

A GENESIS POTENTIAL INDEX FOR ASIAN-AUSTRALIAN MONSOON LOW PRESSURE SYSTEMS

Sarah Dunn Ditchek
Advisor: William Boos
Second Reader: Trude Storelmo

April 30, 2014

A Senior Thesis presented to the faculty of the Department of Geology and Geophysics,
Yale University, in partial fulfillment of the Bachelor's Degree.

In presenting this thesis in partial fulfillment of the Bachelor's Degree from the Department of Geology and Geophysics, Yale University, I agree that the department may make copies or post it on the departmental website so that others may better understand the undergraduate research of the department. I further agree that extensive copying of this thesis is allowable only for scholarly purposes. It is understood, however, that any copying or publication of this thesis for commercial purposes or financial gain is not allowed without my written consent.

Sarah Dunn Ditchek, 30 April, 2014

ABSTRACT

Creating a statistically-based genesis potential index which identifies meteorological variables that are well correlated with monsoon low pressure system genesis is vital as these storm systems contribute over half of the rainfall to monsoonal regions, rainfall which can be catastrophic. Grounded in the work of Emanuel and Nolan (2004), Camargo et al. (2008), and Tippet et al. (2011), this research constructed such an index for the Asian-Australian region by performing a Poisson regression on 1) genesis points at moist locations defined by a total column water vapor value greater than 35 kgm^{-2} , and 2) four meteorological variables retrieved from the European Centre for Medium-Range Weather Forecasts' monthly mean ERA-Interim reanalysis datasets: absolute vorticity at 850- hPa, estimated convective available potential energy calculated by subtracting the 200- to 400- hPa mean moist static energy from that at the surface, total column water vapor, and wind shear between 850- and 200- hPa. The index indicates a negative correlation between genesis and wind shear and positive correlations between genesis and the remaining three meteorological variables, with absolute vorticity being most strongly, positively correlated to genesis. Spatially and climatologically, the index is highly correlated to the actual genesis distributions in the Asian-Australian region. Through similar methodology, it was also found that genesis points can be described by similar statistics for the Indian Continent region and the entire globe. Finally, in an attempt to replicate the Tippet et al. (2011) results, a genesis potential index was created using tropical cyclone "best-track data" to compare the monsoon low pressure system genesis index to that of tropical cyclones. It was found that their respective genesis indices are very similar, and if tropical cyclones could form over land, their distribution would mimic the monsoon low pressure system spatial distribution which stretches inland from the Bay of Bengal.

TABLE OF CONTENTS

| | | |
|--------------|---|----|
| I. | Introduction | 1 |
| II. | Region of Interest: The Asian-Australian (AA) Region | 2 |
| III. | An AA MLPS GPI | 2 |
| | A. Pre-Construction | 2 |
| | 1. Statistical Method for GPI Creation..... | 2 |
| | a. <i>The Emanuel and Nolan (2004) Multiple Regression Method</i> | 3 |
| | b. <i>The Tippet et al. (2011) Poisson Regression Method</i> | 3 |
| | 2. Meteorological Data..... | 4 |
| | a. <i>Estimated Convective Available Potential Energy (ECAPE)</i> | 5 |
| | b. <i>Total Column Water Vapor (TCWV)</i> | 6 |
| | c. <i>Absolute Vorticity (η)</i> | 6 |
| | d. <i>Wind Shear (V)</i> | 6 |
| | 3. Genesis Points (GP)..... | 6 |
| | a. <i>Digitizing Mooley and Shukla (1987) and Sikka (2006)</i> | 7 |
| | b. <i>Compiling MLPS from Hurley (2014) Database</i> | 8 |
| | B. Refinement and Exploration | 9 |
| | C. Construction | 10 |
| | 1. <i>Executing the Poisson Regression: The Statistics</i> | 10 |
| | 2. <i>Spatial Distribution</i> | 13 |
| | 3. <i>Zonal Distribution</i> | 14 |
| | 4. <i>Seasonal Climatology</i> | 15 |
| | 5. <i>Interannual Variability</i> | 15 |
| IV. | An Indian MLPS GPI | 16 |
| | A. <i>Executing the Poisson Regression: The Statistics</i> | 16 |
| | B. <i>Spatial Distribution</i> | 17 |
| | C. <i>Seasonal Climatology</i> | 18 |
| V. | A Global MLPS GPI | 18 |
| | A. <i>Executing the Poisson Regression: The Statistics</i> | 19 |
| | B. <i>Spatial Distribution</i> | 19 |
| | C. <i>Seasonal Climatology</i> | 21 |
| VI. | A Global TC GPI | 22 |
| | A. <i>Executing the Poisson Regression: The Statistics</i> | 22 |
| | B. <i>Spatial Distribution</i> | 23 |
| | C. <i>Seasonal Climatology</i> | 25 |
| VII. | Future Work | 27 |
| VIII. | Conclusion | 27 |
| IX. | Acknowledgments | 28 |
| X. | References | 29 |

I. Introduction

A monsoon circulation is induced by a landmass lying poleward of an equatorial ocean and is accompanied by a strong seasonal precipitation cycle of dry winters and wet summers over land. During winter, air rises over warm tropical oceans and subsides over cooler landmasses, creating equatorward, low-level monsoonal flows. This circulation is known as the typical Hadley Cell circulation (Clift and Plumb, 2008). During summer months, landmasses gain heat at faster rates than oceans due to land's low specific heat content. This creates strong land-ocean temperature gradients that induce reverse Hadley circulations, wherein air rises over landmasses through convection and subsides over now cooler oceans. Poleward, low-level monsoonal flows result (Clift and Plumb, 2008). If the shift in monsoonal flow in a given region is at least 120 degrees then that region is termed a monsoonal region (Ramage, 1971).

Monsoonal rains are vital to the economy, culture, and lives of many living in monsoonal regions as well as the region's ecosystems. Almost half of the total seasonal precipitation in monsoonal regions is said to be the result of embedded, synoptic-scale monsoon low pressure systems (MLPS) within the monsoonal circulation (Yoon and Chen, 2005). Nevertheless, MLPS have the potential to create catastrophic rainfall events dependent on their level of intensity: low, depression, deep depression, or typhoon (Ajayamohan, 2010). Some of the most dramatic examples of the devastation MLPS can cause involved summer Indian monsoons. A few examples include the 2013 monsoon causing over 500 deaths, over 80,000 displaced and over 5,000 missing individuals (BBC News, 2013; Masters, 2013) and the 2005 world-record breaking monsoon dumping 944 mm of rain within a 24 hour period and causing \$1 billion in economic loss (Jenamani et al., 2006).

Yet despite the impact of MLPS, very little is known or understood about how they form. Therefore, this research was undertaken to create a genesis potential index (GPI) for MLPS which would estimate how likely it is that a MLPS will form in a specific region of the globe at a certain time as a function of meteorological variables such as temperature, humidity, and winds. A GPI is vital for several reasons: it can improve our understanding of MLPS, the meteorological variables that control them, and their influence on global climate. It can also lead to improved forecasts of these MLPS, thereby aiding in preparedness operations and preventing or minimizing damage to or loss of human life and activities, animal life, ecosystems, and property. While the genesis of the higher intensity MLPS, such as deep depressions or typhoons, might be of interest to study, their

numbers are too small to yield a robust statistical assessment. Thus, the genesis of all MLPS was investigated, regardless of the eventual level of intensity attained.

II. Region of Interest: The Asian-Australian (AA) Region

Monsoonal regions and their embedded MLPS are found worldwide. The region that is most associated with monsoons is the Indian Continent region. In order to allow for the study of MLPS in both boreal and austral winters and summers and to create one GPI for many different regions rather than creating individual GPIs for each region impacted by MLPS, this research focused on the Indian Continent region as well as the additional regions of southern Asia, the Southern Ocean, and Australia. This entire area, now termed the Asian-Australian (AA) region, stretches across a diverse topographical area (Fig. 1) from 35°S to 35°N and 40°E to 160°E. AA includes mountain ranges such as the Himalayas, plateaus such as the Tibetan Plateau, the deserts of Australia, the rivers of China and India, and the seas and oceans of the equatorial regions.

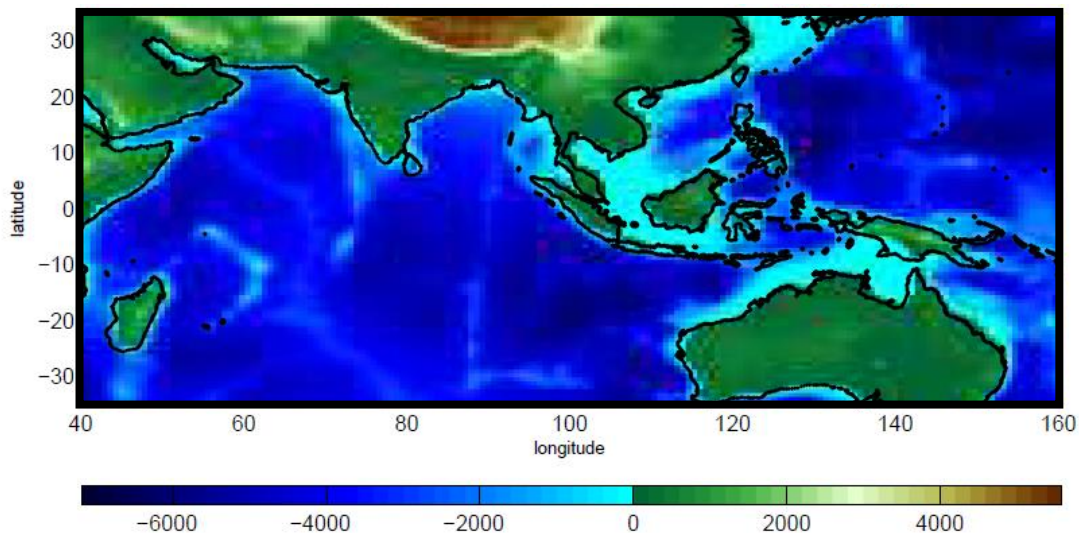


FIG. 1. Topography of the AA region in meters above and below sea level. Data retrieved from NGDC (1998).

III. An AA MLPS GPI

A. Pre-Construction

To create a MLPS GPI, three types of information are needed: 1) an appropriate statistical method to relate meteorological data to genesis points, 2) the meteorological data itself, and 3) the date, latitude, and longitude of genesis points in the region of interest.

1. Statistical Method for GPI Creation

Emanuel and Nolan (2004) created a GPI for tropical cyclones (TC) from empirical results discovered by Gray (1979). The TC GPI was created through the use of a multiple regression in

order to determine the relationship, if any, between TC GP and meteorological variables. Recently, a more rigorously derived and statistically robust GPI for TC was formulated by Tippet et al. (2011) using a Poisson regression technique.

a. The Emanuel and Nolan (2004) Multiple Regression Method

The TC GPI developed by Emanuel and Nolan (2004) upon which this research’s GPI is based is defined as

$$\text{GPI} = |10^5\eta|^{3/2} \left(\frac{\mathcal{H}}{50}\right)^3 \left(\frac{\text{PI}}{70}\right)^3 (1 + 0.1V)^{-2}, \quad (1)$$

where η is the 850 hPa absolute vorticity (s^{-1}), \mathcal{H} is the 600 hPa relative humidity (%), PI is the potential intensity (ms^{-1}), and V is the magnitude of the vertical wind shear between 850 and 200 hPa (ms^{-1}). TC “best-track data” was obtained from Dr. Kerry Emanuel, Professor of Atmospheric Science at the Massachusetts Institute of Technology (Emanuel, 2013). The present research initially attempted a multiple regression technique to fit the TC “best-track data” to meteorological variables as in Emanuel and Nolan (2004). However, their results were unable to be replicated as the variables were too sensitive to small changes in the data and the fit was not robust. Camargo et al. (2007) found similar difficulties. When attempting a multiple regression on MLPS data, similar statistical issues were found. Therefore, the method employed by Tippet et al. (2011) was used to create a GPI for MLPS.

b. The Tippet et al. (2011) Poisson Regression Method

Tippet et al. (2011) performed a Poisson regression on TC. They chose this specific method since Poisson regressions are used to model count data (i.e. whether an event occurs or not). Additionally, each TC GP is independent of one another and the rate of TC genesis has an interannual variation, two factors making a Poisson regression the appropriate statistical test to use (Elsner and Jagger, 2013). Using similar variables as in the Emanuel and Nolan (2004) index, the Tippet et al. (2011) index took the form

$$\mu = \exp(b + b_\eta\eta + b_{\mathcal{H}}\mathcal{H} + b_T T + b_V V + \log\cos\Phi), \quad (2)$$

where μ represents the number of TC GP expected per month, b represents a constant intercept term, b_i represents the coefficient associated with the particular climate variable (where i is η , \mathcal{H} , T , or V), η is the 850 hPa absolute vorticity (10^5s^{-1}), \mathcal{H} is the 600 hPa relative humidity (%), T is the sea surface temperature (SST, $^\circ\text{C}$), V is the magnitude of the vertical wind shear between 850 and 200 hPa (ms^{-1}), and Φ is the latitude ($^\circ$).

Furthermore, to account for the bias of additional predictors used in the regression, the Akaike information criterion (AIC) was used where

$$\text{AIC} = -2L + 2p. \quad (3)$$

In this formula, L is the maximized-log-likelihood of the regression and p is the number of predictors used in the regression (Akaike, 1973). The maximized-log-likelihood determines appropriate beta-coefficients which would allow for the known Poisson distribution of actual MLPS GP to be achieved. As noted in Tippet et al. (2011), the AIC should only be used as a guide to select which combination of variables has the best fit since the AIC, being a function of the data, is random. The lowest AIC value obtained corresponds to the best fit.

In addition to an AIC, the dispersion (σ) – the degree to which the data is spread – was also calculated. As a Poisson regression is one where the variance of the data is equal to its mean, dispersion is simply the ratio of the variance to the mean. If this ratio is greater than one, then the fit is overdispersed and error coefficients need to be increased to compensate for additional errors. Tippet et al. (2011) records σ^2 values, but this paper records σ values to indicate whether a fit is overdispersed or not.

2. Meteorological Data

All climate variable data come from two sources: the monthly mean 40-year European Centre for Medium-Range Weather Forecasts' (ECMWF) Re-Analysis (ERA-40) (Uppala et al., 2005) and the monthly mean ECMWF Re-Analysis ERA-Interim (ERA-INT, Berrisford et al., 2009) datasets with climatological means over the 23 year period 1979-2001 and over the 34 year period 1979-2012, both computed for the 12 month period January through December respectively. These two temporal periods were chosen due to the increased accuracy of satellite-acquired data beginning in 1979 and the availability of complete years of meteorological data from ERA-40 and ERA-INT through 2001 and 2012 respectively. The ERA-40 datasets have a vertical resolution of 60 model levels and 23 pressure levels and a horizontal resolution of $\sim 1.4^\circ \times \sim 1.4^\circ$ while the ERA-INT datasets have a vertical resolution of 60 model levels and 37 pressure levels and a horizontal resolution of $\sim 0.70^\circ \times \sim 0.70^\circ$. The highly topographical nature (Fig. 1) of the Indian Continent region within the AA region required the use of model level datasets for two out of the four variables used in this paper since model levels conform to topography (ECMWF, 2013; Fig. 2).

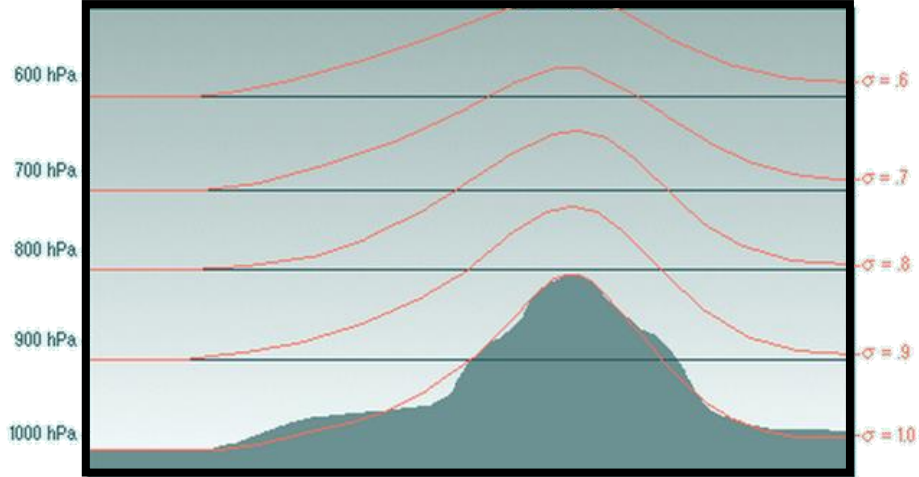


FIG. 2. An example of how model levels conform to topography: the left axis represents pressure levels and the right axis represents model levels. Note how pressure level 900 hPa intersects topography while corresponding model level sigma 0.9 conforms to topography. Sigma is calculated as the ratio of the pressure at any given point to the pressure of the surface directly below it (Allen et al., 2002).

In determining what variables to choose for the GPI, both the Emanuel and Nolan (2004) and the Tippet et al. (2011) variable choices were taken into account. After much trial-and-error, which included permutations of various combinations of variables and the subsequent observations of the spatial and climatological results, four variables were deemed to deliver the best MLPS GPI distribution: estimated convective available potential energy (ECAPE), total column water vapor (TCWV), absolute vorticity (η), and wind shear (V).

a. Estimated Convective Available Potential Energy (ECAPE)

Unlike TC, MLPS have the capability of forming over land. Therefore, the PI variable used in Emanuel and Nolan (2004) and the T variable used in Tippet et al. (2011) are not ideal for this research as they are only defined over the ocean due to their SST dependence. Thus, a simple proxy was used to estimate the energy available for convection over both ocean and land. This proxy estimated the convective available potential energy, here called ECAPE, by subtracting the average saturation moist static energy between 200 hPa and 400 hPa from the moist static energy at the surface so that

$$\text{ECAPE} = (C_p T + L_v Q + GZ)_{\text{surf}} - (C_p T + L_v Q_{\text{sat}} + GZ)_{200\text{hPa}-400\text{hPa}}, \quad (4)$$

where L_v is the latent heat of condensation ($2.501 \times 10^6 \text{ Jkg}^{-1}$), C_p is the specific heat at constant pressure ($1.005.7 \text{ JK}^{-1}\text{kg}^{-1}$), Q is the specific humidity, Q_{sat} is the saturation specific humidity, T is the temperature (K), and GZ is the geopotential height (Jkg^{-1}). ECAPE is expressed in units

of K by dividing ECAPE by C_p . Due to the height at which the saturation value of moist static energy is calculated, model levels were not needed as AA topography does not reach that height.

b. Total Column Water Vapor (TCWV)

This variable represents the column integrated water vapor content (i.e. the height of water that would exist if all water vapor in a column were precipitated as rain). It is commonly called precipitable water in the literature and is expressed in kgm^{-2} . Both the Emanuel and Nolan (2004) and the Tippet et al. (2011) GPI use the variable relative humidity as a predictor. Here, TCWV is used instead of relative humidity since TCWV yielded a lower AIC value in regression trials and more robust spatial and climatological results.

c. Absolute Vorticity (η)

Values for η , expressed in 10^5s^{-1} , were derived by taking the absolute value of the summation of the Coriolis factor ($f = 2\Omega\sin\Phi$) and the ERA-40 derived relative vorticity values. Relative vorticity was retrieved at model level 49, the model level equivalent to the 850 hPa pressure level as used in the GPI established by Emanuel and Nolan (2004) and used in Tippet et al. (2011).

d. Wind Shear (V)

Wind Shear (V), expressed in ms^{-1} , was calculated by solving for the total vector shear defined as

$$V = \sqrt{(u_{49} - u_{30})^2 + (v_{49} - v_{30})^2}, \quad (5)$$

where the entity under the square root symbol stands for the difference between horizontal winds at the 49 and 30 model level. These two model levels are equivalent to the 850 hPa and 200 hPa levels as used in the GPI established by Emanuel and Nolan (2004) and used in Tippet et al. (2011).

3. Genesis Points (GP)

Two different sources of genesis points (GP) were available for this research: a) GP digitized from the subjective analyses by Mooley and Shukla (1987) and Sikka (2006) covering the Indian Continent region from June through September 1888 – 2003 and b) GP compiled from the objective analysis by Yale Geology & Geophysics' Postdoc John Hurley covering the globe from January through December 1979 - 2012. Ultimately, it was determined that the Hurley database was to be used for this research due to the objectivity implemented in its compilation and the completeness of the database, as it covers a global spatial scale and incorporates all months of each year. Still, the Mooley and Shukla (1987) and Sikka (2006) databases are detailed below.

a. Digitizing Mooley and Shukla (1987) and Sikka (2006)

MLPS track data were retrieved from Mooley and Shukla (1987) and Sikka (2006). Identification of track data was achieved through the analysis of “the morning (0300 UTC) sea level synoptic pressure analysis” daily weather charts put forth by the India Meteorological Department (Ajayamohan, 2010). Lim and Simmonds (2007) believed that although subjectivity did occur in these analyses, they provided more accurate accounts of MLPS occurrences than the “reanalyses used by automated cyclone identification procedures.” Track data were categorized according to intensity, with values of 1 through 5 given to low, depression, deep depression, cyclonic storm, and severe cyclonic storm respectively (Ajayamohan, 2010). Other information included with the track data were dates of the MLPS genesis, dates of the subsequent track of the MLPS, and corresponding latitude-longitude coordinates for both genesis and track path. Both Mooley and Shukla (1987) and Sikka (2006) provided data for the months of June through September, the months of the boreal monsoon, in the Indian Continent region. While Mooley and Shukla (1987) provided undigitized MLPS track data from 1888-1983 and Sikka (2006) provided undigitized MLPS track data from 1984-2003, for this research MLPS data needed to be and therefore was digitized for and restricted to the years 1979-2001, the first of the two aforementioned temporal periods. MLPS GP were then isolated (Fig. 3). Other climatologies that exist include Chen and Weng (1999) and Saha et al. (1981). Yet, since they are temporally and spatially similar to the Mooley and Shukla (1987) dataset and produce similar climatologies, they were not used in this research.

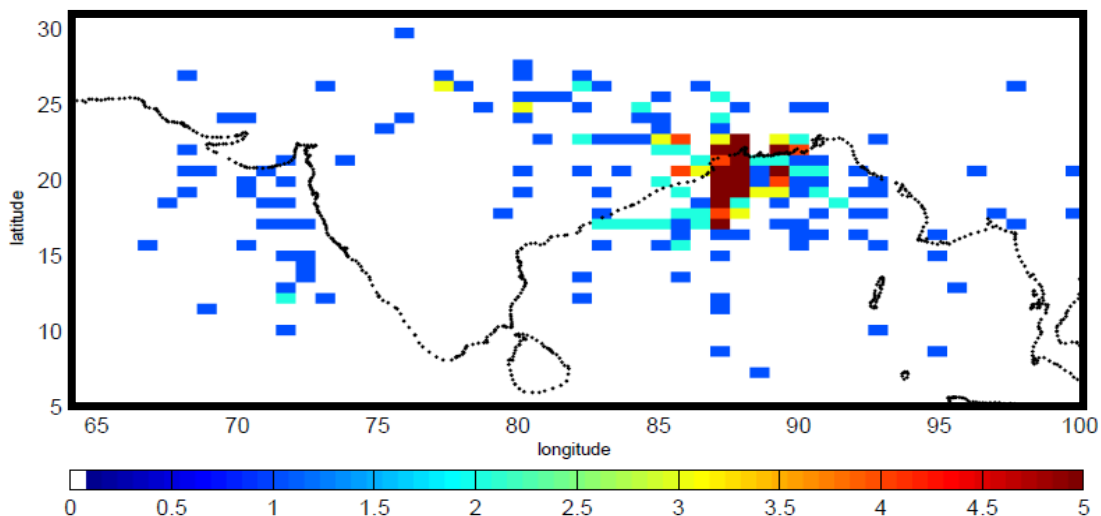


FIG. 3. The GP density distribution of MLPS from June-September 1979-2001 in the Indian Continent region.

b. Compiling MLPS from Hurley (2014) Database

At the start of this research, there was no existing database of MLPS located outside of the Indian Continent region. Therefore, as part of a separate initiative to create such a database, for the reasons previously stated (see Section II), a climatology of global MLPS was created and compiled by Postdoc John Hurley by using an automatic feature tracking algorithm developed by Hodges (1995). Hurley adapted this algorithm, called TRACK, to pinpoint cyclonic vorticity maxima in the ERA-Interim reanalysis on a global scale. From his work, there now exists a global database of MLPS GP and their associated tracks. The resulting GP and track data were categorized according to intensity, with values of 1 through 3 given to low, depression, as well as deep depression and stronger respectively. Other data included dates of the MLPS genesis, dates of the subsequent track of the MLPS, and corresponding latitude-longitude coordinates for both genesis and track path. Data were provided from January 1979 through December 2012. As with the Mooley and Shukla (1987) and Sikka (2006) data, MLPS data were restricted to and GP were isolated for the years 1979-2001 but also for 1979-2012 (Fig. 4) to encompass both of the aforementioned temporal time periods.

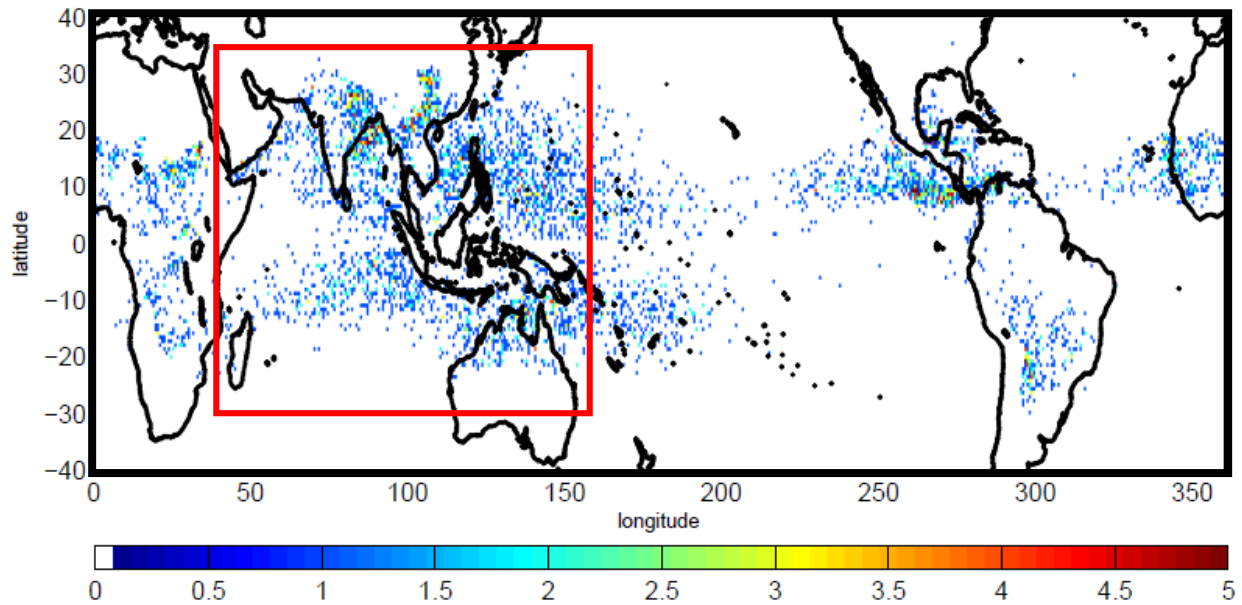


FIG. 4. Global GP density distribution from January 1979-December 2012. The red box indicates the AA region. Represented here are only those storms which have a precipitable water content of 35 kgm^{-2} or higher (Section IV).

B. Refinement and Exploration

Through variable analysis, two different types of MLPS were identified by TRACK: those that formed under dry conditions and those that formed under moist conditions. Since this research is interested in determining genesis for those MLPS which contribute almost half of the total seasonal precipitation to monsoonal regions, a value which robustly separates dry storms from moist storms needed to be found. Histograms were therefore created from points of monthly mean TCWV values which corresponded to points of genesis (Fig. 5). The bimodal distribution created shows the dry and moist MLPS are separated at about 35 kgm^{-2} . Therefore, for this research all GP which formed at points where TCWV were less than 35 kgm^{-2} were removed from the index. This procedure was also performed with daily data and the minimum between the bimodal peaks was also observed at 35 kgm^{-2} .

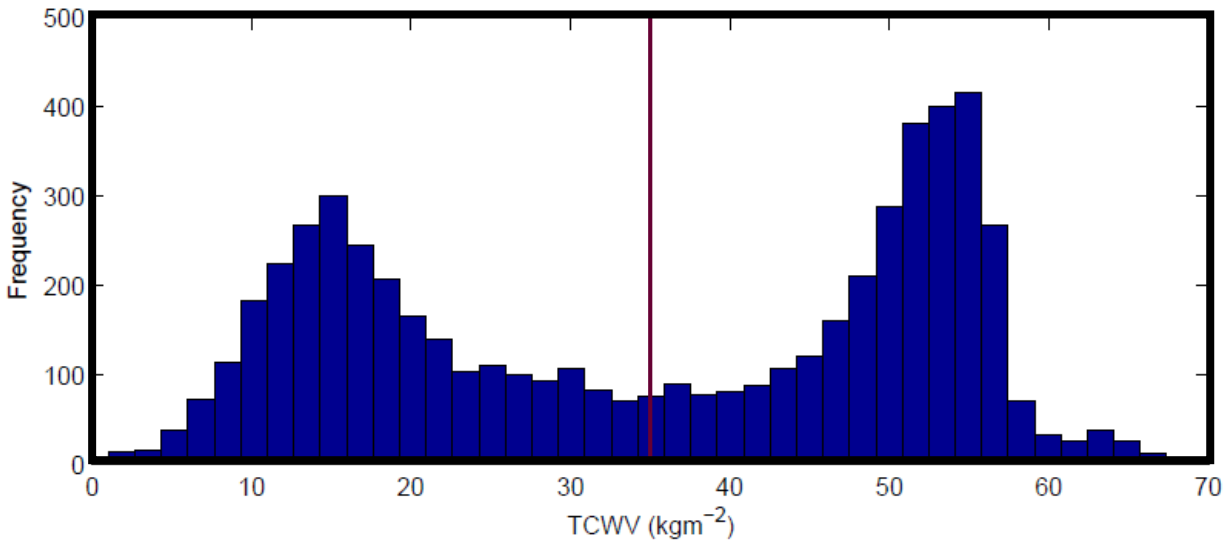


FIG. 5. A PDF displaying the density distribution of monthly mean TCWV in the AA region. Note the bimodal distribution with one peak centered around 15 kgm^{-2} and the other peak centered around 53 kgm^{-2} .

After the dry storms were removed from the index, histograms were once again created, this time for each of the four variables at points of genesis and points of non-genesis separated by eventual attainment of intensity (Fig. 6). From observational analysis of the resulting histograms, it appears as though the genesis points associated with MLPS evolving into deep depressions and stronger form at higher ECAPE and TCWV values and lower WS values than genesis points associated with MLPS evolving into depressions or lows. Depressions are shown to form at higher ECAPE and TCWV values and lower WS values than those genesis points associated with MLPS evolving into lows.

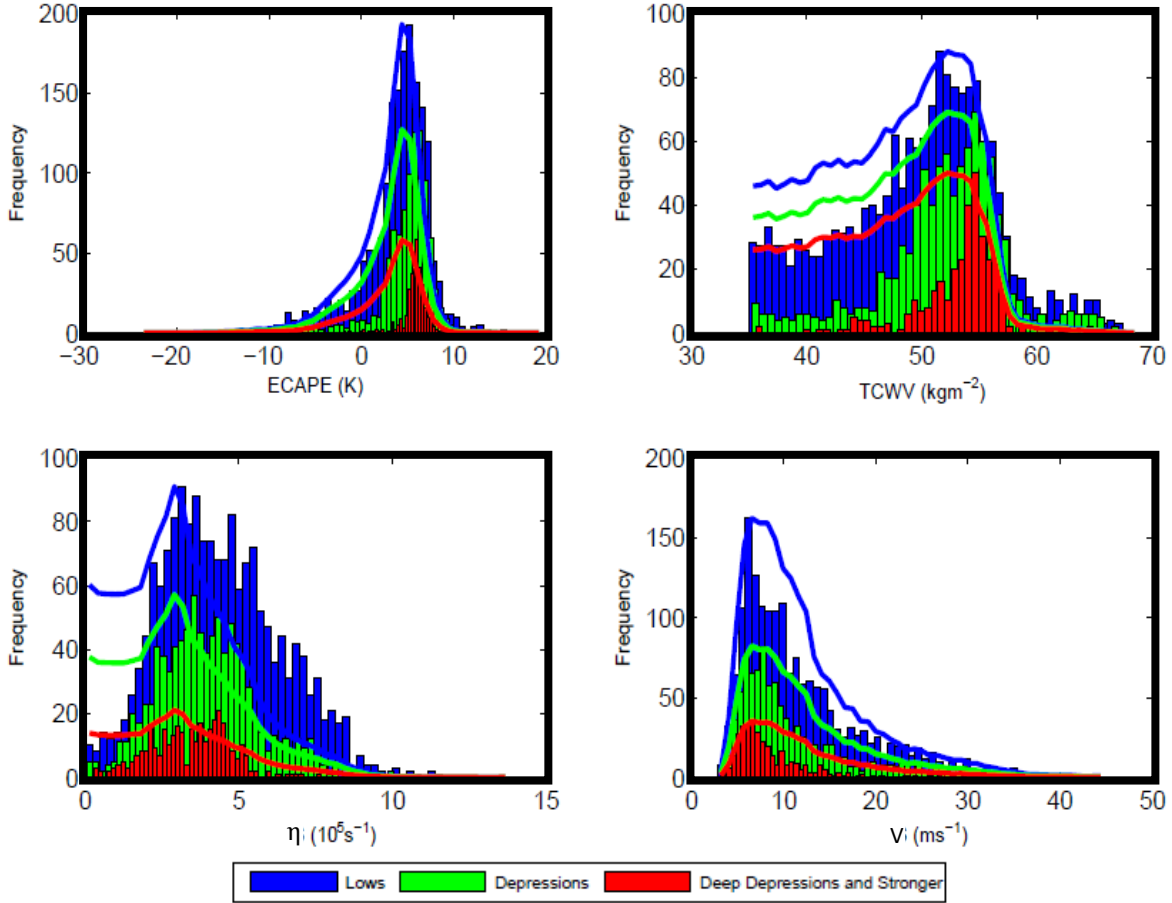


FIG. 6. PDFs displaying the density distribution by intensity of each of the four meteorological variables used. The bars indicate frequency at points of genesis. The lines indicate frequency at points of non-genesis, normalized to the height of the corresponding color's bars.

C. Construction

A GPI for MLPS was constructed by implementing a Poisson regression on the results from the TRACK program's identification of global MLPS and the four aforementioned meteorological variables. All data were geographically restricted to the AA region defined in Section II.

1. Executing the Poisson Regression: The Statistics

A Poisson regression was run on 18 combinations of variables using first the ERA-40 and then the ERA-INT datasets. Combinations included Poisson regression runs on: one variable (Trials 1-4), two variables (Trials 5-10), three variables (Trials 11-14), each level of attained intensity (Trials 15-17), and all four variables (Trial 18). Resulting beta coefficients, AIC, dispersion, and standard errors can be seen below (Tabs. 1A and 1B). Correlation coefficients between estimated GP and actual GP for the Northern and Southern Hemispheres were calculated and provided as well.

| Trial | Constant | Heat | Humidity | Vorticity | Shear | Statistics | |
|------------------------|-----------------------|------------------------------|------------------------------|------------------------------|-----------------------------|----------------------------|----------------|
| # | b | b_{CAPE} | b_{TCWV} | b_{η} | b_{ν} | σ | AIC |
| Coefficients | | | | | | σ | AIC |
| 1 | -4.2607 | 0.1795 | ----- | ----- | ----- | 1.0456 | 12939 |
| 2 | -8.7581 | ----- | 0.1239 | ----- | ----- | 0.9974 | 12234 |
| 3 | -3.1070 | ----- | ----- | -0.0054 | ----- | 1.1753 | 15601 |
| 4 | -1.8033 | ----- | ----- | ----- | -0.0889 | 1.1412 | 14180 |
| 5 | -8.3193 | 0.0352 | 0.1089 | ----- | ----- | 1.0030 | 12218 |
| 6 | -5.6568 | 0.2014 | ----- | 0.3209 | ----- | 0.9272 | 12017 |
| 7 | -4.0309 | 0.1730 | ----- | ----- | -0.0142 | 1.0421 | 12919 |
| 8 | -10.3188 | ----- | 0.1324 | 0.3091 | ----- | 0.8695 | 11239 |
| 9 | -8.3981 | ----- | 0.1207 | ----- | -0.0167 | 0.9958 | 12206 |
| 10 | -2.3585 | ----- | ----- | 0.2009 | -0.1086 | 1.1054 | 13776 |
| 11 | -9.3900 | 0.0837 | 0.0967 | 0.3407 | ----- | 0.8751 | 11115 |
| 12 | -8.1631 | 0.0243 | 0.1111 | ----- | -0.0142 | 0.9993 | 12199 |
| 13 | -5.3472 | 0.1917 | ----- | 0.3275 | -0.0207 | 0.9186 | 11976 |
| 14 | -9.7489 | ----- | 0.1269 | 0.3253 | -0.0293 | 0.8600 | 11148 |
| 15 | -9.1089 | 0.0787 | 0.0781 | 0.4109 | -0.0108 | 0.8800 | 8091 |
| 16 | -12.0351 | 0.0704 | 0.1379 | 0.2678 | -0.0236 | 0.8174 | 4583 |
| 17 | -17.5105 | 0.2518 | 0.2200 | 0.0803 | -0.1603 | 1.0228 | 1871 |
| 18 | -9.2058 | 0.0690 | 0.1000 | 0.3447 | -0.0195 | 0.8664 | 11079 |
| Standard Errors | | | | | | NH Corr | SH Corr |
| 1 | 0.0548 | 0.0053 | ----- | ----- | ----- | 0.8597 | 0.8012 |
| 2 | 0.1525 | ----- | 0.0029 | ----- | ----- | 0.9920 | 0.8798 |
| 3 | 0.0524 | ----- | ----- | 0.0107 | ----- | -0.9823 | -0.8541 |
| 4 | 0.0455 | ----- | ----- | ----- | 0.0034 | 0.6890 | 0.8634 |
| 5 | 0.1877 | 0.0089 | 0.0047 | ----- | ----- | 0.9896 | 0.8704 |
| 6 | 0.0638 | 0.0039 | ----- | 0.0096 | ----- | 0.9566 | 0.8666 |
| 7 | 0.0783 | 0.0057 | ----- | ----- | 0.0034 | 0.8213 | 0.7983 |
| 8 | 0.1184 | ----- | 0.0021 | 0.0083 | ----- | 0.9678 | 0.9101 |
| 9 | 0.1718 | ----- | 0.0031 | ----- | 0.0032 | 0.9868 | 0.8676 |
| 10 | 0.0553 | ----- | ----- | 0.0107 | 0.0034 | 0.9437 | 0.9276 |
| 11 | 0.1375 | 0.0066 | 0.0034 | 0.0088 | ----- | 0.9723 | 0.9017 |
| 12 | 0.1937 | 0.0094 | 0.0048 | ----- | 0.0034 | 0.9853 | 0.8630 |
| 13 | 0.0784 | 0.0043 | ----- | 0.0096 | 0.0030 | 0.9564 | 0.8780 |
| 14 | 0.1344 | ----- | 0.0023 | 0.0084 | 0.0028 | 0.9756 | 0.9127 |
| 15 | 0.1717 | 0.0081 | 0.0040 | 0.0117 | 0.0036 | 0.9611 | 0.8872 |
| 16 | 0.2832 | 0.0149 | 0.0070 | 0.0151 | 0.0052 | 0.9265 | 0.8951 |
| 17 | 0.9919 | 0.0444 | 0.0196 | 0.0363 | 0.0187 | 0.7914 | 0.7323 |
| 18 | 0.1430 | 0.0071 | 0.0035 | 0.0088 | 0.0029 | 0.9763 | 0.9057 |

TAB. 1A. Beta coefficient results and statistics from the Poisson regression between MLPS GP and ERA-40 meteorological variables. Entries with “-----” indicate that the respective variable was not included in the regression.

| Trial | Constant | Heat | Humidity | Vorticity | Shear | Statistics | |
|------------------------|-----------------------|-------------------------------|------------------------------|------------------------------|-------------------------|----------------------------|----------------|
| # | b | b_{ECAPE} | b_{TCWV} | b_{η} | b_V | | |
| Coefficients | | | | | | σ | AIC |
| 1 | -4.2289 | 0.1821 | ----- | ----- | ----- | 0.9660 | 25319 |
| 2 | -9.6964 | ----- | 0.1256 | ----- | ----- | 0.9178 | 24465 |
| 3 | -4.0935 | ----- | ----- | 0.0038 | ----- | 1.0892 | 29659 |
| 4 | -2.7975 | ----- | ----- | ----- | -0.0847 | 1.0589 | 27590 |
| 5 | -8.6736 | 0.0554 | 0.1010 | ----- | ----- | 0.9195 | 24396 |
| 6 | -5.3292 | 0.1866 | ----- | 0.2926 | ----- | 0.8431 | 23957 |
| 7 | -4.1254 | 0.1796 | ----- | ----- | -0.0079 | 0.9689 | 25310 |
| 8 | -10.9494 | ----- | 0.1299 | 0.2832 | ----- | 0.8210 | 23067 |
| 9 | -9.3415 | ----- | 0.1227 | ----- | -0.0173 | 0.9183 | 24418 |
| 10 | -3.3485 | ----- | ----- | 0.1929 | -0.1025 | 1.0307 | 27000 |
| 11 | -9.3007 | 0.0925 | 0.0875 | 0.3174 | ----- | 0.8137 | 22777 |
| 12 | -8.6036 | 0.0472 | 0.1032 | ----- | -0.0122 | 0.9205 | 24374 |
| 13 | -5.1980 | 0.1827 | ----- | 0.2965 | -0.0111 | 0.8422 | 23937 |
| 14 | -10.4144 | ----- | 0.1246 | 0.2948 | -0.0259 | 0.8102 | 22952 |
| 15 | -9.3337 | 0.0860 | 0.0724 | 0.3858 | -0.0035 | 0.8305 | 15494 |
| 16 | -11.6099 | 0.1120 | 0.1170 | 0.2414 | -0.0191 | 0.7720 | 8660 |
| 17 | -14.1791 | 0.2685 | 0.1648 | 0.0443 | -0.1397 | 1.6111 | 3465 |
| 18 | -9.2115 | 0.0848 | 0.0893 | 0.3196 | -0.0131 | 0.8104 | 22754 |
| Standard Errors | | | | | | NH Corr | SH Corr |
| 1 | 0.0238 | 0.0041 | ----- | ----- | ----- | 0.9315 | 0.8651 |
| 2 | 0.1120 | ----- | 0.0022 | ----- | ----- | 0.9888 | 0.8958 |
| 3 | 0.0392 | ----- | ----- | 0.0079 | ----- | 0.9809 | 0.9073 |
| 4 | 0.0338 | ----- | ----- | ----- | 0.0025 | 0.7632 | 0.8940 |
| 5 | 0.1577 | 0.0064 | 0.0035 | ----- | ----- | 0.9902 | 0.8922 |
| 6 | 0.0336 | 0.0028 | ----- | 0.0064 | ----- | 0.9729 | 0.9092 |
| 7 | 0.0405 | 0.0043 | ----- | ----- | 0.0025 | 0.9191 | 0.8610 |
| 8 | 0.0862 | ----- | 0.0016 | 0.0060 | ----- | 0.9481 | 0.9192 |
| 9 | 0.1256 | ----- | 0.0023 | ----- | 0.0024 | 0.9911 | 0.8831 |
| 10 | 0.0412 | ----- | ----- | 0.0078 | 0.0025 | 0.9584 | 0.9479 |
| 11 | 0.1115 | 0.0045 | 0.0024 | 0.0062 | ----- | 0.9573 | 0.9200 |
| 12 | 0.1605 | 0.0067 | 0.0036 | ----- | 0.0025 | 0.9905 | 0.8839 |
| 13 | 0.0421 | 0.0030 | ----- | 0.0064 | 0.0021 | 0.9748 | 0.9122 |
| 14 | 0.0979 | ----- | 0.0017 | 0.0060 | 0.0020 | 0.9586 | 0.9200 |
| 15 | 0.1400 | 0.0054 | 0.0030 | 0.0082 | 0.0027 | 0.9451 | 0.9204 |
| 16 | 0.2250 | 0.0108 | 0.0050 | 0.0113 | 0.0039 | 0.9292 | 0.9028 |
| 17 | 1.0344 | 0.0544 | 0.0211 | 0.0463 | 0.0224 | 0.8455 | 0.7781 |
| 18 | 0.1139 | 0.0048 | 0.0025 | 0.0062 | 0.0021 | 0.9613 | 0.9207 |

TAB. 1B. Beta coefficient results and statistics from the Poisson regression between MLPS GP and ERA-INT meteorological variables. Entries with “-----” indicate that the respective variable was not included in the regression.

An analysis of the resulting spatial and climatological distributions of all runs showed that Trial 18 from both Tabs. 1A and 1B provided the best and closest fit. Additionally, the lowest AIC value was obtained when running the Poisson regression with all four variables. This combined with a dispersion value less than “1” made the results from Trials 18 ideal for a MLPS GPI in the AA region. This paper uses the results obtained with the ERA-INT dataset (Trial 18, Tab. 1B) since compared to the ERA-40 dataset, the ERA-INT dataset is more recent, has a higher spatial resolution, and has corrected many existing problems found within the ERA-40 dataset (Berrisford et al., 2009). Trial 18 from Tab. 1B will hereinafter be referenced as just Trial 18. Trial 18 demonstrates that low-level absolute vorticity has the strongest influence on MLPS genesis. Additionally, only one of the four variables has a negative coefficient – total shear – which implies that shear represses MLPS genesis.

2. Spatial Distribution

Using the coefficients in Trial 18, the corresponding spatial distribution with overlaying GP was produced (Fig. 7). Comparing the results in Fig. 7 to the overlaying GP points and to the density distribution in Fig. 4, the GPI was able to replicate the spatial distribution of GP. Describing the larger AA region by sections allows for a more thorough investigation of how well the generated GPI reproduced the actual spatial distribution of GP.

ASIA: The GPI distribution in the Western Pacific tilts to the southeast, and it is similar in intensity to the GP density distribution. In the western Pacific, the GPI was able to reproduce the tongue-esque distribution of GP. Over northern China, the GPI overestimates intensity while over southern China, the GPI underestimates intensity.

INDIA: The GPI distribution in both the Bay of Bengal and the eastern Arabian Sea replicates the spatial distribution and the intensity of GP. Additionally, the GPI over India replicates the spatial and intensity distribution. The GPI was unable to replicate the western Arabian GP distribution, however.

OCEAN (SOUTHERN): The GPI is able to replicate the approximately ten degree band of GP which arcs from Madagascar to western Australia.

AUSTRALIA: Australia’s northern third is covered with GP. However, the GPI only covers the northern coastline of Australia. In the surrounding waters, the GPI displays a better representation of GP spatial distribution and intensity.

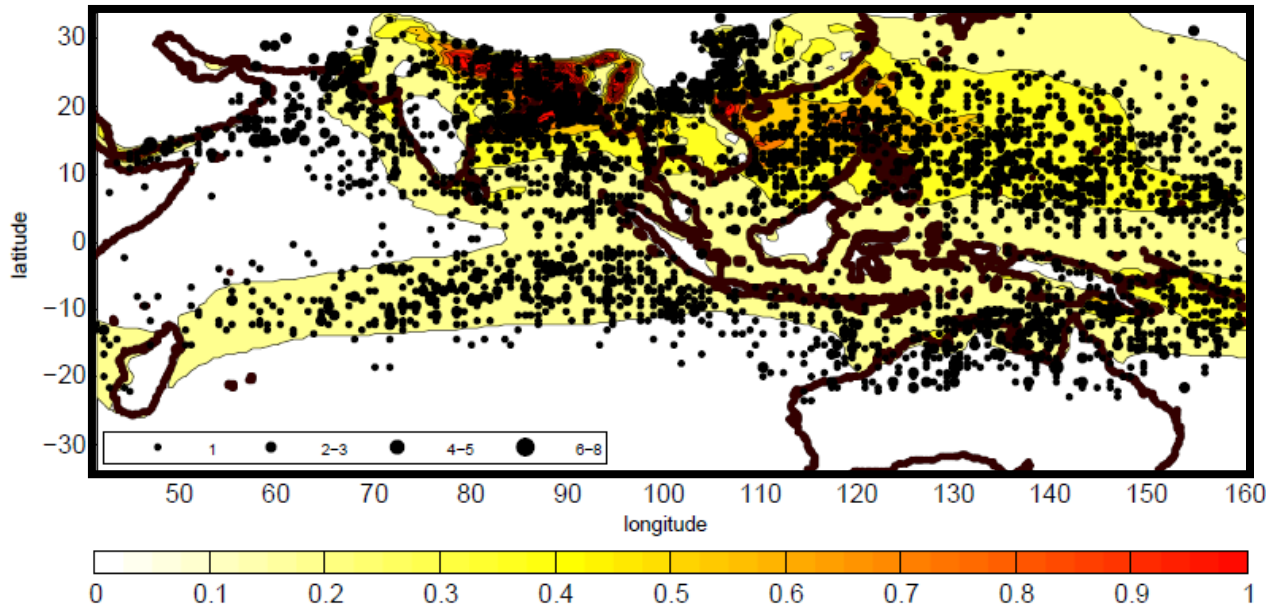


FIG. 7. Spatial distribution of the MLPS GPI in the AA region (shaded) based on Trial 18 coefficients for the 34 year period of January 1979 – December 2012. GP distribution (black dots) is overlaid on top of the GPI distribution with varying sizes corresponding to differing numbers of GP occurring at any given point. The density values for the four different sized black dots are shown in the legend.

Overall, the index does a very good job in reproducing the Indian, Asian, and Southern Ocean distributions and a good job in reproducing the Australian distribution.

3. Zonal Distribution

The spatial distribution observed in Fig. 7 presented itself in two strong zonal bands –north of the equator and south of the equator – with the equator having a noticeable lack of GP. The zonal mean distribution is displayed in Fig. 8 with both the Southern and Northern hemispheric GP means aligning closely to the projected GPI zonal means with a correlation of 0.8802.

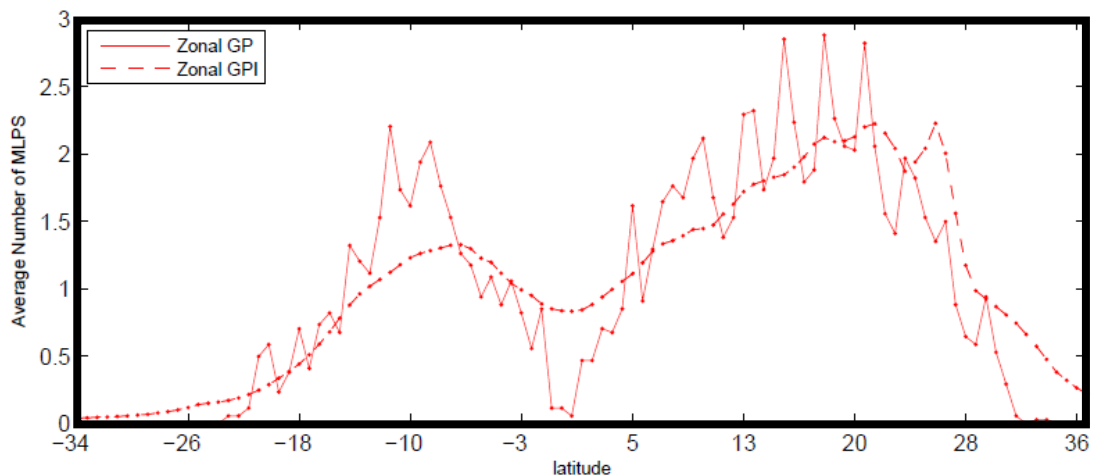


FIG. 8. The zonal mean from the MLPS GPI (red dashed) and the actual TRACK-determined climatology (red solid) for the AA region.

4. Seasonal Climatology

Climatologically, the GPI is able to closely replicate the frequency of MLPS GP over the 34 year period (Fig. 9). It depicts the increase in MLPS GP during boreal summer and the decrease in boreal winter. As seen in Trial 18, correlations between the actual number of GP and the estimated number of GP from the GPI have a 0.9613 correlation for the Northern Hemisphere and a 0.9207 correlation for the Southern Hemisphere.

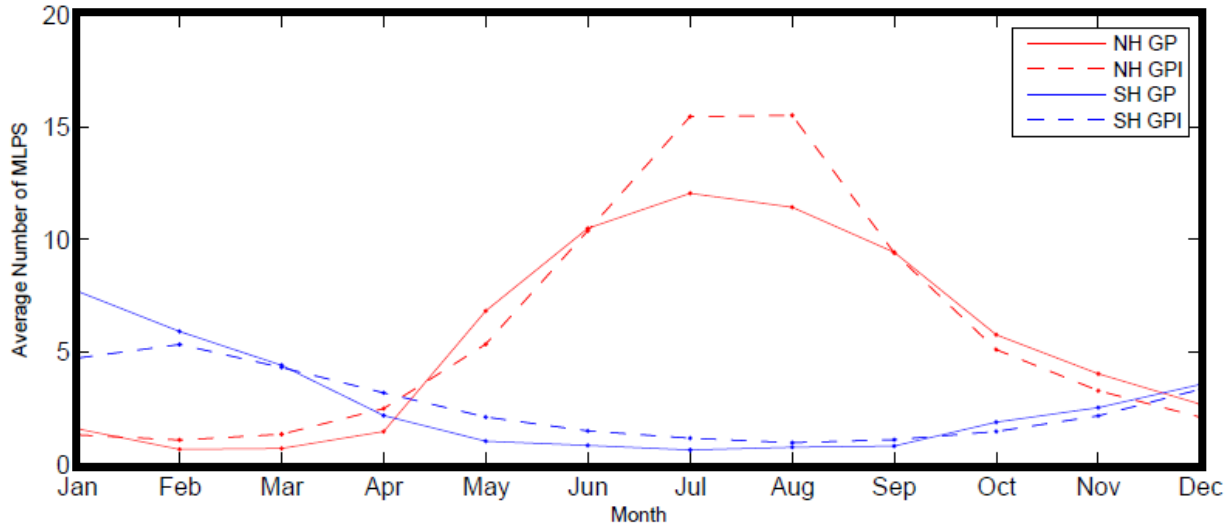


FIG. 9. The climatology from the MLPS GPI in the AA region is shown for the Northern (red dashed) and Southern (blue dashed) hemispheres. These values are compared to the actual TRACK-determined climatology for the Northern (red solid) and Southern (blue solid) Hemispheres.

Additionally, the GPI predicts that the Southern Hemisphere had an average of about 118 GP less than the Northern Hemisphere, close to the actual 99 GP difference between the Northern and Southern Hemispheres. More statistics (Tab. 2) display the magnitude of difference between the actual GP count and the estimated GPI count.

| Difference | Northern Hemisphere | Southern Hemisphere |
|------------|---------------------|---------------------|
| Mean | 38 | 24 |
| Minimum | 1 | 4 |
| Maximum | 139 | 101 |

TAB. 2. Calculated from Fig. 9, the difference column in the above Table represents the mean, minimum, or maximum difference between the actual GP counts (solid lines, Fig. 9) and the estimated GPI counts (dashed lines, Fig. 9).

5. Interannual Variability

One way to test whether the Trial 18 coefficients are robust is to use those coefficients in conjunction with interannual data to try and recreate a similar spatial and a well-produced climatological distribution of MLPS. Three different methodologies were implemented: 1) using

the Trial 18 coefficients on data for all months covering the period January 1979 through December 2012 inclusively rather than on monthly mean data, 2) using the Trial 18 coefficients on data for the boreal and austral five-month summers of 1979 through 2012 inclusively rather than on data for all months, and 3) using the Trial 18 coefficients on the mean boreal and austral five-month summers from 1979-2012 inclusive. Spatially, each of these three methodologies performed well and was able to closely replicate the climatological monthly-mean spatial distribution as shown in Fig. 7. The spatial distribution observed once again presented itself in two strong zonal bands – one north of the equator and one south of the equator. The zonal mean distribution for each of the three methods performed aligned closely to the projected GPI zonal means. However, the interannual climatology in each of the three methods underperformed, as shown by the correlation coefficients in Tab. 3. Although values of the projected GPI were close to the actual GP distribution per year, the GPI lacked the magnitude of oscillations found in the GP distribution.

| Methodology | NH Corr | SH Corr |
|-------------|---------|---------|
| 1 | 0.1261 | 0.5923 |
| 2 | 0.0742 | 0.5847 |
| 3 | 0.0254 | 0.5337 |

TAB. 3. Correlation coefficients calculated from the three methodologies detailed in Section IIIC5. SH GP counts and estimated GPI counts were much more correlated than that of the NH GP counts and estimated GPI counts.

IV. An Indian MLPS GPI

With the successful implementation of a Poisson regression to generate a GPI for MLPS over the AA region, the same regression was once again performed in the same manner as explained above but now for the geographically restricted region of the Indian Continent (5°N to 30°N and 64°E to 100°E). The purpose behind this execution was to determine whether the MLPS which form over India do so under similar or different conditions than the MLPS that form in the greater AA region. This can be determined by examining the similarity between the results generated in the Indian Continent region to the results in Trial 18 corresponding to the AA region (Tab. 4).

A. Executing the Poisson Regression: The Statistics

From the statistics generated, a few important similarities appear. In both the AA region and the Indian Continent region, wind shear has a negative effect on the formation of MLPS, where the stronger the wind shear, the fewer the number of MLPS. Yet, the wind shear component from the Indian GPI is not that statistically robust since the corresponding error coefficient is larger in magnitude than the actual coefficient. For each of the other three variables in the index, positive

relationships were found in both regions, with low-level absolute vorticity having the largest influence on MLPS genesis. The magnitudes of each respective variable from both the AA and Indian GPI are similar, implying that MLPS which form over the Indian Continent region are formed under similar conditions as those that form in the entire AA region.

| <u>Trial</u> | <u>Constant</u> | <u>Heat</u> | <u>Humidity</u> | <u>Vorticity</u> | <u>Shear</u> | <u>Statistics</u> | |
|------------------------|-----------------|-------------|-----------------|------------------|--------------|-------------------|----------------|
| <u>Location</u> | b | b_{ECAPE} | b_{TCWV} | b_{η} | b_V | σ | AIC |
| Coefficients | | | | | | | |
| AA | -9.2115 | 0.0848 | 0.0893 | 0.3196 | -0.0131 | 0.8104 | 22754 |
| India | -8.0973 | 0.0735 | 0.0671 | 0.3212 | -0.0011 | 0.8430 | 4440 |
| Standard Errors | | | | | | NH Corr | SH Corr |
| AA | 0.1139 | 0.0048 | 0.0025 | 0.0062 | 0.0021 | 0.9613 | 0.9207 |
| India | 0.1817 | 0.0112 | 0.0044 | 0.0157 | 0.0038 | 0.9537 | ----- |

TAB. 4. Beta coefficient results from the Poisson regression between MLPS GP and meteorological variables. The first row corresponds to the Trial 18 results from Tab. 1B. The second line depicts the results from executing a Poisson regression on the four variables excluding dry GP in the Indian Continent region.

B. Spatial Distribution

Spatially, the GPI was able to replicate the distribution of MLPS GP over India, as shown in Fig. 4 and by the overlaying GP in Fig. 10. The maximum density boomerang shape over the Northern Bay of Bengal is consistent with the arched, observed distribution. The GPI performs well over the Indian Continent interior and also mimics the southwest band of storms off India’s western coast. Yet, the GPI underperforms to the northeast of Nepal, where it projects GP where there are none.

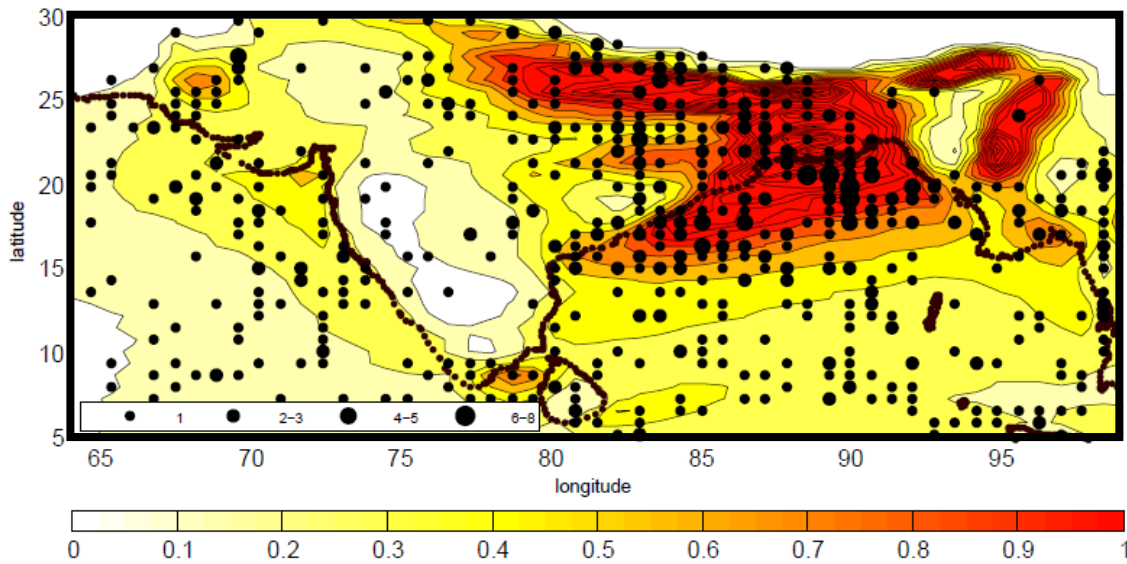


FIG. 10. Spatial distribution of the MLPS GPI in the Indian Continent region (shaded) based on Tab. 3 coefficients for the Indian Continent region over the 34 year period of January 1979 – December 2012. GP distribution is overlaid on top of the GPI distribution (black dots) with varying sizes corresponding to differing numbers of GP occurring at each point. The density values for the four different sized black dots are shown in the legend.

C. Seasonal Climatology

Once again, the GPI is able to climatologically replicate the MLPS GP frequency, this time over the Indian Continent region over the 34 year period (Fig. 11). As seen in Tab. 3, the correlation between the actual number of GP and the estimated number of GP from the GPI is 0.9537.

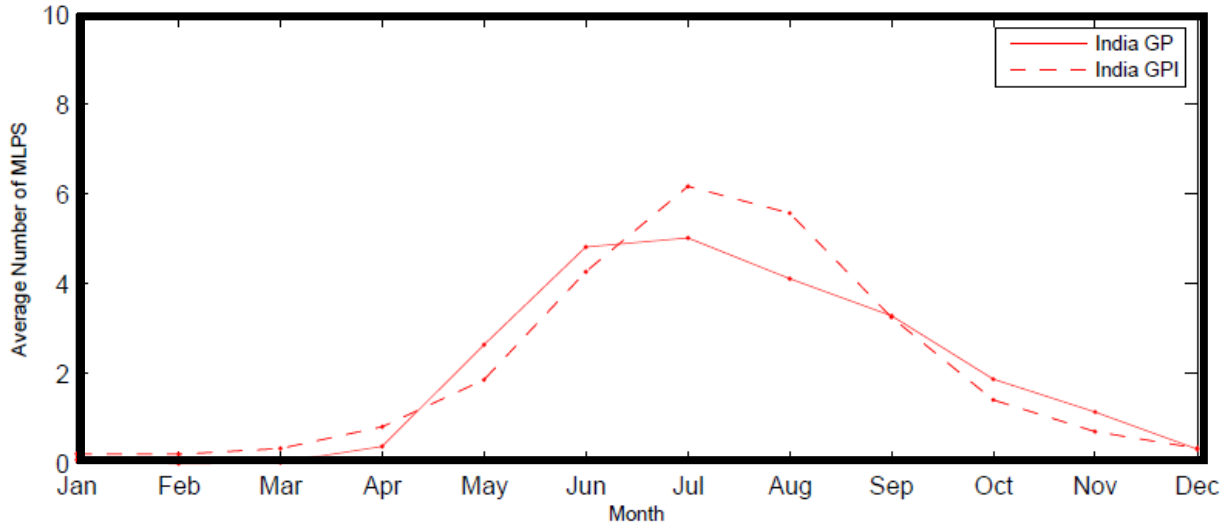


FIG. 11. The climatology from the MLPS GPI in the Indian Continent region is shown for the Northern (red dashed) and Southern (blue dashed) hemispheres. These values are compared to the actual TRACK determined climatology for the Northern (red solid) and Southern (blue solid) Hemispheres.

Additional statistics (Tab. 5) display the magnitude of difference between the actual GP count and the estimated GPI count.

| Difference | Indian Continent region |
|------------|-------------------------|
| Mean | 17 |
| Minimum | 0 |
| Maximum | 50 |

TAB. 5. Calculated from Fig. 10, the difference column in the above Table represents the mean, minimum, or maximum difference between the actual GP counts (solid lines, Fig. 11) and the Estimated GPI counts (dashed lines, Fig. 11).

V. A Global MLPS GPI

This research recognized the value in having a global MLPS GP distribution, a scope not originally undertaken due to the breadth of the region and its different climates, topographical regions, and weather patterns. If a robust global MLPS GP distribution could be achieved and if this distribution were similar to those of the AA and Indian Continent regions, then it could be concluded that MLPS GP are governed by similar meteorological triggers worldwide.

A. Executing the Poisson Regression: The Statistics

On a global scale, the Poisson regression statistics (Tab. 6) return similar beta coefficients as in the AA and Indian region cases. The absolute vorticity contribution was weaker when performed on a global scale but this was compensated by a larger contribution from ECAPE. Once again, total shear seemed to inhibit MLPS formation.

| <u>Trial</u> | <u>Constant</u> | <u>Heat</u> | <u>Humidity</u> | <u>Vorticity</u> | <u>Shear</u> | <u>Statistics</u> | |
|------------------------|-----------------|-------------|-----------------|------------------|--------------|----------------------------|----------------|
| # | b | b_{ECAPE} | b_{TCWV} | b_{η} | b_{ν} | | |
| Coefficients | | | | | | σ | AIC |
| Global | -9.7375 | 0.1273 | 0.0968 | 0.4053 | -0.0263 | 0.5241 | 43666 |
| AA | -9.2115 | 0.0848 | 0.0893 | 0.3196 | -0.0131 | 0.8104 | 22754 |
| India | -8.0973 | 0.0735 | 0.0671 | 0.3212 | -0.0011 | 0.8430 | 4440 |
| Standard Errors | | | | | | NH Corr | SH Corr |
| Global | 0.0570 | 0.0020 | 0.0012 | 0.0038 | 0.0013 | 0.9851 | 0.8873 |
| AA | 0.1139 | 0.0048 | 0.0025 | 0.0062 | 0.0021 | 0.9613 | 0.9207 |
| India | 0.1817 | 0.0112 | 0.0044 | 0.0157 | 0.0038 | 0.9537 | ----- |

TAB. 6. Beta coefficient results from the Poisson regression between MLPS GP and meteorological variables. The first row corresponds to the results from executing a Poisson regression on the four variables excluding dry GP over the entire globe. The second and third rows are the results shown in Tab. 4, with the AA results being those from Trial 18 in Tab. 1B.

B. Spatial Distribution

Global spatial coverage performed almost as well as the spatial coverage in the AA and Indian cases. Breaking down the globe into its respective continents and major oceans from west to east in Fig. 12 allows for a more in-depth analysis of the distribution.

EAST AFRICA: The GPI does not perform well over eastern Africa. It barely displays GP events near the eastern coastline. Rather the focus of the GPI is on the western coastline.

INDIA AND ASIA: The spatial distribution in these regions are robust. There is one band of GP that was not able to be reproduced, located directly to the east of the northern Bay of Bengal and directed northeastward.

INDIAN OCEAN – ARABIAN SEA AND BAY OF BENGAL: With most of the GP hugging the shoreline in the Arabian Sea, the GPI was able to display this distribution while also displaying a broader stretching distribution in the Bay of Bengal.

OCEAN (SOUTHERN): As before, the distribution in the Southern Ocean is confined to a ten degree band, with the inner five degrees having a higher concentration of GP. The GPI index was able to replicate this pattern.

AUSTRALIA: The GPI was unable to reproduce the southern extent of the GP distribution. Yet, it was able to spatially show the GP distribution off the northern coastline.

PACIFIC OCEAN: Reproducing the approximate shape of GP distribution, the GPI was concentrated around a ten degree band located just north of the equator from about the middle to eastern Pacific. In the Pacific Warm Pool region, a larger swath of ocean encompassed GP and the spatial extent of the GPI. In the Eastern Pacific, the GPI reproduced the maximum density distribution off the western coast of Central America.

NORTH AMERICA: The lack of GPI is consistent to the little to no GP in North America.

SOUTH AMERICA: The GPI did not reproduce the spatial distribution of GP in southern South America. Rather, it produced a maximum distribution in northern South America.

ATLANTIC OCEAN: Similar to the Pacific Ocean distribution, the Atlantic Ocean distribution was mainly confined to a band located just north of the equator. Differing from the Pacific, however, the Atlantic band was approximately fifteen degrees wide. The Western Atlantic Ocean GP was overestimated by the GPI, as the GPI produced maximums in the vicinity of the Gulf Stream off the eastern United States. On the other hand, the spatial distribution in the eastern Atlantic was reproduced with a maximum centralized off the western coast of Africa.

WEST AFRICA: The GPI did replicate the spatial distribution off the southern coast of western Africa. However, it could not replicate the GP in the central part of western Africa.

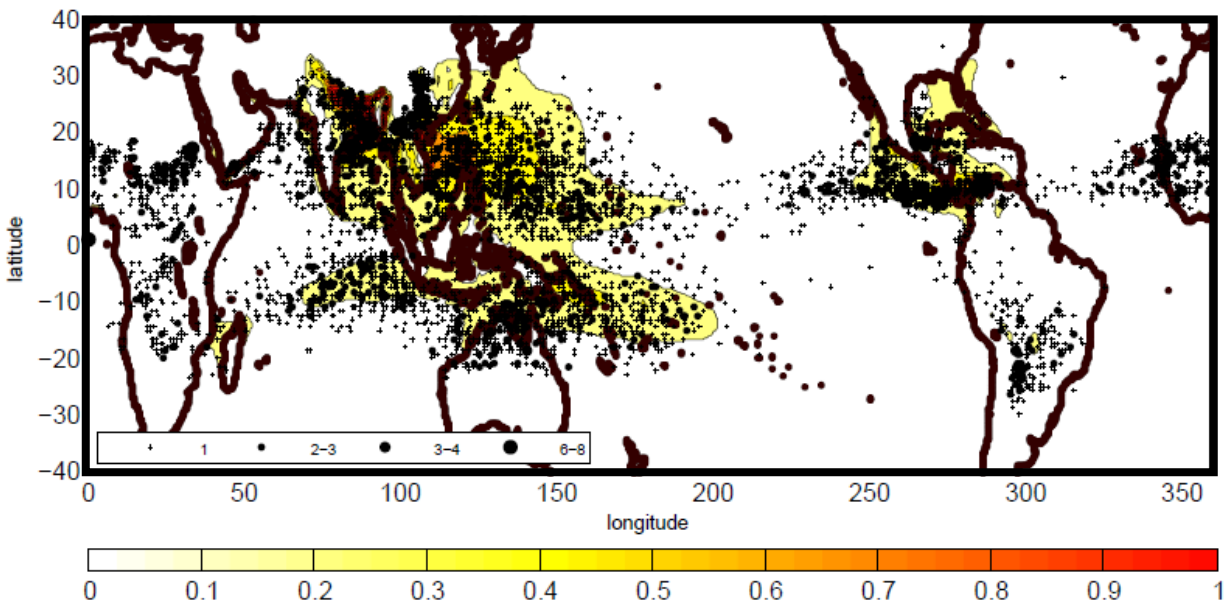


FIG. 12. The global spatial distribution of the MLPS GPI (shaded) based on Tab. 5 Line 1 coefficients for the 34 year period of January 1979 – December 2012. GP distribution is overlaid on top of the GPI distribution (black dots) with varying sizes corresponding to differing numbers of GP occurring at each point. The density values for the four different sized black dots are shown in the legend.

Overall, the index does a very good job in reproducing spatial distributions over India, the Indian Ocean, the Southern Ocean, the Pacific Ocean, North America, and the Atlantic Ocean; a good job in reproducing spatial distributions over Australia; and a generally poor job in reproducing the spatial distributions over Africa and South America.

C. Seasonal Climatology

Once again, the GPI is able to closely replicate the frequency of MLPS GP over the 34 year period (Fig. 13) climatologically. It depicts both the increase and decrease in MLPS GP during summers and winters respectively. As seen in Tab. 6, correlations between the actual number of GP and the estimated number of GP from the GPI have a 0.9851 correlation for the Northern Hemisphere and a 0.8873 correlation for the Southern Hemisphere.

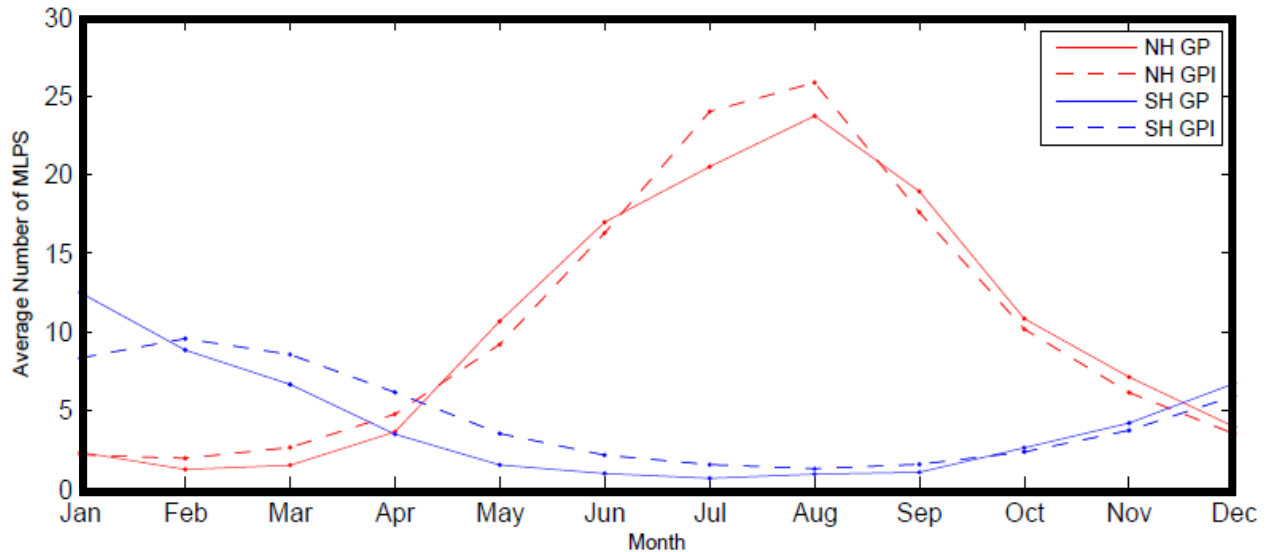


FIG. 13. The global climatology from the MLPS GPI is shown for the Northern (red dashed) and Southern (blue dashed) hemispheres. These values are compared to the actual TRACK determined climatology for the Northern (red solid) and Southern (blue solid) Hemispheres.

Additionally, the GPI projected that the Southern Hemisphere had an average of 197 GP less than the Northern Hemisphere, close to the actual 202 GP difference between the hemispheres. More statistics (Tab. 7) display the magnitude of difference between the actual GP count and the estimated GPI count.

| Difference | Northern Hemisphere | Southern Hemisphere |
|------------|---------------------|---------------------|
| Mean | 40 | 45 |
| Minimum | 4 | 9 |
| Maximum | 119 | 141 |

TAB. 7. Calculated from Fig. 13, the difference column in the above Table represents the mean, minimum, or maximum difference between the actual GP counts (solid lines, Fig. 13) and the Estimated GPI counts (dashed lines, Fig. 13).

VI. A Global TC GPI

To determine whether MLPS genesis is statistically similar to TC genesis, the TC GP dataset obtained from Dr. Kerry Emanuel was used in place of the TRACK dataset in conjunction with the same methodology used in this paper to generate a TC GPI. Then, the TC GP were run with meteorological variables that more closely matched the Tippet et al. (2011) variables. The beta coefficients (Tab. 8, Lines 2 and 3) were compared to both the global MLPS GPI results (Tab. 8, Line 1) and the global TC GPI results found in Tippet et al. (2011; Tab. 8, Line 4).

A. Executing the Poisson Regression: The Statistics

In comparing results from this research's two TC GPI runs to the global MLPS GPI and Tippet et al. (2011) results, a few important similarities and differences appear. First, near identical beta coefficients were obtained for humidity and shear coefficients in all four GPI results, with wind shear having a negative correlation to both TC and MLPS genesis. Also, near identical heat and vorticity coefficients were obtained for the MLPS GPI and the TC GPI which used this research's methodology. On the other hand, both the TC heat and vorticity coefficients differ from the Tippet et al. (2011) index. With identical meteorological variables as used previously in this research, the heat coefficient was three times smaller and the vorticity coefficient was four times smaller than those in Tippet et al. (2011). When using sea surface temperature and clipped absolute vorticity, the heat coefficient was about six fifths times smaller and the vorticity coefficient was about three times smaller than the Tippet et al. (2011) coefficients. This implies that the Tippet et al. (2011) index's heat and vorticity terms have a stronger correlation to TC formation.

| Trial | Constant | Heat | | Humidity | | Vorticity | | Shear | Statistics | | |
|------------------------|-----------|-------------|--------|------------|-------------|-----------|----------------|---------|------------|----------------|----------------|
| | | b_{ECAPE} | b_T | b_{TCWV} | $b_{SSM/I}$ | b_η | $b_{\eta,3.7}$ | | | | b_V |
| Scope | b | | | | | | | | σ | AIC | |
| Estimates | | | | | | | | | | | |
| Global | -9.7375 | 0.1273 | ----- | 0.0968 | ----- | 0.4053 | ----- | -0.0263 | 0.5241 | 43666 | |
| TC | -9.4635 | 0.1510 | ----- | 0.1011 | ----- | 0.3007 | ----- | -0.1421 | 0.4107 | 25232 | |
| TC_Tipp | -120.2764 | ----- | 0.3668 | 0.1060 | ----- | ----- | 0.4040 | -0.1433 | 0.4879 | 24541 | |
| Tippet* | -11.9600 | ----- | 0.4600 | ----- | 0.1200 | ----- | 1.1200 | -0.1300 | 1.7000 | 12070 | |
| Standard Errors | | | | | | | | | | NH Corr | SH Corr |
| Global | 0.0570 | 0.0020 | ----- | 0.0012 | ----- | 0.0038 | ----- | 0.0013 | 0.9851 | 0.8873 | |
| TC | 0.0680 | 0.0027 | ----- | 0.0014 | ----- | 0.0034 | ----- | 0.0019 | 0.9291 | 0.9537 | |
| TC_Tipp | 2.8970 | ----- | 0.0097 | 0.0020 | ----- | ----- | 0.0089 | 0.0024 | 0.9317 | 0.8657 | |
| Tippet ¹ | 0.5400 | ----- | 0.0370 | ----- | 0.0067 | ----- | 0.0440 | 0.0100 | ----- | ----- | |

TAB. 8. Global beta coefficient results from a Poisson regression. The first row corresponds to the results from executing a Poisson regression on the four variables excluding dry MLPS GP over the entire globe. The second row corresponds to the results from executing a Poisson regression on the four variables excluding dry TC GP over the entire globe. The third row corresponds to the results on executing a Poisson regression on four variables for TC GP as found in Tippet et al. (2011).

¹ Tippet et al. (2011) used slightly different variables under the same main headings. They used the heat variable relative sea surface temperature (SST) in C°, the humidity variable column integrated relative humidity from the Special Sensor Microwave Imager (SSM/I), the vorticity variable absolute vorticity modified to a threshold level of 3.5, and the same shear variable used by this paper.

B. Spatial Distribution

Global spatial coverage performed similarly to the results found in the global MLPS GPI. Since TCs form over oceans only, continents are not discussed unless the GPI shows a distribution over a continent. Additionally, each region is discussed for Fig. 14A and Fig. 14B. If differences between plots occur, it is noted by the respective letter before a comment. If spatial distributions are similar in both figures, then no letter will be listed.

INDIA AND ASIA:

A | Despite no presence of TC GP over land, the GPI shows a strong spatial distribution in these regions, similar to the MLPS GP and MLPS GPI distributions.

B | There is no GPI distribution over land, as SST is not defined over land.

INDIAN OCEAN – ARABIAN SEA AND BAY OF BENGAL: The GPI does not reproduce the GP located within the Arabian Sea. It does, however, show a more broad stretching distribution in the Bay of Bengal.

OCEAN (SOUTHERN): The distribution in the Southern Ocean is confined to a band.

A | The GPI index was not able to replicate this pattern well.

B | The GPI index was able to replicate this pattern well.

AUSTRALIA: The GPI was able to spatially show the GP off the northern coastline.

PACIFIC OCEAN: As with the global MLPS GPI, the global TC GPI was able to reproduce the approximate shape of the GP distribution throughout the Pacific Ocean.

A | The distribution was unable to replicate the thin band of storms which stretches from the Pacific Warm Pool eastward across the Pacific Ocean.

B | The distribution was able to replicate the thin band of storms stretching across the Pacific Ocean.

ATLANTIC OCEAN: Similar to the Pacific Ocean distribution, the Atlantic Ocean GP distribution was mainly confined to a band located just north of the equator. However, unlike the global MLPS GPI, neither of the global TC GPI was able to replicate this band.

A | There was no GPI distribution in the Atlantic Ocean or over the Gulf Stream. There was a GPI distribution in the Gulf of Mexico, however, which closely replicated the GP distribution

B | The GPI distribution was too far south of the equator. But, the GPI distribution in the Gulf of Mexico and over the Gulf Stream was able to be reproduced.

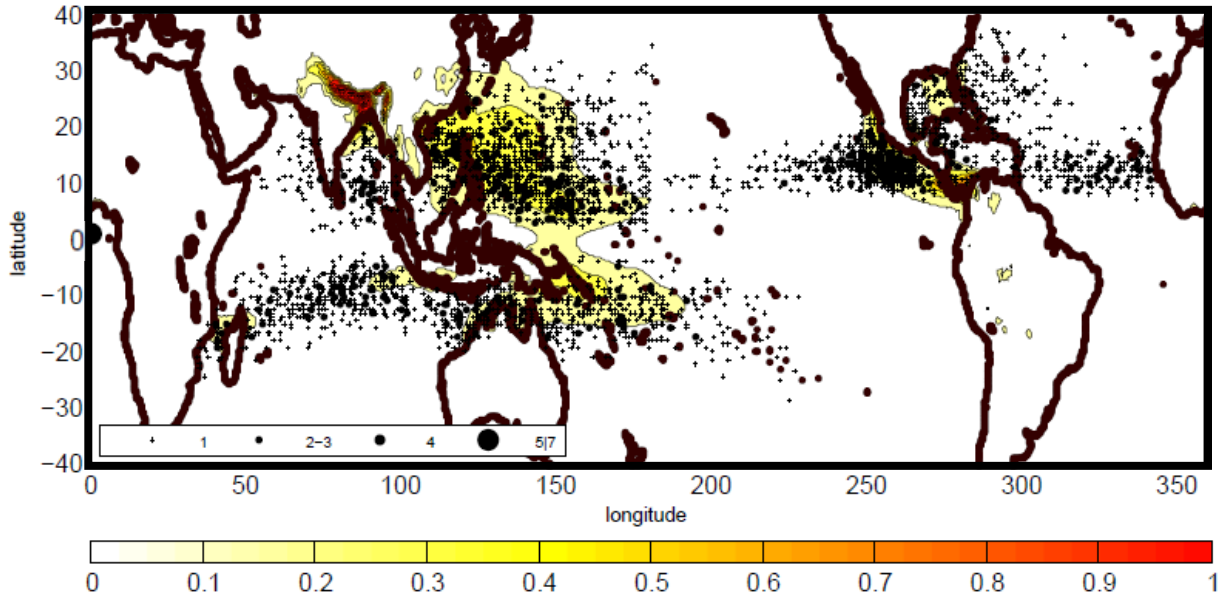


FIG. 14A. The global spatial distribution of the TC GPI (shaded) based on Tab. 8 Line 2 coefficients for the 34 year period of January 1979 – December 2012. GP distribution is overlaid on top of the GPI distribution (black dots) with varying sizes corresponding to differing numbers of GP occurring at each point. The density values for the four different sized black dots are shown in the legend.

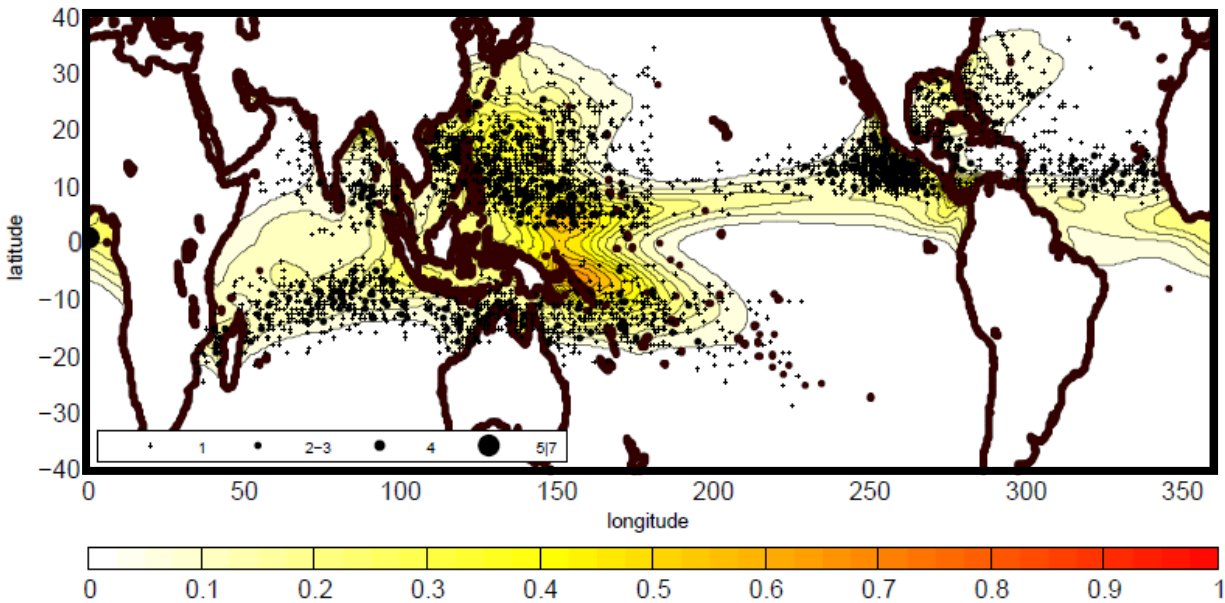


FIG. 14B. The global spatial distribution of the TC GPI (shaded) based on Tab. 8 Line 3 coefficients for the 34 year period of January 1979 – December 2012. GP distribution is overlaid on top of the GPI distribution (black dots) with varying sizes corresponding to differing numbers of GP occurring at each point. The density values for the four different sized black dots are shown in the legend.

Overall, the indexes do a very good job in reproducing spatial distributions over the Indian Ocean, the Southern Ocean, and the Pacific Ocean; a good job in reproducing spatial distributions over eastern Atlantic Ocean; and a generally poor job in reproducing the spatial distributions over the majority of the Atlantic Ocean. It also predicted a spatial distribution over India and Southern Asia, distributions which could not occur for hurricanes, as shown in Fig. 14A since all variables used are defined over both ocean and land. Fig. 14B used variables similar to Tippet et al. (2011) including their heat variable, SST, which is only defined over ocean. Therefore, no GPI distribution could occur over land. This implies that MLPS form under similar conditions as TC, and if TC had the capability of forming over land, they would exhibit a similar distribution as the MLPS in the Bay of Bengal. Perhaps then although the MLPS and TC genesis seem to be governed by the same variables the specific dynamics of genesis still differ as TC do not form over land.

C. Seasonal Climatology

The global TC GPI is able to closely replicate the frequency of TC GP over the 34 year period (Figs. 15) climatologically. It depicts both increase and decrease in TC GP during summers and winters respectively. As seen in Tab. 8, Line 2, correlations between the actual number of GP and the estimated number of GP from the GPI has a 0.9291 correlation for the Northern Hemisphere and a 0.9537 correlation for the Southern Hemisphere.

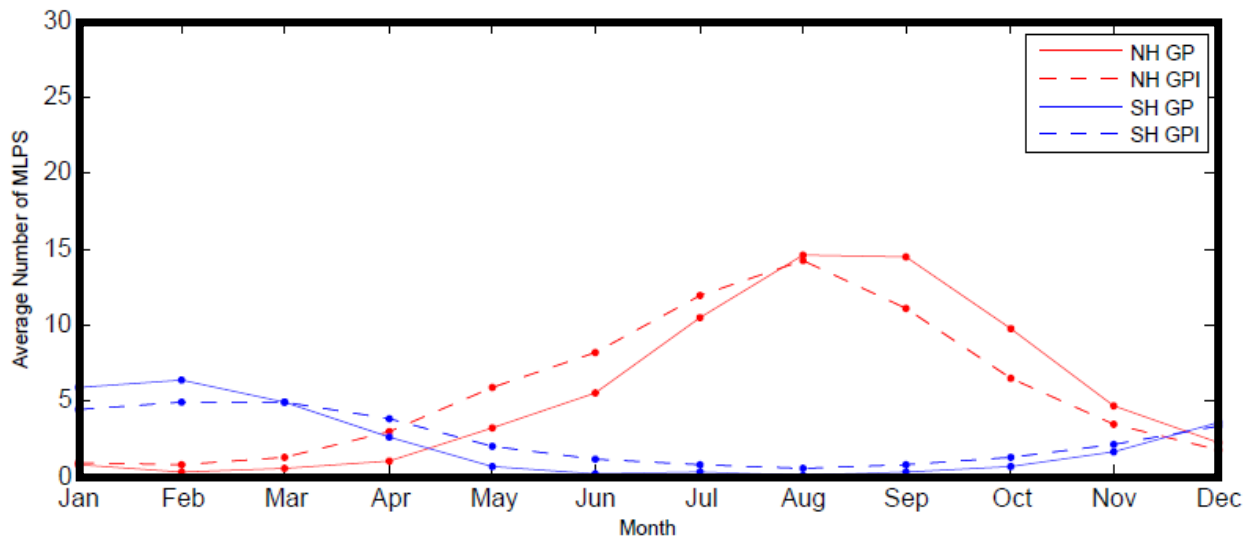


FIG. 15A. The global climatology from the TC GPI (Tab. 8, Line 2) is shown for the Northern (red dashed) and Southern (blue dashed) hemispheres. These values are compared to the actual TC climatology for the Northern (red solid) and Southern (blue solid) hemispheres as obtained from Emanuel.

Additionally, the Southern Hemisphere was projected to have an average of 110 GP less than the Northern Hemisphere, close to the actual 115 GP difference between the hemispheres. More

statistics (Tab. 9) display the magnitude of difference between the actual GP count and the estimated GPI count.

| Difference | Northern Hemisphere | Southern Hemisphere |
|------------|---------------------|---------------------|
| Mean | 53 | 26 |
| Minimum | 8 | 1 |
| Maximum | 117 | 52 |

TAB. 9. Calculated from Fig. 15, the difference column in the above Table represents the mean, minimum, or maximum difference between the actual GP counts (solid lines, Fig. 15) and the Estimated GPI counts (dashed lines, Fig. 15).

As seen in Tab. 8, Line 3, correlations between the actual number of GP and the estimated number of GP from the GPI has a 0.9371 correlation for the Northern Hemisphere and a 0.8657 correlation for the Southern Hemisphere.

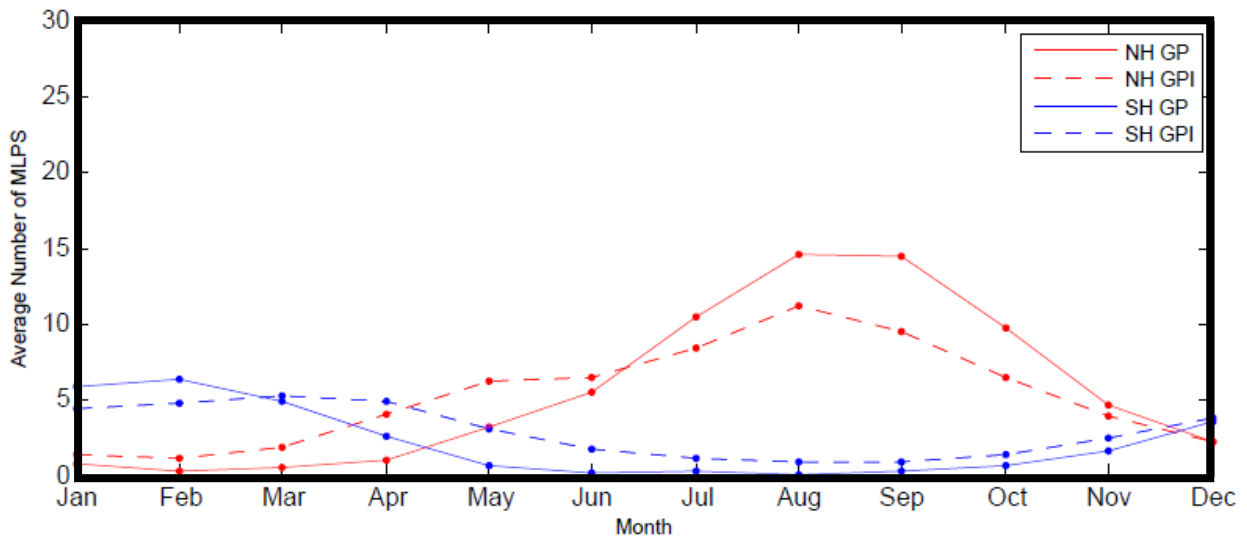


FIG. 15B. The global climatology from the TC GPI (Tab. 8, Line 3) is shown for the Northern (red dashed) and Southern (blue dashed) hemispheres. These values are compared to the actual TC climatology for the Northern (red solid) and Southern (blue solid) hemispheres as obtained from Emanuel.

Additionally, the Southern Hemisphere was projected to have an average of 80 GP less than the Northern Hemisphere, close to the actual 115 GP difference between the hemispheres. More statistics (Tab. 10) display the magnitude of difference between the actual GP count and the estimated GPI count.

| Difference | Northern Hemisphere | Southern Hemisphere |
|------------|---------------------|---------------------|
| Mean | 68 | 39 |
| Minimum | 1 | 9 |
| Maximum | 166 | 82 |

TAB. 10. Calculated from Fig. 15, the difference column in the above Table represents the mean, minimum, or maximum difference between the actual GP counts (solid lines, Fig. 15) and the Estimated GPI counts (dashed lines, Fig. 15).

VII. Future Work

This research was performed with monthly mean data averaged over a 34 year period to create a GPI for MLPS. It was queried whether the use of daily data from the ERA-INT reanalysis datasets with a similar methodology would yield comparable results as those found when using monthly mean data. The one change to the methodology was the use of a logistic regression instead of a Poisson regression, since daily data cause values of GP to range between 0 and 1 for any given day. Yet, the resulting spatial distribution of MLPS in the GPI always clustered around the sharp topographical gradient of the Tibetan Plateau and Himalayas. As a result, a Poisson regression was used in conjunction with the daily data in an attempt to exactly match methodology used with monthly mean data, despite MLPS GP varying between 0 and 1 for any given day. Still, spatial bias towards topography was present. Further investigation into why these biases are occurring is needed and will be attempted in the future.

VIII. Conclusion

This research was successfully able to create a MLPS GPI for those MLPS GP that occurred where the TCWV was greater than 35 kgm^{-2} on three spatial scales: the AA region, the Indian Continent region, and the entire globe. Each of the three GPI was able to spatially and climatologically replicate the actual distribution and trends of GP provided by Hurley over the January 1979 through December 2012 period. Since the coefficients obtained for the three GPI created were very similar, this implies that the genesis of moist MLPS worldwide is governed by the same properties, those being wind shear inhibiting and absolute vorticity having the highest contribution to MLPS genesis. Furthermore, by creating a TC GPI and comparing those results to that of the global MLPS GPI and the Tippet et al. (2011) results, it was shown that if TC could form over land, then the TC spatial distribution would mimic that of the MLPS distribution over the Bay of Bengal. Additionally, the similarity between the TC GPI and the MLPS GPI coefficients implies that the underlying causes of genesis for TC and MLPS are similar.

IX. Acknowledgments

Thank you to my friends, suitemates, and parents. Although you rarely understood what I was talking about, as mostly I mumbled in MATLAB code, you always listened and cheered me on.

I want to thank the Yale University's Geology & Geophysics' Karen L. Von Damm '77 Undergraduate Research Fellowship for funding my research during the summer of 2012 and the support by Office of Naval Research Young Investigator Program award N00014-11-1-0617 for funding my research from the summer of 2011 through the summer of 2012.

Thank you to John Hurley, formerly of Yale University and currently at the University of Albany, for compiling the global, comprehensive, objective analysis of MLPS GP and tracks. Additionally, the Boos Research Group, comprised of Professor Boos, a few graduate students, and a few post-doctoral students, was instrumental in providing suggestions, answering questions, and listening inquisitively over the past three years when I discussed and presented my research updates and paper reviews.

For academic assistance, guidance, and advice, I want to thank my academic advisor, Professor Ronald B. Smith, the Director of Undergraduate Studies, Professor David Evans, and all those professors that I have had as instructors and mentors over the past four years. Additionally, I want to thank Trude Storelvmo for being my second reader.

Last, but certainly not least, thank you to my fantastic and supportive mentor, Professor William Boos, for allowing me to begin this research as a freshman and always encouraging, guiding, and teaching me how to think inquisitively, grow scientifically, and learn adaptively. Without his outstanding care, constant patience, and attentive personality I would not have been able to complete this research.

X. References

- Ajayamohan, R.S., Merryfield, W.J., & Kharin, V.V. (2010). Increasing Trend Of Synoptic Activity And Its Relationship With Extreme Rain Events Over Central India. *Journal of Climate*. 23 (4), 1004–1013.
- Akaike, H. (1973). Information Theory And An Extension Of The Maximum Likelihood Principle. *Second International Symposium on Information Theory*. 267-281.
- Allen, C., Kramer, D., Smith, R., & Stults, A. (2002). Vertical Resolution and Coordinates. *Texas A&M University Meteorology Department*. Retrieved December 26, 2013 from <http://atmo.tamu.edu/class/metr452/models/2001/vertres.html>.
- BBC News. (2013). India Floods: Death Toll in Uttarakhand ‘Passes 500’. *BBC News India*. Retrieved December 26, 2013 from <http://www.bbc.co.uk/news/world-asia-india-23007121>.
- Berrisford, P., Dee, D., Fielding, K., Fuentes, M., Kallberg, P., Kobayashi, S. & Uppala, S. (August, 2009). The ERA-Interim Archive. *ERA Report Series*. (1), 1-16.
- Camargo, S.J., Emanuel, K.A., & Sobel, A.H. (2007). Use of a Genesis Potential Index to Diagnose ENSO Effects on Tropical Cyclone Genesis. *Journal of Climate*. 20(19).
- Chen, T. & Weng, S. (1999). Interannual And Intraseasonal Variations In Monsoon Depressions And Their Westward-Propagating Predecessors. *Monthly Weather Review*. 127 (6), 1005–1020.
- Clift, P.D., & Plumb, R.A. (2008). The Asian Monsoon: Causes, History And Effects. *Cambridge University Press*. Vol. (288).
- ECMWF. (2013). 60 Model Level Definitions. *ECMWF*. Retrieved December 26, 2013 from http://www.ecmwf.int/products/data/technical/model_levels/model_def_60.html.
- Elsner, J.B., & Jagger, T.H. (2013). Hurricane Climatology: A Modern Statistical Guide Using R. *Oxford University Press*.
- Emanuel, K.A. (13 November, 2013). Global Tropical Cyclone Data in NETCDF Format. *MIT*. Retrieved December 23, 2013 from ftp://texmex.mit.edu/pub/emanuel/HURR/tracks_netcdf/.
- Emanuel, K.A., & Nolan, D.S. (May, 2004). Tropical Cyclone Activity And The Global Climate System. *26th Conf. on Hurricanes and Tropical Meteorology*. Miami, FL.
- Gray, W. (1979). Hurricanes: Their Formation, Structure And Likely Role In The Tropical Circulation. *Meteorology over the Tropical Oceans*. 77, 155-218.
- Hodges, K. (1995). Feature Tracking On The Unit Sphere. *Monthly Weather Review*. 123, 3458– 3465.

- Jenamani, R.K., Bhan, S.C., & Kalsi, S.R. (2006). Observational/Forecasting Aspects Of The Meteorological Event That Caused A Record Highest Rainfall In Mumbai. *Current Science-Bangalore*, 90(10), 1344.
- Lim, E.P., & Simmonds, I. (2007). Southern Hemisphere Winter Extratropical Cyclone Characteristics And Vertical Organization Observed With The ERA-40 Data In 1979–2001. *Journal of Climate*. 20, 2675–2690.
- Masters, J. (2013). Over 500 Killed in India's Monsoon Floods. *Dr. Jeff Masters' WunderBlog*. Retrieved December 26, 2013 from <http://www.wunderground.com/blog/JeffMasters/over-500-killed-in-indias-monsoon-floods>.
- Mooley, D.A. & Shukla, J. (1987). Characteristics Of The Westward-Moving Summer Monsoon Low Pressure Systems Over The Indian Region And Their Relationship With The Monsoon Rainfall. *Center for Ocean-Land-Atmosphere Interactions*.
- National Geophysical Data Center (NGDC). (1988). ETOPO-5 Bathymetry/Topography Data. Data Announcement 88–MGG–02. *National Oceanic and Atmospheric Administration*. U.S. Department of Commerce, Boulder, CO.
- Ramage, C.S. (1971). Monsoon Meteorology. *New York: Academic Press*. Vol. (296).
- Saha, K., Sanders, F., & Shukla, J. (1981). Westward Propagating Predecessors Of Monsoon Depressions. *Monthly Weather Review*. 109 (2), 330–343.
- Sikka, D.R. (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, Center for the Application of Research on the Environment*. Technical Report, 217.
- Tippett, M., Camargo, S., & Sobel, A. (2011). A Poisson Regression Index For Tropical Cyclone Genesis And The Role Of Large-Scale Vorticity In Genesis. *Journal of Climate*. 24, 2335-2357.
- Uppala, S.M., Kållberg, P.W., Simmons, A.J., Andrae, U., Bechtold, V., Fiorino, M., & Woollen, J. (2005). The ERA-40 Re-Analysis. *Quarterly Journal of the Royal Meteorological Society*, 131(612), 2961-3012.
- Yoon, J. & Chen, T. (2005). Water Vapor Budget of The Indian Monsoon Depression. *Tellus A*. 57, 770–782.