geophys.* **53**: 225–246.
- Dimri AP, Yasunari T, Kotlia BS, Mohanty UC, Sikka DR. 2016. Indian winter monsoon: Present and past. *Earth Sci. Rev.* **163**: 297–322.
- Dinku T, Connor SJ, Ceccato P. 2010. Comparison of CMORPH and TRMM-3B42 over mountainous regions of Africa and South America. In *Satellite Rainfall Applications for Surface Hydrology*, Dinku T, Hossain F. (eds.): 193–204. Springer: Dordrecht, the Netherlands. [https://doi.org/10.1007/978-90-481-2915-7\\_11](https://doi.org/10.1007/978-90-481-2915-7_11).
- Emanuel K. 2005. Genesis and maintenance of ‘Mediterranean hurricanes’. *Adv. Geosci.* **2**: 217–220.
- Guo H, Chen S, Bao A, Hu J, Gebregiorgis AS, Xue X, Zhang X. 2015. Inter-comparison of high-resolution satellite precipitation products over Central Asia. *Remote Sens.* **7**: 7181–7211.
- Gupta A, Rathore LS, Singh SV, Mendiratta N. 1999. Performance of a global circulation model in predicting the winter systems and associated precipitation over North West India during 1994–1997. In *The Himalayan Environment*, Dash SK, Bahadur J, Tremblay LB. (eds.): 123–138. New Age International Publishers: New Delhi.
- Hara M, Kimura F, Yasunari T. 2004. ‘The generation mechanism of the western disturbances over the Himalayas’. In *6th International GAME Conference*, 3–5 December 2004, Kyoto, Japan.
- Hardy A. 1994. An examination of procedures for determining the number of clusters in a data set. In *New Approaches in Classification and Data Analysis*, Diday E, Lechevallier Y, Schader M, Bertrand P, Burtschy B. (eds.): 178–185. Studies in Classification, Data Analysis, and Knowledge Organization. Springer: Berlin, Heidelberg.
- Hartigan JA, Wong MA. 1979. Algorithm AS 136: A *k*-means clustering algorithm. *J. R. Stat. Soc. C-Appl.* **28**: 100–108.

- Huffman GJ, Bolvin DT, Nelkin EJ, Wolff DB, Adler RF, Gu G, Hong Y, Bowman KP, Stocker EF. 2007. The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.* **8**: 38–55. <https://doi.org/10.1175/JHM560.1>.
- Huffman GJ, Adler RF, Bolvin DT, Nelkin EJ. 2010. The TRMM multisatellite precipitation analysis (TMPA). In *Satellite Rainfall Applications for Surface Hydrology*, Gebremichael M, Hossain F. (eds.): 3–22. Satellite Rainfall Applications for Surface Hydrology. Springer: Dordrecht. [https://doi.org/10.1007/978-90-481-2915-7\\_1](https://doi.org/10.1007/978-90-481-2915-7_1).
- Hunt KMR, Turner AG. 2017. The representation of Indian monsoon depressions at different horizontal resolutions in the Met Office Unified Model. *J. Clim.* **143**: 1756–1771. <https://doi.org/10.1002/qj.3030>.
- Hunt KMR, Turner AG, Inness PM, Parker DE, Levine RC. 2016. On the structure and dynamics of Indian monsoon depressions. *Mon. Weather Rev.* **144**: 3391–3416. <https://doi.org/10.1175/MWR-D-15-0138.1>.
- Hurley JV, Boos WR. 2015. A global climatology of monsoon low pressure systems. *Q. J. R. Meteorol. Soc.* **141**: 1049–1064. <https://doi.org/10.1002/qj.2447>.
- Kotal SD, Roy SS, Roy Bhowmik SK. 2014. Catastrophic heavy rainfall episode over Uttarakhand during 16–18 June 2013 – observational aspects. *Curr. Sci.* **107**: 234–245.
- Kulkarni A, Kripalani RH, Singh SV. 1992. Classification of summer monsoon rainfall patterns over India. *Int. J. Climatol.* **12**: 269–280.
- Kummerow C, Barnes W, Kozu T, Shiue J, Simpson J. 1998. The Tropical Rainfall Measuring Mission (TRMM) sensor package. *J. Atmos. Oceanic Technol.* **15**: 809–817. [https://doi.org/10.1175/1520-0426\(1998\)015<0809:ttrmmt>2.0.CO;2](https://doi.org/10.1175/1520-0426(1998)015<0809:ttrmmt>2.0.CO;2).
- Lang TJ, Barros AP. 2004. Winter storms in the central Himalayas. *J. Meteorol. Soc. Jpn.* **82**: 829–844.
- Madhura RK, Krishnan R, Revadekar JV, Mujumdar M, Goswami BN. 2015. Changes in western disturbances over the western Himalayas in a warming environment. *Clim. Dyn.* **44**: 1157–1168.
- Mohanty UC, Madan OP, Rao PLS, Raju PVS. 1998. ‘Meteorological fields associated with western disturbances in relation to glacier basins of western Himalayas during winter season’, Technical report. Center for Atmospheric Sciences, Indian Institute of Technology: New Delhi.
- Mohanty UC, Madan OP, Raju PVS, Bhatla R, Rao PLS. 1999. A study on certain dynamic and thermodynamic aspects associated with western disturbances over north-west Himalaya. In *The Himalayan Environment*, Dash SK, Bahadur J. (eds.): 113–122. New Age International Pvt. Ltd: New Delhi, India.
- Molinari J, Skubis S, Vollaro D. 1995. External influences on hurricane intensity. Part III: Potential vorticity structure. *J. Atmos. Sci.* **52**: 3593–3606.
- Mooley DA. 1957. The role of western disturbances in the production of weather over India during different seasons. *Indian J. Meteorol. Geophys.* **8**: 253–260.
- Nair S, Srinivasan G, Nemani R. 2009. Evaluation of multi-satellite TRMM derived rainfall estimates over a western state of India. **87**: 927–939. <https://doi.org/10.2151/jmsj.87.927>.
- Palazzi E, Hardenberg J, Provenzale A. 2013. Precipitation in the Hindu-Kush Karakoram Himalaya: Observations and future scenarios. *J. Geophys. Res. Atmos.* **118**: 85–100. <https://doi.org/10.1029/2012JD018697>.
- Patil R, Kumar PP. 2016. WRF model sensitivity for simulating intense western disturbances over North West India. *Model. Earth Syst. Env.* **2**: 1–15.
- Rajeevan M, Bhat J, Jaswal AK. 2008. Analysis of variability and trends of extreme rainfall events over India using 104 years of gridded daily rainfall data. *Geophys. Res. Lett.* **35**: L18707. <https://doi.org/10.1029/2008GL035143>.
- Rakesh V, Singh R, Yuliya D, Pal PK, Joshi PC. 2009. Impact of variational assimilation of MODIS thermodynamic profiles in the simulation of western disturbance. *Int. J. Remote Sens.* **30**: 4867–4887.
- Ramanathan V, Saha KR. 1972. Application of a primitive equation barotropic model to predict movement of ‘Western Disturbances’. *J. Appl. Meteorol.* **11**: 268–272.
- Rangachary N, Bandyopadhyay BK. 1987. An analysis of the synoptic weather pattern associated with extensive avalanching in Western Himalaya. *Int. Assoc. Hydrol. Sci. Publ.* **162**: 311–316.
- Robinson WA. 1989. On the structure of potential vorticity in baroclinic instability. *Tellus A* **41**: 275–284.
- Roy SS, Roy Bhowmik SK. 2005. Analysis of thermodynamics of the atmosphere over northwest India during the passage of a western disturbance as revealed by model analysis field. *Curr. Sci.* **88**: 947–951.
- Saha S, Moorthi S, Pan HL, Wu X, Wang J, Nadiga S, Tripp P, Kistler R, Woollen J, Behringer D, Liu H, Stokes D, Grumbine R, Gayno G, Wang J, Hou Y-T, Chuang H, Juang H-MH, Sela J, Iredell M, Treadon R, Kleist D, Van Delst P, Keyser D, Derber J, Ek M, Meng J, Wei H, Yang R, Lord S, van den Dool H, Kumar A, Wang W, Long C, Chelliah M, Xue Y, Huang B, Schemm J-K, Ebisuzaki W, Lin R, Xie P, Chen M, Zhou S, Higgins W, Zou C-Z, Liu Q, Chen Y, Han Y, Cucurull L, Reynolds RW, Rutledge G, Goldberg M. 2010. The NCEP climate forecast system reanalysis. *Bull. Am. Meteorol. Soc.* **91**: 1015–1057.
- Schiemann R, Lüthi D, Schär C. 2009. Seasonality and interannual variability of the westerly jet in the Tibetan Plateau region. *J. Clim.* **22**: 2940–2957.
- Semwal G, Dimri AP. 2012. Impact of initial and boundary conditions on simulations of western disturbances and associated precipitation. *Nat. Hazards* **64**: 1405–1424.
- Sharma R, Subramaniam D. 1983. The western disturbance of 22 December 1980: A case study. *Mausam* **34**: 117–120.
- Simpson J, Kummerow C, Tao WK, Adler RF. 1996. On the Tropical Rainfall Measuring Mission (TRMM). *Meteorol. Atmos. Phys.* **60**: 19–36. <https://doi.org/10.1007/BF01029783>.
- Singh MS, Agnihotri CL. 1977. Baroclinity over India in winter and its relation to western disturbances and jet streams. I. *Indian J. Meteorol. Hydro. Geophys.* **28**: 303–310.
- Srinivasan V. 1971. Some case studies of cirriform clouds over India during the winter period. *Indian J. Meteorol. Geophys.* **22**: 421–428.
- Veeraraghavan K, Nath T. 1989. A satellite study of an active western disturbance. *Mausam* **40**: 303–306.
- Yatagai A, Arakawa O, Kamiguchi K, Kawamoto H, Nodzu MI, Hamada A. 2009. A 44-year daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Sola* **5**: 137–140.
- Yatagai A, Kamiguchi K, Arakawa O, Hamada A, Yasutomi N, Kitoh A. 2012. APHRODITE: Constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bull. Am. Meteorol. Soc.* **93**: 1401–1415. <https://doi.org/10.1175/BAMS-D-11-00122.1>.
- Yianilos PN. 1993. Data structures and algorithms for nearest neighbor search in general metric spaces. *SODA* **93**: 311–321.
- Zulkafli Z, Buytaert W, Onof C, Manz B, Tarnavsky E, Lavado W, Guyot JL. 2014. A comparative performance analysis of TRMM 3B42 (TMPA) versions 6 and 7 for hydrological applications over Andean-Amazon river basins. *J. Hydrometeorol.* **15**: 581–592.