A Genesis Index for Monsoon Disturbances

SARAH D. DITCHEK[a] AND WILLIAM R. BOOS

Yale University, New Haven, Connecticut

SUZANA J. CAMARGO

Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York

MICHAEL K. TIPPETT

Department of Applied Physics and Applied Mathematics, Columbia University, New York, New York, and Center of Excellence for Climate Research, Department of Meteorology, King Abdulaziz University, Jeddah, Saudi Arabia

ABSTRACT

Synoptic-scale monsoon disturbances produce the majority of continental rainfall in the monsoon regions of South Asia and Australia, yet there is little understanding of the conditions that foster development of these low pressure systems. Here a genesis index is used to associate monsoon disturbance genesis in a global domain with monthly mean, climatological environmental variables. This monsoon disturbance genesis index (MDGI) is based on four objectively selected variables: total column water vapor, low-level absolute vorticity, an approximate measure of convective available potential energy, and midtropospheric relative humidity. A Poisson regression is used to estimate the index coefficients. Unlike existing tropical cyclone genesis indices, the MDGI is defined over both land and ocean, consistent with the fact that monsoon disturbance genesis can occur over land. The index coefficients change little from their global values when estimated separately for the Asian–Australian monsoon region or the Indian monsoon region, suggesting that the conditions favorable for monsoon disturbance genesis, and perhaps the dynamics of genesis itself, are common across multiple monsoon regions. Vertical wind shear is found to be a useful predictor in some regional subdomains; although previous studies suggested that baroclinicity may foster monsoon disturbance genesis, here genesis frequency is shown to be reduced in regions of strong climatological vertical shear. The coefficients of the MDGI suggest that monsoon disturbance genesis is fostered by humid, convectively unstable environments that are rich in vorticity. Similarities with indices used to describe the distribution of tropical cyclone genesis are discussed.

1. Introduction

At least half of the total seasonal precipitation in most continental monsoon regions is estimated to be produced by synoptic-scale disturbances embedded within the larger-scale monsoon circulation (Hurley and Boos 2015). These systems deliver rainfall that is essential for the lives of billions of people, but also have the potential to cause catastrophic hydrological extremes (e.g., Ajayamohan et al. 2010). For example, thousands of people died and tens of thousands were displaced in 2013 because of floods in northern India associated with a monsoon disturbance that migrated inland from the Bay of Bengal (e.g., Joseph et al. 2015), and in 2010 a monsoon disturbance produced floods that killed over 1700 people and submerged 20% of the surface area of Pakistan (e.g., Aon Benfield 2010; Webster et al. 2011). These are but two of many examples of catastrophic floods produced by monsoon disturbances in South Asia and other monsoon regions.

Although monsoon disturbances have been studied most intensively in the Indian region, transient synoptic-scale disturbances are common in all monsoon regions. Such

[a] Current affiliation: Department of Atmospheric and Environmental Sciences, University at Albany, State University of New York, Albany, New York.

Corresponding author address: Sarah D. Ditchek, Department of Atmospheric and Environmental Sciences, University at Albany, State University of New York, Earth Science 341, 1400 Washington Ave., Albany, NY 12222.
E-mail: sarahditchek@gmail.com

propagating systems are commonly referred to as "monsoon low pressure systems" in the Indian monsoon; here we use the term "monsoon disturbance" for brevity. Each year, about 13–14 such disturbances are observed during June–September in the Indian monsoon region (Yoon and Chen 2005; Krishnamurthy and Ajayamohan 2010). It is estimated that they produce about half of continental India's summer rainfall (Saha et al. 1981; Yoon and Chen 2005; Hurley and Boos 2015). In the Australian summer monsoon, an average of 10–15 monsoon disturbances occur each month (Berry et al. 2012), and a roughly similar count is observed over the west Pacific during boreal summer (Hurley and Boos 2015).

Although regional differences surely exist in monsoon disturbance structure and dynamics, Hurley and Boos (2015) showed that monsoon disturbances in India, Australia, and the west Pacific all share a common structure consisting of a warm-over-cold core, a top-heavy column of potential vorticity extending from the surface to the upper troposphere, and outer wind diameters of about 2000 km. Monsoon disturbances are abundant in Africa, where they are called African easterly waves. Precipitating African easterly waves are typically larger (by about 50%) and have much stronger vertical tilts than Indian monsoon depressions (e.g., Kiladis et al. 2006). Shallow and dry monsoon disturbances occur frequently over the desert regions of northwestern Africa and southwestern Australia and are highly distinct from the deep vortices that occur in the moist parts of the Indian and Australian monsoons (Hurley and Boos 2015).

Despite the importance of monsoon disturbances, their formation mechanisms are poorly understood. Some dynamical studies have focused on the mechanisms of growth and propagation of an intense subset of Indian monsoon disturbances called "monsoon depressions." The India Meteorological Department currently classifies a monsoon disturbance as a monsoon depression if it has sustained surface wind speeds exceeding $8.5 \, \mathrm{m \, s^{-1}}$ (they apply this criterion over ocean) or a central surface pressure anomaly with a magnitude of at least 4 hPa (they apply this criterion over land; Saha et al. 1981; Chen and Weng 1999; Sikka 2006; Ajayamohan et al. 2010).[1] Indian monsoon depressions

form in regions with substantial horizontal shear of the zonal wind, and several studies accordingly have suggested that these storms may grow by barotropic instability (Shukla 1977; Goswami et al. 1980; Lindzen et al. 1983). Other studies have argued that some sort of baroclinic instability or nonmodal growth process, modified by moist convection, is required to explain the observed structure and growth rates of monsoon depressions (Krishnamurti et al. 1983; Farrell 1985; Krishnamurti and Gadgil 1985). Monsoons occur in regions of strong meridional temperature gradients, so it seems plausible that monsoon depressions might derive energy from this baroclinicity. Yet the theory of moist baroclinic instability in a basic state with easterly vertical shear has only been explored in highly idealized linear models (e.g., Moorthi and Arakawa 1985) and recent observational analysis indicates it cannot explain the spinup of Indian monsoon depressions (Cohen and Boos 2016).

Before embarking on further detailed theoretical studies of growth mechanisms for monsoon depressions and, more generally, for monsoon disturbances, it seems natural to ask what can be learned by examining the association of monsoon disturbance occurrence with the environment in which they occur. For example, does the frequency of monsoon disturbance formation generally increase or decrease as vertical wind shear increases? We know of no studies that quantitatively answer the seemingly simple question of how monsoon disturbance frequency is associated with vertical wind shear, low-level vorticity, sea surface temperature (SST), or other properties of their environment. Sikka (1977) noted that Indian monsoon depressions form in a region of high SST and high low-level vorticity, but this falls short of a quantitative assessment of the statistical association of monsoon disturbance formation with environmental variables. Indeed, until the recent compilation of monsoon disturbance track datasets for Australia (Berry et al. 2012) and for the global monsoon domain (Hurley and Boos 2015), such an analysis would have been impossible outside of the Indian monsoon region.

Here we examine the statistical association of monsoon disturbance genesis with the climatological mean state. In particular, we construct an index that relates the likelihood of monsoon disturbance genesis to monthly climatologies of a number of environmental parameters. This sort of exercise has been conducted numerous times for tropical cyclones (e.g., Gray 1979) and has enhanced understanding of the dependence of tropical cyclone occurrence on the El Niño–Southern Oscillation (ENSO; Camargo et al. 2007; Bell et al. 2014). The same approach has also been applied to U.S. tornado and hail occurrence and association with ENSO (Tippett et al.

---

[1] There is no obvious reason to suspect that monsoon disturbance dynamics change qualitatively beyond this intensity threshold. Unlike mature tropical cyclones, which have been argued to result from a finite-amplitude instability involving wind-induced ocean surface evaporation (Emanuel 1991), monsoon disturbances routinely propagate from ocean to land without loss of intensity and so cannot rely on ocean heat content for their primary energy source (e.g., Sikka 2006; Hurley and Boos 2015).

2012; Allen et al. 2015a,b). We use the Poisson regression methodology presented by Tippett et al. (2011), which allows for objective assessment of the ability of different sets of variables to explain the observed distribution of genesis points. This approach contrasts with the more subjective derivation employed in construction of other genesis index formulas (e.g., Emanuel and Nolan 2004).

By constructing a genesis index for monsoon disturbances, we aim to determine whether the geographic distribution and seasonal cycle of monsoon disturbance occurrence can be explained by properties of the mean monsoon state. For example, can the mean state explain why the genesis of Indian monsoon disturbances occurs most frequently over the northern and western edge of the Bay of Bengal (e.g., Krishnamurti et al. 1977; Boos et al. 2015)? We use a global domain for our analysis, which includes synoptic-scale monsoon disturbances that form during local summer in all major monsoon regions, and also assess whether there are regional differences in the statistical association of monsoon disturbance genesis and mean environment.

Interpreting statistical associations between monsoon disturbance genesis and the climatological mean state is complicated by the fact that monsoon disturbance frequency is large enough to alter the mean state. This is likely a greater issue for monsoon disturbances than for tropical cyclones because of the greater frequency of the former; the large contribution of monsoon disturbances to total continental rainfall in many monsoon regions indicates their great influence on the mean state. This problem is difficult to surmount in a statistical analysis of observations, although some approaches have used instantaneous or time-smoothed data at the time of storm genesis (e.g., DeMaria et al. 2001) instead of climatological mean quantities. We considered using this approach, but opted to first answer the simpler question of how monsoon disturbances relate to properties of a monthly climatology. Similar analyses with climatological mean variables have been conducted for tropical cyclones (e.g., Camargo et al. 2007; Tippett et al. 2011) but, until now, not for monsoon disturbances.

The paper is organized as follows. Section 2 describes the methodology and data sources used in this work and is followed by presentation of the monsoon disturbance genesis index (MDGI) and associated results for the global domain. A subsequent section explores the regional sensitivity of the regression and a comparison to a regression with tropical cyclone genesis points. The paper closes with a summary and a discussion of how this work advances understanding of monsoon disturbances.

2. Data and methods

a. Storm genesis points

Genesis points for monsoon disturbances were obtained from the track climatology compiled by Hurley and Boos (2015), hereafter referred to as the Yale dataset (http://earth.geology.yale.edu/depressdata/). This dataset was compiled using the automated feature tracking algorithm developed by Hodges (1995) to identify propagating 850-hPa cyclonic vorticity maxima in the 6-hourly ERA-Interim dataset, the most recent reanalysis of the European Centre for Medium-Range Weather Forecasts (Dee et al. 2011). The tracks of vorticity maxima were categorized according to the maximum intensity achieved within a global monsoon domain, with intensity determined by the simultaneous values of surface pressure anomaly and maximum surface wind speed occurring within 500 km of the 850-hPa vorticity maximum. A full description of the identification and classification algorithms is provided in Hurley and Boos (2015). Here we begin with all monsoon disturbances contained in the dataset, which includes all disturbances with surface pressures reduced by more than 2 hPa relative to a 21-day running mean. The climatological mean distribution of genesis points for Indian monsoon disturbances in the Yale dataset is highly similar to that documented in other track datasets for the Indian region (e.g., Sikka 2006). The dynamical structures and propagation characteristics of Indian monsoon depressions in the Yale dataset furthermore compare well with those obtained from field campaigns and other analyses (e.g., Godbole 1977; Sanders 1984) as discussed in Boos et al. (2015).

Monsoon disturbance genesis points in the Yale dataset extend around the globe, with maxima concentrated in the South Asian, West African, and Australian monsoon regions, as well as in the east Pacific intertropical convergence zone (ITCZ; Figs. 1a,b, Table 1). Other monsoon disturbance track datasets are limited to more confined geographical areas such as India (Mooley and Shukla 1987; Sikka 2006; Praveen et al. 2015) and Australia (Berry et al. 2012).

The Yale dataset contains two principle types of monsoon disturbances: those that have a warm-over-cold core and a column of potential vorticity (PV) extending through the entire depth of the troposphere, and others that have a uniformly warm core confined below 500 hPa and cyclonic PV trapped within 100 hPa of the surface (Hurley and Boos 2015). The shallow, warm-core disturbances occur most frequently over desert regions; they do not produce heavy precipitation, otherwise, the latent heat release would enhance cyclonic PV in a deeper layer of the troposphere and produce a PV structure that is not surface trapped.
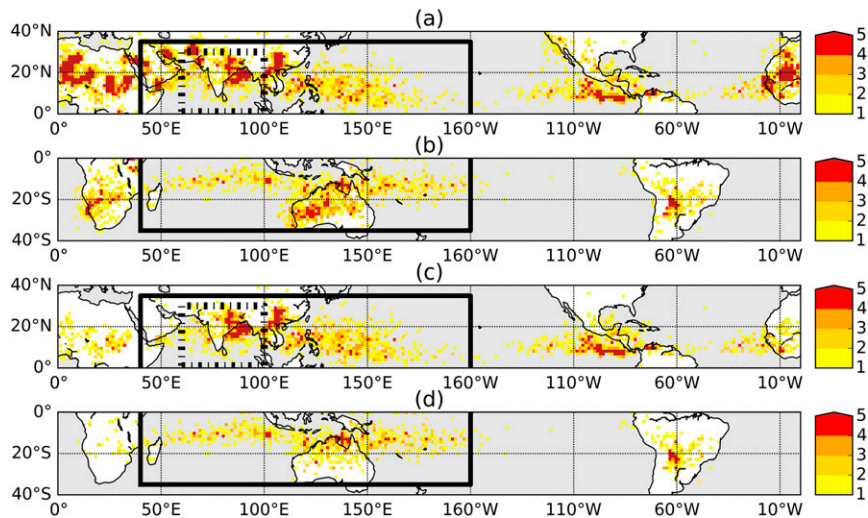
FIG. 1. Global monsoon disturbance genesis point distribution from January 1979–December 2012 for (a) all genesis points in boreal summer (JJAS), (b) all genesis points in austral summer (DJFM), (c) moist genesis points in boreal summer, and (d) moist genesis points in austral summer. The solid black box indicates the Asian–Australian region. The dashed–dotted box indicates the Indian continent region. Each grid box represents the number of storms over 34 years that occur in the specified 4-month season.

Different genesis mechanisms are likely responsible for deep and shallow monsoon disturbances. Here we focus on deep, precipitating monsoon disturbances and use a precipitable water criterion to exclude shallow monsoon disturbances from our analysis. In particular, for each monsoon disturbance track, we compute the lifetime maximum of storm-centered total column water vapor (TCWV) from daily averages of 6-hourly ERA-Interim data. This quantity has a bimodal distribution, suggesting that storms can be classified as "dry" or "moist" storms (Fig. 2). Between the two peaks, the distribution has a minimum value at $53\,\mathrm{kg\,m^{-2}}$, and we only include disturbances in our analysis whose lifetime maximum TCWV exceeds this threshold of $53\,\mathrm{kg\,m^{-2}}$. This threshold is roughly the value of TCWV at

which daily mean precipitation begins to increase sharply in analyses of tropical oceanic rainfall; mean precipitation is approximately $10\,\mathrm{mm\,day^{-1}}$ when TCWV is $53\,\mathrm{kg\,m^{-2}}$ in those analyses (e.g., Bretherton et al. 2004). We expect this criterion to exclude storms that never become moist, precipitating vortices. This criterion eliminates most monsoon disturbance genesis points in desert regions such as the Sahara, the Middle East, southwestern Asia, and southwestern Australia (Figs. 1c,d, Table 1). Hereafter, by monsoon disturbance we refer only to monsoon disturbances from the Yale dataset whose TCWV exceeds $53\,\mathrm{kg\,m^{-2}}$ during their lifetime.

b. Environmental parameters

The climatological mean state was computed from ERA-Interim by averaging monthly mean data over 34 years

TABLE 1. Observed genesis counts for all disturbances and for moist disturbances, as defined in the text. Genesis counts in various subregions are defined under their respective hemisphere.

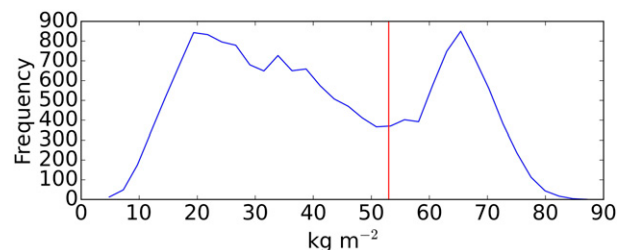| Region	Boundaries	All	Moist
Globe	90°S–90°N, 0°–360°	6757	3556
Tropics	35°S–35°N, 0°–360°	6676	3551
Northern Hemisphere	0°–35°N	4805	2442
North Africa	40°W–60°E	2413	385
Asian region	40°–200°E	2064	1488
India	60°–100°E	812	578
Northwest Pacific	100°E–160°W	898	866
East Pacific	160°–40°W	682	613
Southern Hemisphere	0°–35°S	1871	1109
South Africa	40°W–40°E	343	35
Australian region	40°–200°E	1227	870
Australia	100°–160°E	870	522
South America	160°–40°W	301	204

FIG. 2. Histogram of the maximum TCWV achieved along each track at the storm center, for all monsoon disturbances. The red, vertical line denotes the chosen separation between dry storms and moist storms: approximately $53\,\mathrm{kg\,m^{-2}}$. Present is a bimodal distribution with one peak centered around $18\,\mathrm{kg\,m^{-2}}$ and the other centered around $67\,\mathrm{kg\,m^{-2}}$.

TABLE 2. Candidate variables for the genesis index. Nondimensional numeric values listed for levels refer to the terrain-following native model levels of ERA-Interim.

Variable	Abbreviation	Units	Vertical level
Absolute vorticity	η	$10^{-5}\,s^{-1}$	49 ($\sim$ 850 hPa)
Column relative humidity	CRH	%	Column integrated
Estimated CAPE	ECAPE	$10^3\,J\,kg^{-1}$	See Eq. (1)
Lapse rate	LR	$K\,mb^{-1}$	49–36 ($\sim$ 850–400 hPa)
Relative humidity	RH	%	42 ($\sim$ 600 hPa)
Relative surface temperature	RST	K	Surface
Surface temperature	ST	K	Surface
Total column water vapor	TCWV	$kg\,m^{-2}$	Column integrated
Vertical shear	V	$m\,s^{-1}$	49–30 ($\sim$ 850–200 hPa)
Vertical velocity	Omega	$Pa\,s^{-1}$	39 ($\sim$ 500 hPa)

(1979–2012) for each of the 12 months of the year. The ERA-Interim dataset is available on 60 terrain-following model levels or 37 pressure levels with a horizontal resolution of $0.7° \times 0.7°$. Many of the analyses described here were also conducted using the ERA-40 reanalysis (Uppala et al. 2005) with qualitatively similar results.

We chose candidate variables for the MDGI based on previous studies. As discussed in the introduction, vertical wind shear has been commonly implicated in the genesis of Indian monsoon depressions, motivated by theoretical arguments that monsoon disturbances grow by drawing energy from the poleward temperature gradient of the monsoon mean state (e.g., Moorthi and Arakawa 1985; Rao et al. 2004). However, other studies have suggested that the genesis of monsoon disturbances may be fostered by high SST, large cyclonic low-level vorticity, and a humid midtroposphere, under the assumption that monsoon disturbance genesis is influenced by the same factors that control the genesis of canonical tropical cyclones (e.g., Ramesh Kumar and Sankar 2010; Prajeesh et al. 2013). With this justification, our pool of candidate variables includes those typically used in the construction of tropical cyclone genesis indices (Table 2). We include several measures of humidity: a column relative humidity defined as the ratio of the TCWV to the saturation value of TCWV, a midtropospheric relative humidity, and the TCWV. Dynamical variables include the midlevel vertical velocity, the magnitude of the low-level absolute vorticity,[2] and vertical wind shear.

Monsoon disturbances form over both land and ocean, so the environmental variables used in our

statistical model must also be defined over both land and ocean. Additionally, since some monsoons occur in regions with high topography, it is necessary to consider how large variations in surface pressure might affect our analyses. Therefore we sample the "low-level" variables listed above on ERA-Interim's terrain-following native model (eta) levels rather than on fixed pressure levels. For example, we define low-level absolute vorticity on model level 49, which lies at about 850 hPa over ocean, rather than the 850-hPa pressure level, and we define vertical wind shear as the magnitude of the vector difference between horizontal winds at model level 49 (again, around 850 hPa over ocean) and model level 30 (around 200 hPa over ocean). Model levels associated with other variables can be found in Table 2. For simplicity, we will henceforth refer to model levels using the approximate pressure they have over ocean.

A more difficult issue is how to deal with the fact that SST, relative SST, or potential intensity—all of which have been found to be important predictors in tropical cyclone genesis indices—are not defined over land. We first naively define a surface temperature field that consists of SST over ocean and 2-m air temperature over land. From this global surface temperature we also derive a relative surface temperature, analogous to the relative SST discussed by Vecchi and Soden (2007), by subtracting the contemporaneous tropical (30°S–30°N) mean surface temperature from each local value. However, we do not expect either of these surface temperature variables to be locally related to moist convective activity over land (e.g., the highest surface air temperatures typically occur over deserts). This motivates our definition of a new thermodynamic variable equal to the difference between the surface air moist static energy and the 200–400 hPa vertically averaged saturation moist static energy. We call this variable the estimated convective available potential energy (ECAPE),

---

[2] The Yale dataset identified monsoon disturbances using low-level relative vorticity, raising the concern that one of our candidate predictors is nearly the same variable as that used to generate the genesis data. However, the Yale dataset was derived using 6-hourly data while here we use the climatological mean state data. Showing that transient disturbances in reanalysis data can be statistically described by the climatological mean variables is a useful result.

$$\text{ECAPE} = h(2\,\text{m}) - \frac{1}{(p_2 - p_1)} \int_{p_1}^{p_2} h^*(p)\,dp, \quad (1)$$

where the moist static energy $h = c_p T + L_v q + gz$, L_v is the latent heat of vaporization, and other variables have their usual meteorological meanings. The saturation moist static energy h^* is the saturation value of h, and is averaged from $p_1 = 200\,\text{hPa}$ to $p_2 = 400\,\text{hPa}$. The integral is evaluated using pressure level data since 400 hPa is well above nearly all terrain.

This new variable is similar in spirit to relative SST, which has been used to explain variations in tropical cyclone activity under the premise that the energy available for moist convection depends not on an absolute SST but on the local vertical thermodynamic sounding (e.g., Vecchi and Soden 2007). The tropical atmosphere cannot maintain strong horizontal temperature gradients above the boundary layer, so an increase in tropical-mean SST will produce a roughly uniform increase in the temperature (or h^*) of the tropical troposphere, and a given location will be more convectively unstable only if it lies over SST that is warmer than the tropical-mean SST. Since convective stability directly depends not on SST but on thermodynamic properties of the subcloud layer, (1) is a generalization of relative SST without the limiting assumption that horizontal temperature gradients are weak. Retaining the possibility for free-tropospheric temperature gradients to affect convective stability is potentially important for monsoon disturbances, which have genesis maxima lying as much as 25° off the equator (Fig. 1). More simply, (1) is justified by the fact that convective available potential energy can be approximated by the difference between h of a lifted parcel and h^* of its environment, vertically integrated from the lifted condensation level to the level of neutral buoyancy (e.g., Emanuel 1994). Note that in (1), we would expect similar results if h and h^* were replaced with the moist entropy (or equivalent potential temperature) of a lifted parcel and its saturation value. In that case, (1) would represent the difference between the low-level equivalent potential temperature (which incorporates low-level moisture) and the upper-level saturation equivalent potential temperature (which depends only on temperature). Over ocean, a boreal summer climatology of ECAPE does look remarkably like SST and potential intensity (PI), with the advantage that ECAPE is defined over land (Fig. 3).

We also include in our collection of candidate variables a measure of the thermal stratification of the middle to lower troposphere, motivated by hypotheses that the spinup of tropical depressions is enhanced when the lower troposphere is cooler and the upper troposphere is
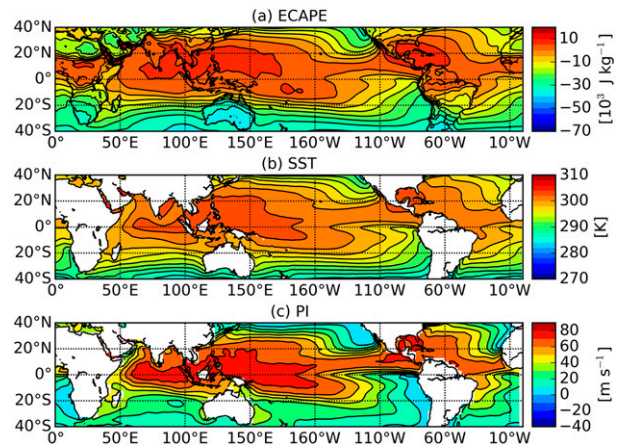

FIG. 3. JJAS climatologies of (a) estimated convective available potential energy (ECAPE, $10^3\,\text{J}\,\text{kg}^{-1}$), (b) sea surface temperature (SST, K), and (c) potential intensity expressed in wind speed (PI, $\text{m}\,\text{s}^{-1}$). Note the spatial similarities that exist across all three variables.

warmer than usual (see the review in Raymond et al. 2015). Such thermal anomalies would enhance the dry static stability of the middle troposphere and foster a more bottom-heavy vertical mass flux in a convectively coupled disturbance. Following Raymond et al. (2011), we define a discretized lapse rate as the difference in saturated equivalent potential temperature between model level 49 (around 850 hPa) and model level 36 (around 400 hPa), divided by the difference in pressure between those two levels. This variable is called the "lapse rate" in Table 2.

c. Regression methods

To construct our statistical model of the climatological distribution of monsoon disturbance genesis, we use the Poisson regression methodology described by Tippett et al. (2011) and Tippett et al. (2012). The expected number of monsoon disturbance genesis points occurring at any point in space–time is written as a log-linear model for storm counts:

$$\mu = \exp[\mathbf{b}^T \mathbf{x} + \log(\Delta x \Delta y\, T \cos\phi)]. \quad (2)$$

Here, μ is the expected number of monsoon disturbance genesis points during the climatological period and $\mathbf{b}$ is the vector of regression coefficients multiplying the climatological mean variables in the vector $\mathbf{x}$. The constant (intercept) term of the regression is included by adding elements to $\mathbf{x}$ that are identically equal to one. The last term (the offset) is included to make $\mathbf{b}$ independent of grid resolution and record length: Δx and Δy are the longitude and latitude grid spacing in degrees, respectively; ϕ is latitude in radians; and T is the number of years in the climatological period (here 34 covering the
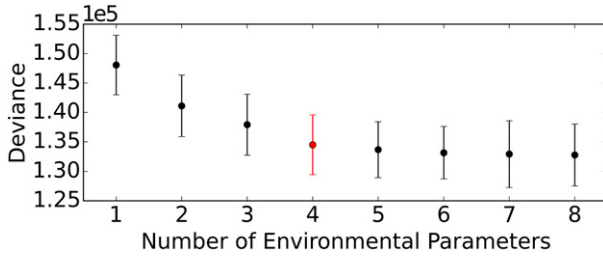
FIG. 4. Cross-validated deviance of the statistical model as a function of the number of environmental predictors. The mean (dots) and $\pm$ 1 standard deviation (error bars) of the deviation from 10 iterations of 10-fold cross validation are shown. Deviance decreases little when the number of predictors is increased from four (red) to five, supporting the decision to use four environmental predictors in the statistical model.

period 1979–2012). The quantity $\exp(\mathbf{b}^{\mathrm{T}}\mathbf{x})$ is therefore proportional to the number of monsoon disturbance genesis points per unit area per year. In this formulation, the regression coefficient values are interpreted as the relative change in the expected genesis frequency for a unit change in an environmental variable:

$$\mathbf{b}^{\mathrm{T}}\delta\mathbf{x} \simeq \frac{\delta\mu}{\mu}. \qquad (3)$$

We estimate the regression coefficients $\mathbf{b}$ from the storm counts in the Yale dataset and the ERA-Interim climatological monthly mean environmental variables using standard maximum likelihood estimation.

We use forward sequential feature selection (SFS) to objectively choose the variables included in the Poisson regression. SFS begins with the constant-only model and adds one variable at a time, choosing at each step the variable that minimizes an error criterion. Here the error criterion is deviance computed from 10 iterations of 10-fold cross validation (Solow 1989; Allen et al. 2015a). Deviance is a goodness-of-fit measure analogous to the sum of squared error and is oriented so that smaller values indicate better fit (McCullagh and Nelder 1989). Cross validation is used to avoid overfitting. In 10-fold cross validation, the data are randomly divided into 10 approximately equal subsets. Nine of the data subsets are used for training, and the remaining data subset is used for testing. The training and testing process is repeated so that each data subset is used once for testing, and 10 values of the cross-validated deviance are computed. In addition, we repeat the data division step 10 times so that 100 values of the deviance are available for each model. SFS results in a sequence of models with an increasing number of variables. We find that the mean cross-validated deviance decreases as more variables are added until four variables are used in the regression, and

TABLE 3. Poisson regression coefficients for the fit between global monsoon disturbance genesis points and mean state environmental variables. Standard errors are estimated from 10 000 bootstrap samples.

	b	b_{TCWV}	b_{η}	b_{ECAPE}	b_{RH}
Coefficients	−12.8	0.025	0.34	0.14	0.074
Errors	0.11	0.0027	0.0051	0.0031	0.0023

the addition of more variables does not substantially reduce the deviance (Fig. 4). The standard errors of the regression coefficients are estimated from 10 000 bootstrap samples, and the coefficients are highly significant.

3. Results

a. The global genesis index

The four environmental variables selected for use in our global MDGI are (in the order chosen by the SFS procedure): TCWV, low-level absolute vorticity (η), ECAPE, and 600-hPa relative humidity (RH). The associated regression coefficients and their standard errors are listed in Table 3. Although it was somewhat unexpected that two humidity variables—600-hPa RH and TCWV—were selected for the model, there is a robust improvement in fit when all four parameters are used instead of just three. Relative humidity and absolute measures of water content (e.g., TCWV) are physically distinct quantities, and humidity at a particular vertical level need not covary with the vertically integrated humidity. We thus take the selection of these two moisture variables as having some physical significance. The regression coefficients are all positive, indicating that monsoon disturbance genesis is more likely in regions with high low-level absolute vorticity, high ECAPE, and high humidity (in two different senses). Note that these coefficients have dimensions (equal to the inverse of the dimensions for the associated parameter listed in Table 2), so it is not possible to infer the importance of a parameter from its associated coefficient alone. However, the fractional change in MDGI associated with a unit change in each variable (using the units in Table 2) is provided by the coefficients. For example, a $1\,\mathrm{kg\,m^{-2}}$ increase in TCWV would produce a 2.5% increase in the MDGI. More insight on the sensitivity to individual parameters will be provided later when marginal distributions are presented.

The MDGI during local summer in each hemisphere reproduces the general features of both the horizontal structure and amplitude of the observed distribution of monsoon disturbance genesis frequency (Fig. 5). Like the observations, the MDGI
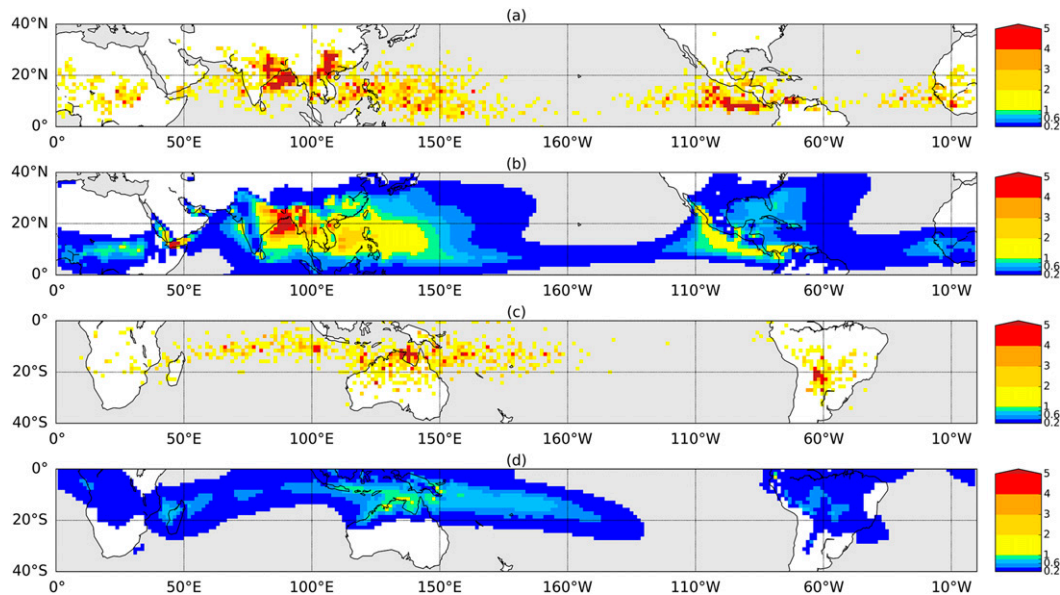
FIG. 5. Global spatial distribution of the number of (a) observed genesis points in boreal summer (JJAS), (b) the MDGI in boreal summer, (c) observed genesis points in austral summer (DJFM), and (d) the MDGI in austral summer. Each grid point represents the number of storms that occur over 34 years in the given 4-month season.

has a global maximum centered over the northern Bay of Bengal in boreal summer with a tail extending over land into the Indo-Gangetic plain. The MDGI even reproduces some of the observed minor features in the Indian monsoon region, such as the secondary maximum of genesis over the eastern Arabian Sea. The MDGI reasonably reproduces the broad features of observed local maxima over the west Pacific warm pool and northern Africa in boreal summer, and over southern Africa, the southern Indian Ocean, and the southwestern Pacific in austral summer.

When making these comparisons, it should be remembered that the MDGI provides an estimate of the expected value of the genesis frequency in a given time period and thus can take on noninteger values, unlike observed storm counts. Therefore, it is not surprising that the MDGI takes on values between zero and unity over large regions where the observed genesis counts fluctuate between zero and one. The MDGI clearly underestimates genesis maxima in the east Pacific ITCZ, just north of Australia (in the Gulf of Carpenteria), and over southwest China (in the lee of the Tibetan Plateau). The observed genesis maximum over South America may also be too horizontally diffuse in the MDGI. Genesis frequencies over the marginal seas surrounding the Arabian Peninsula are overestimated by the MDGI, especially in the Gulf of Aden. Overall, though, the MDGI successfully describes the major features of a fairly complicated spatial pattern of genesis frequency across multiple monsoon regions using just four climatological mean parameters.

The MDGI agrees fairly well with observed genesis frequencies in the zonal mean, when we evaluate the zonal mean for each hemisphere during its local summer season (Fig. 6). The MDGI underestimates the intensity of the equatorial minimum and the off-equatorial maxima when compared to the observed genesis distribution, but the MDGI captures the larger frequency of genesis in the Northern Hemisphere compared to the Southern Hemisphere. The maximum of the zonal mean MDGI in the Northern Hemisphere is too weak and located a few degrees too far poleward when compared to the observed maximum, a bias related to the behavior of the MDGI in the eastern Pacific. Note that the observed absolute maximum near 10°N is associated with high genesis frequencies in the east and west Pacific and the east Atlantic; the Bay of Bengal maximum, although especially prominent in plan view (Fig. 5a), does not stand out in the zonal mean because of relative minima in genesis frequency at other longitudes. The Southern Hemisphere maximum is also underestimated by the MDGI, due primarily to the MDGI being too low in the southern Indian Ocean and southwest Pacific (Figs. 5c,d). The observed zonal mean has higher meridional variability on length scales of a few hundred kilometers, which may be due to some combination of random noise and real geographic features that are not captured by the MDGI. Yet overall the MDGI provides a reasonably
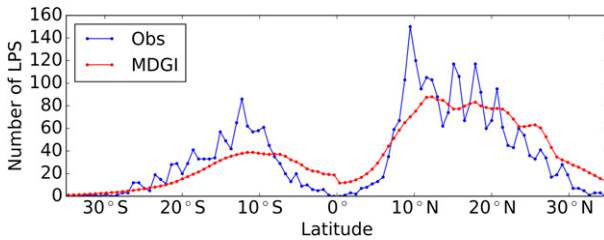
FIG. 6. Zonal integral of the number of observed genesis points (blue) and the MDGI (red) occurring by latitude across the globe over 34 years in boreal summer (JJAS) and austral summer (DJFM).
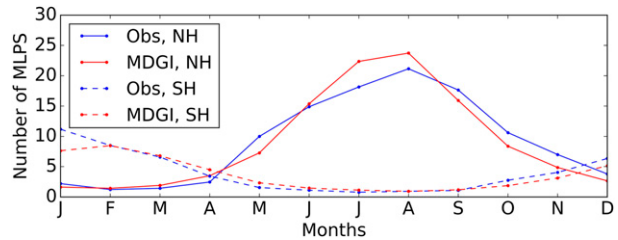

FIG. 7. Climatological distribution of the average number of observed genesis points (blue lines) and the MDGI (red lines) occurring per month per year across the globe over the Northern Hemisphere (solid lines) and Southern Hemisphere (dashed lines).

good fit, with the two zonal mean distributions having a correlation coefficient of 0.85.

The seasonal cycle of genesis in each hemisphere is also well reproduced by the MDGI. The number of genesis events occurring each month in our 34-yr period separately summed over the Northern and Southern Hemispheres compares well with hemispheric integrals of the MDGI (Fig. 7). As noted by Hurley and Boos (2015), 2–3 times as many monsoon disturbances form in the Northern Hemisphere as in the Southern Hemisphere during their local summers. These relative amplitudes of the seasonal cycle and the timing of the extrema are reproduced with only minor bias. Correlations between the time series of spatial sums of the MDGI and observed counts are 0.97 for the Northern Hemisphere and 0.95 for the Southern Hemisphere.

We now discuss how the four variables set the spatial distribution and seasonal cycle of the MDGI. The contribution of each variable to the spatial distribution of the MDGI presented in Fig. 5 is obtained by computing the local anomaly, relative to the spatial mean, of each variable multiplied by its respective regression coefficient. In boreal summer, most of the zonal asymmetry in the MDGI is produced by the asymmetries in ECAPE and midtropospheric RH, while the strong equatorial minimum of the MDGI is produced by the low-level absolute vorticity (Fig. 8). The vorticity has a fair amount of finescale structure around high topography and a distribution that is to first-order zonally symmetric, but there are notable deviations from zonal symmetry over the northern Bay of Bengal and Arabian Sea (where η reaches values that are achieved roughly 15° latitude farther poleward at other longitudes). This helps in understanding some of the prominent features of the seasonal mean MDGI. In particular, the high frequency of genesis over the Bay of Bengal (Figs. 5a,b) is associated with high values of all four variables. Although vorticity is high over the northern Arabian Sea and in the extratropical Pacific, the MDGI is reduced in those regions because of the substantially lower RH,

TCWV, and ECAPE. The large MDGI in the west Pacific warm pool is associated with midtropospheric humidities and ECAPE that are just as large as in the Bay of Bengal, but lower values of vorticity and TCWV. The high value of the MDGI just west of Central America is due primarily to the structure of ECAPE, while the dry midtroposphere over the Caribbean prevents the MDGI from reaching high values there even though the ECAPE is locally high. Similar associations between the MDGI and the environmental variables are found during austral summer (not shown).

To understand the seasonal cycle of the MDGI shown in Fig. 7, we found the relative contribution of each variable to the hemispheric seasonal cycle of the MDGI.
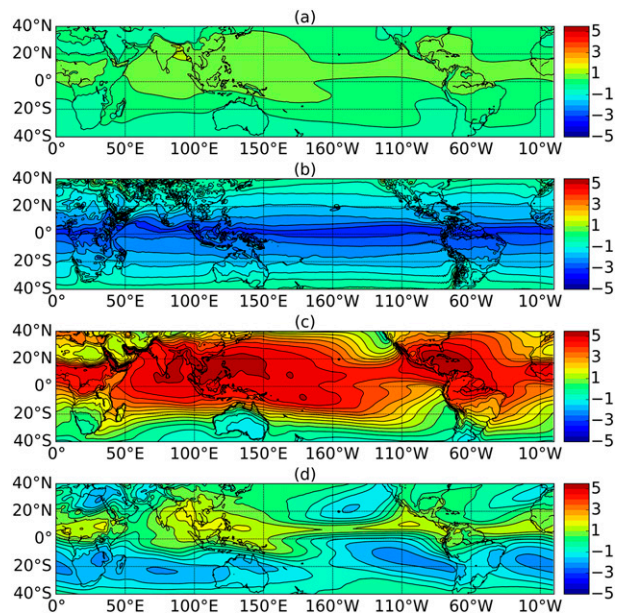

FIG. 8. Plan view of the anomalous relative contributions of (a) TCWV, (b) η, (c) ECAPE, and (d) RH to the JJAS mean of the MDGI. Anomalous relative contributions were calculated by computing the anomaly of each variable, relative to a spatial mean, and multiplying that anomaly by the regression coefficient corresponding to that variable.
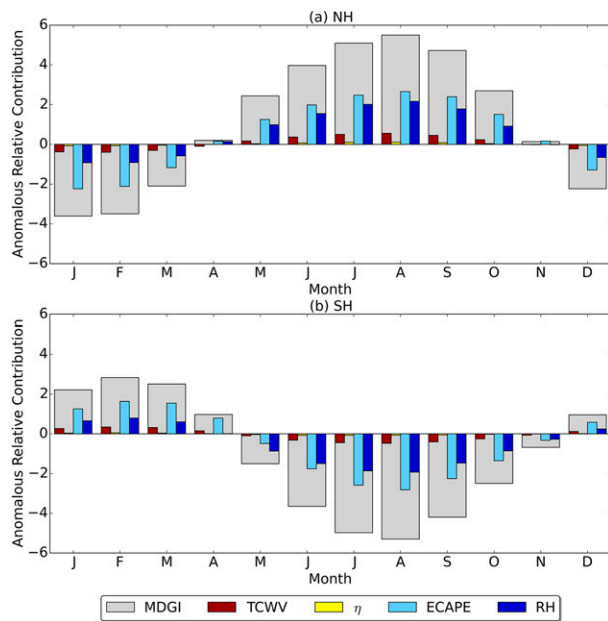
FIG. 9. Anomalous relative contributions of each individual climate variable to the overall climatological distribution for the (a) Northern Hemisphere and (b) Southern Hemisphere by month. Anomalous relative contributions were calculated by computing the difference between the climatological monthly mean and annual mean of each variable, multiplying that difference by the corresponding regression coefficient, and then summing over space. The vertical axis is scaled by 10 000.

Specifically, we computed the spatial integral of the anomaly, relative to the annual mean, of each variable multiplied by its respective regression coefficient. As expressed in (3), the sum of each of these integrals approximates the fractional change in the expected genesis frequency. The largest contributor to the seasonal cycle in both hemispheres is ECAPE, with midlevel RH being a close second (Fig. 9). Given the positive sign of the regression coefficients, this signifies the atmosphere becoming more convectively unstable and more humid in the midtroposphere in the seasons when genesis frequency is highest. In comparison, the TCWV makes a small contribution to the seasonal cycle, providing a clear example of the distinct behavior of the two moisture variables. To be clear, although vorticity makes a negligible contribution to the seasonal cycle of the hemispherically averaged MDGI, it plays an important role in setting the spatial distribution (e.g., Fig. 8). In contrast, TCWV makes a modest but nonnegligible contribution to both the seasonal cycle and spatial distribution of the MDGI.

The individual association of each environmental variable with the MDGI and with observed genesis frequencies can be better understood by examining the marginal distributions. A marginal distribution was
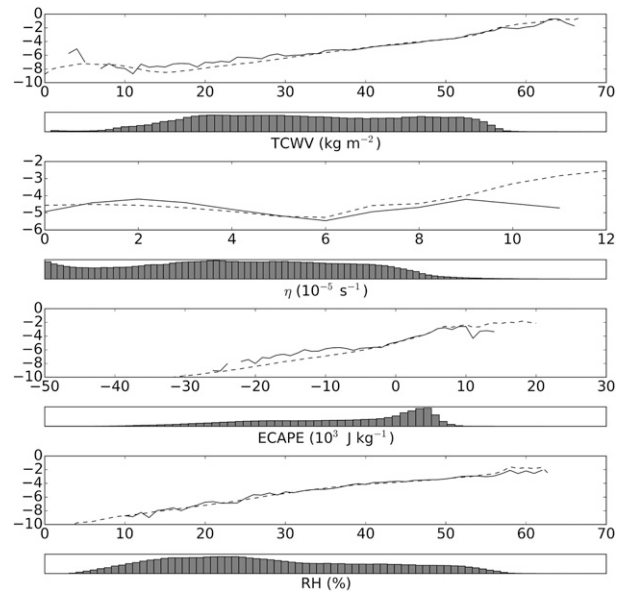

FIG. 10. Marginal distributions of the log of observed genesis points (solid line) and the MDGI (dashed line) per year as a function of total column water vapor (TCWV), absolute vorticity (η), estimated convective available potential energy (ECAPE), and relative humidity (RH). Histograms of each variable are divided into 70 bins and show the distribution of values of the mean state environmental variables.

constructed for each of the four environmental variables by dividing the full range of that environmental variable into 70 bins, then averaging the MDGI values and the observed genesis counts associated with all the latitude–longitude–time occurrences of that variable falling within each bin. The log of these bin-averaged MDGI values and genesis counts is presented (Fig. 10), together with the number of latitude–longitude–time points falling within each bin. For example, the log of the average MDGI and the log of the average number of genesis points that occurred at each RH bin from 0% to 100% are shown by the solid and dashed lines in the top part of the bottom panel of Fig. 10, and the histogram of RH values is shown in the bottom part of the bottom panel. The marginal distributions for ECAPE, TCWV, and 600-hPa RH are all monotonic, with the MDGI and observed genesis counts both increasing by over five orders of magnitude over the range of each of these variables (Fig. 10). The full range of storm frequencies furthermore stretches across nearly the full range of each of these three variables, indicating that there is no regime in which genesis becomes strongly controlled by one variable while another ceases to be relevant, except for, perhaps, vorticity. The marginal distribution for vorticity is highly distinct, with the MDGI and observed frequencies having a highly nonmonotonic dependence on vorticity. This behavior
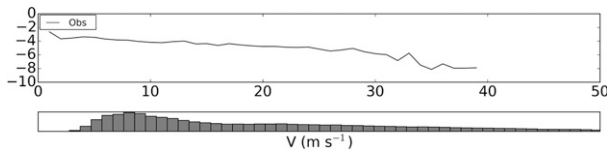
FIG. 11. Marginal distribution of the log of observed genesis points (solid line) per year as a function of vertical shear (V). The histogram is divided into 70 bins and shows the distribution of values of the mean state vertical shear.
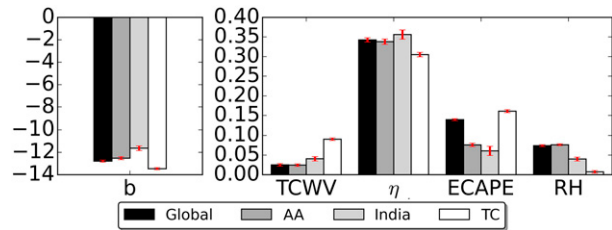

FIG. 12. Statistics from Poisson regressions between monsoon disturbance genesis points and mean state environmental variables over the entire globe, the Asian–Australian region, the Indian continent region, and between TC genesis points and the same mean state environmental variables over the entire globe. Red error bars indicate bootstrap errors.

seems to occur primarily because genesis frequencies are less zonally symmetric than absolute vorticities; both initially increase as one moves poleward from the equator to around 10° latitude in each hemisphere, but genesis frequencies drop to nearly zero as one moves farther poleward over much of the eastern Pacific (e.g., Fig. 5). ECAPE (Fig. 3) and TCWV are low in this region, and are associated with reduced genesis in that environment of high ambient vorticity. In contrast, the highest values of the MDGI are found in the Bay of Bengal, where the vorticity is large poleward of the low-level monsoon westerlies and where ECAPE, TCWV, and RH are all high (Fig. 8).

We also examined the marginal distribution of observed genesis frequencies for one variable that was not selected for our statistical model: vertical wind shear. As discussed in the introduction, some studies have argued that monsoon depressions are amplified by the process of baroclinic instability or some form of nonmodal baroclinic growth, drawing energy from the baroclinicity of their environment. This argument seems qualitatively consistent with the fact that monsoon disturbances are observed in regions of strong meridional temperature gradient, or equivalently of large vertical wind shear. However, observed genesis frequencies decrease monotonically as the vertical shear grows (Fig. 11). Although vertical shear is large in monsoon regions where monsoon disturbances form frequently, genesis most frequently occurs slightly poleward of the regions of strongest vertical shear. This was seen in the composites of Indian monsoon depressions constructed by Boos et al. (2015), which show that those storms exist 5°–10° latitude poleward of the strongest vertical shear of the zonal wind. While the statistical association of monsoon disturbance occurrence with vertical wind shear does not rule out baroclinic instability as a growth mechanism, it does disprove the idea that genesis is fostered by large vertical shear, at least in a climatological sense. Although our marginal distributions were created using monthly mean, climatological data, the composites shown by Boos et al. (2015) used 4 × daily data and so we do not expect that the sign of the dependence on shear was changed by our use of climatological means.

b. Comparison of global and regional statistics

It is remarkable that the major spatial features and seasonal cycle of the distribution of monsoon disturbances across all monsoon regions can be reproduced using just four environmental parameters, because it is not obvious that monsoon disturbance genesis in each monsoon region results from the same physical mechanism. To assess the regional variation of the statistical associations of genesis with environmental variables, we refit the index using the same four globally selected predictors (Table 3) using only data in the Asian–Australian region and then only in the Indian monsoon region (boxes in Fig. 1 delineate these domains). Since monsoon disturbances are most commonly associated with the Indian and Australian monsoons (e.g., Mooley and Shukla 1987; Berry et al. 2012), this procedure allows us to assess whether there is something unique about monsoon disturbance genesis in those classic monsoon regions. We emphasize that we did not repeat the process of variable selection with the SFS procedure in this analysis, but found new regression coefficients for each of these two subdomains.

The regression coefficients are quantitatively similar for regressions performed in the global, Asian–Australian, and Indian domains (Fig. 12). The largest difference is in the coefficient for ECAPE, which is twice as large for the global domain as for the Asian–Australian and Indian domains. This implies that ECAPE is more important for monsoon disturbance genesis over regions such as Africa and the Americas than in the Asian–Australian region. Spatial, zonal, and seasonal distributions of the MDGI produced using the regression coefficients calculated for the Indian and Asian–Australian regions are highly similar to those using the regression coefficients calculated for the globe (not shown). In fact, the spatial distribution of MDGI in the Asian–Australian region computed using global coefficients is highly correlated ($R = 0.97$)
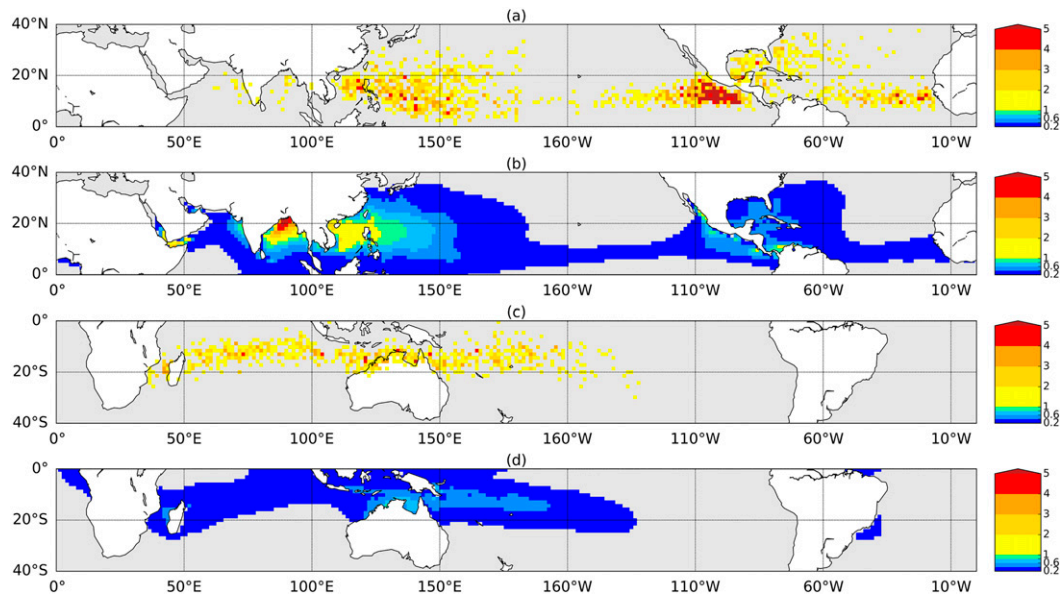
FIG. 13. Global spatial distribution of (a) TC genesis points in boreal summer (JJAS), (b) tropical cyclone genesis index (TCGI) density in boreal summer (JJAS), (c) TC genesis points in austral summer (DJFM), and (d) TCGI density in austral summer (DJFM). Each grid point represents the number of storms that occur over 34 years in the given 4-month season.

with the MDGI in the same region computed using co-efficients tuned for that region. The same is true for the Indian region.

We performed another analysis in which the variable selection (i.e., SFS) procedure was repeated for each of the subregions listed in Table 1, except for Africa and South America because of the small sample sizes in those regions. For all subregions, the deviance stopped decreasing after two or three predictors were added to the model, which is not surprising given the smaller amount of data available in the smaller domains. Chosen predictors in almost every subregion were a subset of the set of predictors chosen for the global domain (TCWV, η, ECAPE, and RH). An additional predictor, vertical shear, was selected by the SFS procedure for the subregions of Australia and the west North Pacific, although for the latter subregion it came in as the third predictor. Thus, while there is some regional variability in the statistical association of monsoon disturbances with environmental parameters, genesis in any region can be described by the four variables used in our global model, with the addition of vertical wind shear in select regions.

c. Comparison with tropical cyclone statistics

Finally, we compare our MDGI with genesis indices derived for tropical cyclones (TCs). Many of the candidate variables considered for inclusion in our statistical model (Table 2) were chosen because a few previous studies assumed that monsoon disturbance genesis is influenced by the same environmental variables that control genesis of TCs. If monsoon disturbances and TCs have highly similar statistical associations with environmental variables, this would support the idea that genesis of these two types of disturbances may be governed by similar mechanisms.

We regress the distribution of TC genesis points on the same four environmental variables chosen for our MDGI, using the same Poisson regression model. Note that we did not repeat the variable selection procedure; we expect that doing so would simply reproduce the TC genesis index presented by Tippett et al. (2011). Since there are no TC genesis points over land, grid cells containing land were not included in this regression. The TC genesis points were obtained from NOAA best-track data for the Atlantic and northeast Pacific, and from the U.S. Navy Joint Typhoon Warning Center for all other regions, as obtained from the website of K. Emanuel (ftp://texmex.mit.edu/pub/emanuel/HURR/tracks_netcdf/). The resulting regression coefficients are similar to those obtained for monsoon disturbances with two notable exceptions: compared to monsoon disturbance genesis, TC genesis is 2–3 times more sensitive to TCWV and much less sensitive to midlevel RH (Fig. 12). The coefficient for ECAPE, which is our generalization of relative SST, has a similar value for TCs as for the global distribution of monsoon disturbances.

The spatial distribution of the genesis index that uses coefficients obtained from the TC data has substantial bias when compared with the observed distribution[3] of TC genesis (Fig. 13). Although the genesis index qualitatively reproduces the off-equatorial maxima and equatorial minimum in the west Pacific, it greatly underestimates the frequency of genesis in the east Pacific and greatly overestimates frequency in the northern Indian Ocean, especially in the Bay of Bengal. The zonally elongated genesis maximum that stretches from Madagascar to the southwest Pacific is captured in its spatial structure, but is too weak in amplitude.

The poorer fit to the observed distribution of TC genesis may be caused by the lack of a vertical wind shear variable in the regression model. Wind shear is large over the north Indian Ocean during local summer, and could reduce the genesis index in that region if it was included in the model. The four variables used for our MDGI all have a positive relationship with genesis frequency, and all are large over the north Indian Ocean. Thus, even when the regression of TC genesis points on the same four environmental variables chosen for our MDGI is performed, the model is unable to capture important features of the observed distribution of TC genesis. Vertical shear is also the only variable lacking from our model that was included in the statistical model derived by Tippett et al. (2011) for TC genesis—their other three variables were low-level absolute vorticity, midtropospheric relative humidity, and relative SST (which we have confirmed behaves similarly to ECAPE over ocean). Incidentally, the comparatively poor fit of the TC genesis index provides a useful reference point that lends confidence to our correct choice of variables for fitting the observed distribution of monsoon disturbance genesis.

4. Summary and discussion

Although the tracks of Indian monsoon disturbances have been documented for over a century (Mooley and Shukla 1987), their genesis had not been systematically associated with properties of the environment in which they form. Furthermore, the genesis locations of monsoon disturbances in other regions were compiled only recently (Berry et al. 2012; Hurley and Boos 2015), and also had not been statistically associated with properties of their environment. Here we showed that the spatial distribution and seasonal cycle of the global distribution of monsoon disturbance genesis can be generally

captured by a genesis index based on monthly climatologies of four variables: total column water vapor, low-level absolute vorticity, an approximate measure of convective available potential energy (ECAPE), and midlevel relative humidity. Observed genesis frequency increases with all four variables, consistent with the simple expectation that a precipitating vortex will be more likely to spin up in a moist, convectively unstable environment with high ambient vorticity. The locally higher values of the four variables in the Bay of Bengal explain why there is a maximum in genesis events there; the more moderate frequency of genesis in the west Pacific is associated with similar levels of humidity and ECAPE as in the Bay of Bengal, but lower vorticities. The seasonal cycle of hemispherically integrated storm counts is due primarily to the seasonal cycle in ECAPE and midlevel RH.

Although previous studies have argued that vertical wind shear influences monsoon disturbance genesis in the Indian monsoon (e.g., Sikka 1977; Prajeesh et al. 2013), vertical wind shear was not selected by the objective procedure used to choose the index variables. Moreover, the climatological strength of vertical wind shear was shown to be negatively associated with genesis frequency, contradicting the idea that monsoon disturbance genesis is enhanced in regions of strong climatological shear. This is consistent with the finding by Cohen and Boos (2016) that baroclinic instability cannot explain the dynamical structures observed during the spinup of Indian monsoon depressions. A measure of midtropospheric static stability was also considered as a candidate variable, motivated by previous work suggesting that the spinup of tropical depressions may be fostered by enhanced static stability in the midtroposphere, but it did not significantly improve the index performance.

Our MDGI shares some variables with indices previously proposed to describe tropical cyclogenesis (Camargo et al. 2007; Tippett et al. 2011), but it is unique in that it is defined over both land and ocean. This feature is necessary because monsoon disturbance genesis occurs over land in monsoon regions, and is made possible by our use of ECAPE as a predictor instead of relative SST. ECAPE may well have applications beyond its use here in our MDGI. Like relative SST, ECAPE provides a simple measure of moist convective instability but, unlike relative SST, does not neglect horizontal temperature gradients in the free troposphere or inhomogeneities in the air–sea thermodynamic disequilibrium. And unlike potential intensity, ECAPE is defined over land, which may be advantageous even for indices of TC genesis if one wishes to consider the possibility that the early stages of TC

---

[3] While Figs. 13a,c show some TC genesis points overlapping with coastlines, this is merely an artifact of the gridding process.

genesis may sometimes occur over land (e.g., perhaps tropical depression spinup sometimes occurs over land even though best-track datasets only provide genesis points over ocean).

Two regional genesis indices based on the same environmental variables were created for Asian–Australian monsoon disturbances and for Indian monsoon disturbances. The coefficients obtained for these regional regressions are very similar to those obtained using the global distribution of monsoon disturbances, implying that similar processes may govern monsoon disturbance genesis in all regions. In contrast, regressing observed TC genesis counts on the same four variables yielded a model that greatly overpredicted the number of TCs in the north Indian Ocean and underpredicted numbers in the east Pacific. These biases are likely due to the absence of vertical wind shear as a model variable; although monsoon disturbances and TCs both have genesis frequencies that decrease with the strength of vertical shear, this relationship seems to be important for monsoon disturbances only in the regional domains of Australia and the west Pacific, at least in a climatological mean sense.

A logical next step would be to examine the statistical associations of monsoon disturbance genesis with temporally local environmental variables (e.g., monthly mean or higher-frequency data instead of monthly climatologies), or to determine whether our MDGI has any predictive skill. Our MDGI might also be used to improve understanding of the covariation of monsoon disturbance genesis with various climate oscillations, such as the Madden–Julian oscillation and ENSO. Yet even without these extensions, the MDGI improves our understanding of monsoon disturbance genesis: it has statistical associations that are common across all monsoon regions and that are consistent with the idea of genesis being fostered by humid, vorticity-rich environments with high convective available potential energy.

REFERENCES

Ajayamohan, R., W. J. Merryfield, and V. V. Kharin, 2010: Increasing trend of synoptic activity and its relationship with extreme rain events over central India. *J. Climate*, **23**, 1004–1013, doi:10.1175/2009JCLI2918.1.

Allen, J. T., M. K. Tippett, and A. H. Sobel, 2015a: An empirical model relating U.S. monthly hail occurrence to large-scale meteorological environment. *J. Adv. Model. Earth Syst.*, **7**, 226–243, doi:10.1002/2014MS000397.

——, ——, and ——, 2015b: Influence of the El Niño/Southern Oscillation on tornado and hail frequency in the United States. *Nat. Geosci.*, **8**, 278–283, doi:10.1038/ngeo2385.

Aon Benfield, 2010: Pakistan flood event recap report. Tech. Rep., 11 pp. [Available online at http://www.aon.com/attachments/reinsurance/201008_pakistan_flood.pdf.]

Bell, R., K. Hodges, P. L. Vidale, J. Strachan, and M. Roberts, 2014: Simulation of the global ENSO–tropical cyclone teleconnection by a high-resolution coupled general circulation model. *J. Climate*, **27**, 6404–6422, doi:10.1175/JCLI-D-13-00559.1.

Berry, G. J., M. J. Reeder, and C. Jakob, 2012: Coherent synoptic disturbances in the Australian monsoon. *J. Climate*, **25**, 8409–8421, doi:10.1175/JCLI-D-12-00143.1.

Boos, W. R., J. V. Hurley, and V. S. Murthy, 2015: Adiabatic westward drift of Indian monsoon depressions. *Quart. J. Roy. Meteor. Soc.*, **141**, 1035–1048, doi:10.1002/qj.2454.

Bretherton, C. S., M. E. Peters, and L. E. Back, 2004: Relationships between water vapor path and precipitation over the tropical oceans. *J. Climate*, **17**, 1517–1528, doi:10.1175/1520-0442(2004)017<1517:RBWVPA>2.0.CO;2.

Camargo, S. J., K. A. Emanuel, and A. H. Sobel, 2007: Use of a genesis potential index to diagnose ENSO effects on tropical cyclone genesis. *J. Climate*, **20**, 4819–4834, doi:10.1175/JCLI4282.1.

Chen, T. C., and S. P. Weng, 1999: Interannual and intraseasonal variations in monsoon depressions and their westward-propagating predecessors. *Mon. Wea. Rev.*, **127**, 1005–1020, doi:10.1175/1520-0493(1999)127<1005:IAIVIM>2.0.CO;2.

Cohen, N. Y., and W. R. Boos, 2016: Perspectives on moist baroclinic instability: Implications for the growth of monsoon depressions. *J. Atmos. Sci.*, **73**, 1767–1788, doi:10.1175/JAS-D-15-0254.1.

Dee, D. P., and Coauthors, 2011: The ERA-interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, doi:10.1002/qj.828.

DeMaria, M., J. A. Knaff, and B. H. Connell, 2001: A tropical cyclone genesis parameter for the tropical Atlantic. *Wea. Forecasting*, **16**, 219–233, doi:10.1175/1520-0434(2001)016<0219:ATCGPF>2.0.CO;2.

Emanuel, K. A., 1991: A scheme for representing cumulus convection in large-scale models. *J. Atmos. Sci.*, **48**, 2313–2329, doi:10.1175/1520-0469(1991)048<2313:ASFRCC>2.0.CO;2.

——, 1994: *Atmospheric Convection.* Oxford University Press, 580 pp.

——, and D. Nolan, 2004: Tropical cyclone activity and the global climate system. *26th Conf. on Hurricanes and Tropical*

Meteorology, Miami, FL, Amer. Meteor. Soc., 10A.2. [Available online at https://ams.confex.com/ams/26HURR/techprogram/paper_75463.htm.]

Farrell, B., 1985: Transient growth of damped baroclinic waves. *J. Atmos. Sci.*, **42**, 2718–2727, doi:10.1175/1520-0469(1985)042<2718:TGODBW>2.0.CO;2.

Godbole, R. V., 1977: The composite structure of the monsoon depression. *Tellus*, **29A**, 25–40, doi:10.1111/j.2153-3490.1977.tb00706.x.

Goswami, B. N., R. N. Keshavamurty, and V. Satyan, 1980: Role of barotropic, baroclinic and combined barotropic-baroclinic instability for the growth of monsoon depressions and midtropospheric cyclones. *J. Earth Syst. Sci.*, **89**, 79–97.

Gray, W. M., 1979: Hurricanes: Their formation, structure and likely role in the tropical circulation. *Meteorology over the Tropical Oceans*, D. Shaw, Ed., Royal Meteorological Society, 155–218.

Hodges, K. I., 1995: Feature tracking on the unit-sphere. *Mon. Wea. Rev.*, **123**, 3458–3465, doi:10.1175/1520-0493(1995)123<3458:FTOTUS>2.0.CO;2.

Hurley, J. V., and W. R. Boos, 2015: A global climatology of monsoon low-pressure systems. *Quart. J. Roy. Meteor. Soc.*, **141**, 1049–1064, doi:10.1002/qj.2447.

Joseph, S., and Coauthors, 2015: North Indian heavy rainfall event during June 2013: Diagnostics and extended range prediction. *Climate Dyn.*, **44**, 2049–2065, doi:10.1007/s00382-014-2291-5.

Kiladis, G. N., C. D. Thorncroft, and N. M. Hall, 2006: Three-dimensional structure and dynamics of African easterly waves. Part I: Observations. *J. Atmos. Sci.*, **63**, 2212–2230, doi:10.1175/JAS3741.1.

Krishnamurthy, V., and R. Ajayamohan, 2010: Composite structure of monsoon low pressure systems and its relation to Indian rainfall. *J. Climate*, **23**, 4285–4305, doi:10.1175/2010JCLI2953.1.

Krishnamurti, T. N., and S. Gadgil, 1985: On the structure of the 30 to 50 day mode over the globe during FGGE. *Tellus*, **37A**, 336–360, doi:10.1111/j.1600-0870.1985.tb00432.x.

——, J. Molinari, H.-L. Pan, and V. Wong, 1977: Downstream amplification and formation of monsoon disturbances. *Mon. Wea. Rev.*, **105**, 1281–1297, doi:10.1175/1520-0493(1977)105<1281:DAAFOM>2.0.CO;2.

——, V. Wong, H.-L. Pan, R. Pasch, J. Molinari, and P. Ardanuy, 1983: A three-dimensional planetary boundary layer model for the Somali jet. *J. Atmos. Sci.*, **40**, 894–908, doi:10.1175/1520-0469(1983)040<0894:ATDPBL>2.0.CO;2.

Lindzen, R. S., B. Farrell, and A. J. Rosenthal, 1983: Absolute barotropic instability and monsoon depressions. *J. Atmos. Sci.*, **40**, 1178–1184, doi:10.1175/1520-0469(1983)040<1178:ABIAMD>2.0.CO;2.

McCullagh, P., and J. A. Nelder, 1989: *Generalized Linear Models.* 2nd ed. Chapman and Hall, 532 pp.

Mooley, D. A., and J. Shukla, 1987: Characteristics of the westward-moving summer monsoon low pressure systems over the Indian region and their relationship with the monsoon rainfall. University of Maryland Center for Ocean–Land–Atmosphere Interactions Tech. Rep., 128 pp.

Moorthi, S., and A. Arakawa, 1985: Baroclinic instability with cumulus heating. *J. Atmos. Sci.*, **42**, 2007–2031, doi:10.1175/1520-0469(1985)042<2007:BIWCH>2.0.CO;2.

Prajeesh, A. G., K. Ashok, and D. Rao, 2013: Falling monsoon depression frequency: A Gray-Sikka conditions perspective. *Sci. Rep.*, **3**, 1–8, doi:10.1038/srep02989.

Praveen, V., S. Sandeep, and R. S. Ajayamohan, 2015: On the relationship between mean monsoon precipitation and low pressure systems in climate model simulations. *J. Climate*, **28**, 5305–5324, doi:10.1175/JCLI-D-14-00415.1.

Ramesh Kumar, M. R., and S. Sankar, 2010: Impact of global warming on cyclonic storms over North Indian Ocean. *Indian J. Mar. Sci.*, **39**, 516–520.

Rao, B. R., D. V. Rao, and V. B. Rao, 2004: Decreasing trend in the strength of Tropical Easterly Jet during the Asian summer monsoon season and the number of tropical cyclonic systems over Bay of Bengal. *Geophys. Res. Lett.*, **31**, L14103, doi:10.1029/2004GL019817.

Raymond, D., S. L. Sessions, and C. López Carrillo, 2011: Thermodynamics of tropical cyclogenesis in the northwest Pacific. *J. Geophys. Res.*, **116**, D18101, doi:10.1029/2011JD015624.

——, Ž. Fuchs, S. Gjorgjievska, and S. Sessions, 2015: Balanced dynamics and convection in the tropical troposphere. *J. Adv. Model. Earth Syst.*, **7**, 1093–1116, doi:10.1002/2015MS000467.

Saha, K., F. Sanders, and J. Shukla, 1981: Westward propagating predecessors of monsoon depressions. *Mon. Wea. Rev.*, **109**, 330–343, doi:10.1175/1520-0493(1981)109<0330:WPPOMD>2.0.CO;2.

Sanders, F., 1984: Quasi-geostrophic diagnosis of the monsoon depression of 5–8 July 1979. *J. Atmos. Sci.*, **41**, 538–552, doi:10.1175/1520-0469(1984)041<0538:QGDOTM>2.0.CO;2.

Shukla, J., 1977: Barotropic-baroclinic instability of mean zonal wind during summer monsoon. *Pure Appl. Geophys.*, **115**, 1449–1461, doi:10.1007/BF00874418.

Sikka, D. R., 1977: Some aspects of the life history, structure and movement of monsoon depressions. *Pure Appl. Geophys.*, **115**, 1501–1529, doi:10.1007/BF00874421.

——, 2006: A study on the monsoon low pressure systems over the Indian region and their relationship with drought and excess monsoon seasonal rainfall. Center for Ocean–Land–Atmosphere Studies Tech. Rep. 217, 145 pp.

Solow, A. R., 1989: Statistical modeling of storm counts. *J. Climate*, **2**, 131–136, doi:10.1175/1520-0442(1989)002<0131:SMOSC>2.0.CO;2.

Tippett, M. K., S. J. Camargo, and A. H. Sobel, 2011: A Poisson regression index for tropical cyclone genesis and the role of large-scale vorticity in genesis. *J. Climate*, **24**, 2335–2357, doi:10.1175/2010JCLI3811.1.

——, A. H. Sobel, and S. J. Camargo, 2012: Association of U.S. tornado occurrence with monthly environmental parameters. *Geophys. Res. Lett.*, **39**, L02801, doi:10.1029/2011GL050368.

Uppala, S. M., and Coauthors, 2005: The ERA-40 re-analysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012, doi:10.1256/qj.04.176.

Vecchi, G. A., and B. J. Soden, 2007: Increased tropical Atlantic wind shear in model projections of global warming. *Geophys. Res. Lett.*, **34**, L08702, doi:10.1029/2006GL028905.

Webster, P. J., V. E. Toma, and H.-M. Kim, 2011: Were the 2010 Pakistan floods predictable? *Geophys. Res. Lett.*, **38**, L04806, doi:10.1029/2010GL046346.

Yoon, J. H., and T. C. Chen, 2005: Water vapor budget of the Indian monsoon depression. *Tellus*, **57A**, 770–782, doi:10.1111/j.1600-0870.2005.00145.x.