Regional Typhoon Activity¹ as Revealed by Track Patterns and Climate Change

Pao-Shin Chu, Xin Zhao[#], and Joo-Hong Kim^{*}

Department of Meteorology, School of Ocean and Earth Science and Technology, University of Hawaii, Honolulu, Hawaii (<u>chu@hawaii.edu</u>)

#Sanjole Inc., Honolulu, Hawaii (xzhao@sanjole.com)

*Department of Atmospheric Sciences, National Taiwan University, Taipei, Taiwan (jhkim004@gmail.com)

Abstract

With an expectation-maximization (EM) algorithm solving model parameters, a clustering method built on a mixture Gaussian model is applied to the Joint Typhoon Warning Center (JTWC) best-track records to objectively classify historical tropical cyclone (TC) tracks (1945-2007) over the western North Pacific into eight types. The first three types are labeled as straight movers (A, B, and C), followed by four recurved types (D, E, F, and G), and one mixed straight-recurved type (H). For each type, a log-linear regression model is then applied to detect abrupt shifts in the time series of TC attributes including frequency, lifespan, intensity, and accumulated cyclone energy (ACE). Results indicate that the major climate shift in 1976/1977 may have affected storm's counts for two track patterns (types F and H). All eight types exhibit at least one abrupt shift in their duration since 1945, with three types (A, C, and H) showing a common shift in 1972. For a majority of the eight types, the storms' mean lifetime became longer after the shift. TC intensity shows a prevalence of abrupt shifts in the 1970s. For type D, its intensity has undergone several changes (1972, 1988, and 1998) with stronger intensity since the last shift. Because of its proximity to the East Asian landmasses and its abundance in numbers, the increasing intensity of type D since 1998 is a concern for Taiwan, the east China coast, the Philippines, Japan, and Korea. For ACE, the signal is mixed. To draw more definitive conclusions, a consistency check with another best-track record is called for.

Keywords: Typhoon track clusters, regional typhoon activity, climate change

Corresponding author: Pao-Shin Chu, Department of Meteorology, University of Hawaii, 2525 Correa Road, Honolulu, Hawaii, 96822.

1. Introduction

There is a growing interest in knowing the impact of climate change on tropical cyclone (TC) activity. This interest stems from the fact that tropical storms are among the most destructive of natural disasters. As the world becomes warmer, it is conceivable that the intensity, frequency, tracks, and location of occurrences of tropical storms will be altered from the present-day climate. Using the Joint Typhoon Warning Center (JTWC) best-track records or satellite data, Emanuel (2005), Webster et al. (2005), and more recently Elsner et al. (2008) demonstrated that historical storm intensity has increased dramatically in the western North Pacific (WNP) as well as the North Atlantic. In particular, Webster et al. (2005) noted a large increase in the frequency for the strongest TC categories (4 and 5) in the WNP over the last 30 years: 25 such storms were found in the first 15 years (1975-1989) but the 41 strongest storms occurred during the second 15 years (1990-2004). However, this large trend has been questioned due to large interdecadal variations in the number of intense TCs (Chan, 2006) or possible measurement errors in the dataset (Knaff and Zehr, 2007). In addition, based on the best-track records over a more recent 20-yr (1986-2006), Klotzbach (2006) found small or no trends using alternative analysis methods. Besides the variation in intensity, TC induced rainfall in the WNP also appears to have undergone long-term variations with decadal signatures (Kim et al., 2006; Lau et al., 2008).

Given this dramatic increase over the entire basin as noted by Webster et al. (2005), it is of interest to examine whether there is any change in TC activity at the regional level within the WNP as this is a huge basin and TCs have preferred tracks. Some of them form over the Philippine Sea and track westward to the South China Sea, southern China, and/or Vietnam. Others form over the tropical Pacific and move northwestward toward Taiwan and/or China's coast or recurve toward Japan and Korea (Tu et al., 2009). This study examines the temporal changes in regional typhoon activity over the WNP during the past 60 years using the 6-h best track data from the JTWC. A mixture Gaussian model is built for WNP typhoon tracks based on which historical typhoon tracks are categorized using a clustering algorithm. Eight track patterns are individually examined for temporal changes over the last 60 years.

2

2. Methods

2.1 Clustering methodology

Our TC track clustering method is based on the mixture Gaussian model. A key feature of the mixture Gaussian model is its ability to model multimodal densities while adopting a small set of basic component densities. Finite mixture models have been widely used for clustering data in a variety of areas such as large-scale atmospheric circulation (Camargo et al., 2007). In this study, we assume that there are a few distinct path track types characterizing TC tracks in the WNP. For each TC track path, we model it as a second-order polynomial function of the lifetime of this TC. The basic assumption we impose here is that for each specific track type, the set of coefficient of this polynomial function is jointly Gaussian distributed. Each TC track type has its unique distribution parameter. Therefore, the space spanned by the parameters of this track type model is a linear combination of a set of Gaussian distribution, or a mixture Gaussian distribution model.

We assume there are n observed track records at 6-h intervals for a given TC. For each record, there will be three features reported—latitude, longitude, and the time. We will denote the path record of a TC by

$$\mathbf{z} = [\mathbf{z}_{lat}, \mathbf{z}_{long}] = \begin{bmatrix} z_{1,lat} & z_{1,long} \\ \dots & \dots \\ z_{n,lat} & z_{n,long} \end{bmatrix}$$
(1a)

where $z_{i,lat}$ and $z_{i,long}$ for i = 1,..,n represent the *i*-th latitude and longitude record, respectively. We then denote the relative observed time vector for the second order polynomial function by

$$\mathbf{T} = \begin{bmatrix} 1 & t_1 & t_1^2 \\ \dots & \dots & \dots \\ 1 & t_n & t_n^2 \end{bmatrix}$$
(1b)

where t_i for i = 1,..,n represents the time for the *i*-th records of this TC relative to the first record. We further assume that there are *K* distinct TC track types in

the WNP, where K is assumed to be a constant throughout this study. With definitions (1a) and (1b), provided that this TC is categorized as type k, $1 \le k \le K$, the linkage between the passage and relative time is governed by the following formula

$$\mathbf{z} = \mathbf{T}\boldsymbol{\beta}^{\mathbf{k}} + \boldsymbol{\varepsilon} \text{, where } \boldsymbol{\beta}^{\mathbf{k}} = \begin{bmatrix} \boldsymbol{\beta}_{0,lat}^{k} \boldsymbol{\beta}_{0,long}^{k} \\ \boldsymbol{\beta}_{1,lat}^{k} \boldsymbol{\beta}_{1,long}^{k} \\ \boldsymbol{\beta}_{2,lat}^{k} \boldsymbol{\beta}_{2,long}^{k} \end{bmatrix} \text{ and } \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{\Sigma}^{\mathbf{k}}) \text{.}$$
(1c)

In (1c), the parameter set $\boldsymbol{\beta}^{k}$ is distinct from other TC types. With the model given in (1) for type k, intuitively we can see that the zero-order coefficient dual provides the mean genesis location of this type; the first-order term features the characteristic direction of this path type; the second-order type will determine the recurving shape of the typical path of this type; and the covariance matrix (Σ) in (1c) determines the spread of a particular type. The noise term in (1), $\boldsymbol{\varepsilon}_{i}$, is multivariate Gaussian with zero mean and a 2 by 2 covariance matrix, $\boldsymbol{\Sigma}_{k}$. The covariance matrix $\boldsymbol{\Sigma}_{k}$ contains diagonal elements $\sigma_{0,k}^{2}$ and $\sigma_{1,k}^{2}$, which are the noise variances for each latitude and longitude observation, respectively, in the cluster k. For simplicity the cross covariance of $\boldsymbol{\Sigma}_{k}$ is set to 0.

The conditional density for the *i*th cyclone, conditioned on membership in the cluster k, is thus defined as

$$P(\mathbf{z}_i | \mathbf{T}_i, \mathbf{\theta}_k) = (2\pi)^{-n_i} |\mathbf{\Sigma}_k|^{-n_i/2} \exp\left\{-\operatorname{tr}[(\mathbf{z}_i - \mathbf{T}_i \mathbf{\beta}_k) \mathbf{\Sigma}_k^{-1} (\mathbf{z}_i - \mathbf{T}_i \mathbf{\beta}_k)']/2\right\} (2)$$

In (2), we adopt the notation $\mathbf{\theta}_k = {\{\mathbf{\beta}_k, \mathbf{\Sigma}_k\}}$. This distribution leads to the marginal regression mixture model with K = 8 clusters,

$$P(\mathbf{z}_i | \mathbf{T}_i) = \sum_{k=1}^{K} \alpha_k P(\mathbf{z}_i | \mathbf{T}_i, \mathbf{\theta}_k).$$
(3)

In (3), α_k is the probability of cluster k, and $\sum_{k=1}^{K} \alpha_k = 1$. If we let $\mathbf{Z}' = [\mathbf{z}'_1, \mathbf{z}'_2, ..., \mathbf{z}'_N]$ be the complete set of all observed TC trajectories and $\mathbf{T}' = [\mathbf{T}'_1, \mathbf{T}'_2, ..., \mathbf{T}'_N]$ be the associated measurement times, then the full probability density of \mathbf{Z} given \mathbf{T} , which is also known as the conditional likelihood, is

$$P(\mathbf{Z} | \mathbf{T}) = \prod_{i=1}^{N} \sum_{k=1}^{K} \alpha_{k} P(\mathbf{z}_{i} | \mathbf{T}_{i}, \boldsymbol{\theta}_{k}).$$
(4)

Given a set of a total N TC records, $\{\mathbf{z}_i, \mathbf{T}_i \mid i = 1,...,N\}$, to derive the maximum likelihood estimation of all model parameters and class type, we resort to the Expectation-Maximization (EM) algorithm. In the E-step, we calculate the membership probability of each type for each TC as follows

$$w_{i,k} = \frac{\alpha_k f(\mathbf{z}_i \mid \mathbf{T}_i \boldsymbol{\beta}^k, \boldsymbol{\Sigma}^k)}{\sum_{j=1}^{K} \alpha_j f(\mathbf{z}_i \mid \mathbf{T}_i \boldsymbol{\beta}^j, \boldsymbol{\Sigma}^j)}, \ i = 1, ..., N$$
(5)

where $\alpha_k = f(\mathbf{k})$ is the prior probability of type *k*. Apparently, the membership probability of a TC in (5) is virtually the Bayes' posterior probability of each track type given all model parameter sets. If $\mathbf{w}_{i,k} = w_{i,k} \mathbf{1}_{n_i}$, where n_i denotes the record length of the *i*-th TC and $\mathbf{1}_{n_i}$ represents the n_i vector of ones, we define a new diagonal matrix $\mathbf{W}^{\mathbf{k}} = \text{diag}([\mathbf{w}_{1,k}, \mathbf{w}_{2,k}, ..., \mathbf{w}_{N,k}])$ for each track type *k*. In the M-step, we then calculate the following estimation for the model parameter set of each type.

$$\hat{\boldsymbol{\beta}}^{\mathbf{k}} = (\mathbf{T}'\mathbf{W}^{\mathbf{k}}\mathbf{T})^{-1}\mathbf{T}'\mathbf{W}^{\mathbf{k}}\mathbf{Z}$$
(6a)

$$\hat{\alpha}_k = \frac{1}{S} \sum_{i=1}^N n_i w_{i,k} \tag{6b}$$

$$\Sigma^{k} = \frac{1}{S} \frac{(\mathbf{Z} - \mathbf{T}\boldsymbol{\beta}^{k})' \mathbf{W}^{k} (\mathbf{Z} - \mathbf{T}\boldsymbol{\beta}^{k})}{\hat{\alpha}_{k}}$$
(6c)

In (6), $\mathbf{Z}' = [\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_N]'$, $\mathbf{T}' = [\mathbf{T}_1, \mathbf{T}_2, ..., \mathbf{T}_N]'$, where \mathbf{z} , \mathbf{T} are defined in (1) and the subscript represents the index of a TC; and $S = \sum_{i=1}^N n_i$ is the total

number of observations.

If the number of clusters, K, is provided, given an initial setting of the model parameters and with multiple iterations of (5) and (6), the proposed EM algorithm will converge to a fixed set of parameter estimation. These convergence values are not necessarily the global best estimation and are determined by the initial starting values. Therefore, we apply multiple different initial values and choose the set of estimation with the maximum likelihood. It is also noteworthy that the cluster problem presented hereby is essentially a missing value problem in which the indicators of class identification are not available.

2.2 Change-point analysis

In this study we analyze the time series of annual TC counts, average life spans, average ACE, and average intensity to determine any abrupt shifts. All these series are positive numbers and can be approximated by a gamma distribution. On the other hand, it is well known that a logarithm of a gamma distribution can be well approximated by a Gaussian distribution. In view of this fact, we adopt a (generalized) log linear regression model to determine abrupt shifts (or change-points) of the aforementioned time series (Elsner et al., 2000; Chu, 2002). More elaborately, change-point analysis can be cast in the Bayesian framework to explicitly provide the posterior probabilities of the change points as a function of time (e.g., Chu and Zhao, 2004; Zhao and Chu, 2006). Assuming that a time series is denoted by X_i , i = 1, ..., n, we calculate the latent set $Y_i = \log(1 + X_i)$. We then define the following function

$$S_i(l) = \begin{cases} 0 & i < l \\ 1 & i \ge l \end{cases}, \text{ for } 1 < l < n .$$
(7a)

We thereby adopt the following linear regression model

$$Y_{i} = c_{0}(l) + c_{1}(l)S_{i}(l) + \varepsilon_{i}(l).$$
(7b)

For each l, we use a Least Square Error (LSE) algorithm to compute the intercept $c_0(l)$ and slope $c_1(l)$. Subsequently we define t-ratio as follows.

$$t(l) = \hat{c}_1(l) / se(\hat{c}_1(l))$$
(7c)

In (7c), $\hat{c}_1(l)$ is the estimated slope term and $se(\hat{c}_1(l))$ the estimated standard error of this term. If $t(m) = \max\{t(l), l = 2, ..., n-1\}$ is larger than a critical value, we set index m as a change-point. The critical value is obtained from a t-distribution. For example, this critical value is set as 2.65 for a 99% confidence level. After finding the first change-point m, the process (7) is repeated with a new response series $Y_i^* = Y_i - c_1(m)S_i(m)$, which yields the second change-point if there is any. We repeat this iterative process until no more change-points are found.

3. Data

The TC data over the WNP come from the U.S. JTWC in Honolulu. The data cover the period 1945 to 2007. The data sets contain measurements of TC center location in latitude, longitude, one-minute sustained maximum wind speed, and central pressure at 6-h intervals for all TCs in the WNP. Here TC refers to tropical storms and typhoons. Tropical storms are defined as maximum sustained surface wind speeds between 17.5 and 33 m s⁻¹, and typhoons are defined as wind speeds at least 33 m s⁻¹. The ACE of a year is defined by summing the squares of the estimated maximum sustained velocity of every tropical storm or typhoon at six-hour intervals during the lifetime of a storm through the course of a year. This summed value is then divided by the number of TCs in that year to get the ACE per storm for each year. Because the original numbers are large, they are divided by 10,000 to make them more manageable. ACE is related to storm kinetic energy.

4. Results

The methods presented in Section 2 are applied to the data outlined in Section 3. Using an empirical argument, Camargo et al. (2007) suggested that the optimum number of track patterns in the WNP varies from six to eight. Based on our own simulation results with longer data sets, it seems that eight clusters offer the best explanation for the data.

Figure 1 shows eight major track patterns over the WNP and the South China Sea, with three straight movers (types A, B and C), four recurved ones (types D, E, F, and G), and one mixed pattern of both straight moving and recurved (type H). The type A and B clusters are similar in nature in that they both move more or less straight across Philippines to the South China Sea and/or Hong Kong, Hainan, and Vietnam. The major difference is that type B storms tend to form farther eastward and southward than type A storms. As a result, the mean track for type B storms is longer than that of Type A. Type C cyclones form in the South China Sea and are landlocked by the Indochina peninsula and southern China's coast, with very short path. Similar to types A and B, type D and E systems form in the Philippine Sea but they follow a northward path and many of them made landfall on Taiwan, the eastern China coast, Japan, and Korea. Type F storms tend to form in lowlatitudes and away from Asia. Type G storms also form far away from the Asian continent but at higher latitudes (~15°N). They move northwestward and then northward to the east of Japan over the open ocean. Storms associated with type H are generally formed near the equator and to the east of 165°E, and have a long trajectory over the water. In terms of the frequency of occurrence, Type D has the highest number (316 out of a total of 1621 cases). This is followed by type C (270), E (231), and G (225). Type H has the least number of occurrences (84) among eight types.

8



Fig.1. Eight TC track types identified by the mixture Gaussian model. The number in each panel indicates the number of cases in each type. Black circles denote the mean track for each type.

For each pattern, temporal variations of typhoon related attributes such as frequency, intensity, lifetime, and energy (ACE) are examined. A change-point analysis is applied to detect abrupt shifts in the time series of such attributes. For TC frequency (Fig. 2), five out of eight types (B, D, F, G, and H) show at least one step-like change since 1945. For types F and H, the abrupt shift occurs in 1976, at a time when major climate regime underwent a phase shift (e.g., Trenberth, 1990). After 1976, TCs became more active for both types F and H. For type D storms, which are the leading pattern among all eight groups, typhoon activity has very likely undergone a decadal variation with two abrupt shifts occurring around 1987 and 1998 with three epochs characterized by the active 1945-1986 epoch, the inactive 1987-1997 epoch, and the active 1998-2007 epoch. The reincreasing activity for type D since 1998 is a concern because storms associated with this type are formed near the eastern Asian landmass and their preferred tracks are likely to cause damage to Taiwan, east China coast, Japan, and Korea.



Fig. 2. Time series of annual tropical cyclone counts in the western North Pacific for eight track types. Years in each panel indicate possible change-points or abrupt shifts. Dotted lines are time-mean for each sub-period. The y-axis is the tropical cyclone number.

In Fig. 3, six of the track types (A, B, C, D, F, and H) exhibit one abrupt shift over the last 60 years with the exception of type B, which experienced two shifts. Interestingly, after the shift all five types (A, C, D, F, and H) show an increasing level of storm days, indicative of longer mean storm days after the identified shifts. It is possible that steering flows over the WNP become weaker, making storms traverse more slowly and increasing the lifetime of the storms. This possibility will be examined in future research. Also noteworthy in Fig. 3 is the phase shift in 1972, which occurs in types A, C and H.

The change in TC intensity is also quite interesting (Fig. 4), although this quantity from the JTWC is subject to larger uncertainty. Over the last 60 years, seven (A, B, C, D, E, F, and G) out of the eight types exhibit abrupt shifts in the 1970s. For types B, C, D and G, the shift occurs in 1971-73 and for types E and F

in 1974. For type D, which is not only the leading cluster but also bears threats to Taiwan, east China, Japan, and Korea, a step-like change also occurred recently (1998) with higher intensities, signifying stronger TC intensities in the last few years. This increase in intensity, together with the concurrent increase in frequency (Fig. 2), deserves further investigation.



Fig. 3. Same as Fig. 2 but for tropical cyclone lifespan. The y-axis is number of days

For ACE (Fig. 5), types C, E, and F do not show any apparent change in the past 60 years. The lifetime, intensity, and number of tropical cyclones all contribute to the magnitude of ACE. During El Niño and La Niña years, Camargo and Sobel (2005) found that the lifetime effect appears to be more important than the other two factors to ACE variations in the WNP. This may be reflected by a shift in type H in the 1972 El Niño (Figs. 3 and 5), but not on other types (e.g., A and C). Emanuel (2005) noted a substantial increase in the power dissipation index (PDI) for the WNP since 1970s. The PDI is similar to ACE, except that maximum wind speed is cubed over the storm's lifetime. Emanuel (2005) attributed this increase to the longer lifetime of the storm and greater storm intensities. In our case, only five types show an increase in their lifetimes (Types A, C, D, F, and H) and there is a lack of general intensity increase. These results are different from Emanuel which is based on the basin-wide values.



Fig.4. Same as Fig. 2 but for tropical cyclone intensity. The y-axis is knots.

5. A suggestion for future research

It should be noted that the results presented in this study are based on the JTWC best-track dataset. The veracity of JTWC data during the period prior to satellite observation (before 1970) is called into question (Knaff and Zehr, 2007; Lowry et al., 2009). While the track patterns, TC counts, and lifetime are probably reliable, TC intensity estimates and ACE may be subject to uncertainty. Therefore, extreme caution must be exercised in the intensity and ACE analysis and interpretations. In the future, we plan to apply the same methodology outlined in foregoing section to other best-track data to check for consistency.



Fig.5. Same as Fig. 2 but for the average ACE per storm (10^4 kt^2) .

Acknowledgments

This study was partially funded by the Pacific Disaster Center on Maui, Hawaii, the Central Weather Bureau in Taiwan, and the Korean Ocean Research Development Institute. P.-S. Chu also benefited from discussions with Mr. M. Lowry.

References

- Camargo, S.J., and A.H. Sobel, 2005: Western North Pacific tropical cyclone intensities and ENSO. J. Climate, 18, 2996-3006.
- Camargo, S.J., A.W. Robertson, S.J. Gaffney, P. Smyth, and M. Ghil, 2007: Cluster analysis of typhoon tracks. Part I: general properties. J. Climate, 20, 3635-3653.
- Chan, J.C.L., 2006: Comment on "Changes in Tropical Cyclone Number, Duration, and Intensity in a Warming Environment", *Science*, **311**, 1713, doi:10.1126/science.1121522.
- Chu, P.-S., 2002: Large-scale circulation features associated with decadal variations of tropical cyclone activity over the central North Pacific. J. Climate, 15, 2678-2689.
- Chu, P.-S., and X. Zhao, 2004: Bayesian change-point analysis of tropical cyclone activity: The central North Pacific case. J. Climate, **17**, 4893-4901.
- Emanuel, K., 2005: Increasing destructiveness of tropical cyclones over the past 30 years. *Nature*, **436**, 686-688, doi:10.1038/nature03906.
- Elsner, J.B., T.H. Jagger, and X.-F. Niu, 2000: Changes in the rates of North Atlantic major hurricane activity during the 20th century. *Geophys. Res. Lett.*, **27**, 1743-1746.
- Elsner, J.B., J.P. Kossin, and T.H. Jagger, 2008: The increasing intensity of the strongest tropical cyclones. *Nature*, **455**, 92-95, doi:10.1038/nature07234.
- Kim, J.-H., C.-H. Ho, M.-H. Lee, J.-H. Jeong, and D. Chen, 2006: Large increase in heavy rainfall associated with tropical cyclone landfalls in Korea after the late 1970s. *Geophys. Res. Lett.*, **33**, L18706, doi:10.1029/2006GL027430.
- Klotzbach, P.J., 2006: Trends in global tropical cyclone activity over the past twenty years (1986-2005). *Geophys Res. Lett.*, **33**, L10805, doi:10.1019/2006GL025881.
- Knaff, J.A., and R.M. Zehr, 2007: Reexamination of tropical cyclone windpressure relationships. Wea. & Forecasting, 22, 71-88.
- Lau, K.-M., Y.P. Zhou, and H.T. Wu, 2008: Have tropical cyclones been feeding more extreme rainfall? J. Geophys. Res., 113, D23113, doi:10.1029/2008JD009963.
- Lowry, M.R., J.J. O'Brien, and M.L. Griffin, 2009: The role of observational changes on empirical trends in tropical cyclone intensity across the western

North Pacific. 2nd International Summit on Hurricanes and Climate Change, Aegean Conference Series, Vol. 41, May 31-June 5, Corfu, Greece, p. 56.

- Trenberth, K., 1990: Recent observed interdecadal climate changes in the Northern Hemisphere. *Bull. Amer. Meteor. Soc.*, **71**, 988-992.
- Tu, J.-Y., C. Chou, and P.-S. Chu, 2009: The abrupt shift of typhoon activity in the vicinity of Taiwan and its association with Western North Pacific-East Asian climate change. J. Climate, 22, 3617-3628.
- Webster, P.J., G.J. Holland, J.A. Curry, and H.-R. Chang, 2005: Changes in tropical cyclone number, duration, and intensity in warming environment. *Science*, 309, 1844-1846.
- Zhao, X., and P.-S. Chu, 2006: Bayesian multiple change-point analysis of hurricane activity in the eastern North Pacific: A Markov chain Monte Carlo approach. J. Climate, 19, 564-578.