Asian Journal of Water, Environment and Pollution, Vol. 5, No. 1, pp. 1-6.

Monthly Frequency Distribution of Severe Tropical Cyclones—A Statistical Approach

Sutapa Chaudhuri and Anindita De Sarkar

Department of Atmospheric Sciences, University of Calcutta
51/2 Hazra Road, Kolkata – 700 019, India
✉ chaudhuri_sutapa@yahoo.com

Received January 10, 2005; revised and accepted March 30, 2007

Abstract: Tropical cyclones are one of the nature's most violent manifestations and potentially the deadliest of all meteorological phenomena. The casualty associated with major cyclones in the Indian sub-continent in the recent past give an idea about its enormous destructive capability. Purpose of the present paper is to identify the months of maximum possibility of occurrence of severe tropical cyclones over Bay of Bengal and Arabian Sea. The methodology adopted in this study is based upon the process of "Statistical Hypothesis Testing". The statistical procedure is applied to fit a Poisson distribution to the monthly frequency distribution of occurrence of severe tropical cyclones over the coasts of Bay of Bengal and Arabian Sea. It is observed that Poisson distribution is a good fit to the monthly frequency distribution of severe tropical cyclones over Bay of Bengal and Arabian Sea. The concept of Shannon entropy is then applied to distinguish the months in accordance with the measure of vulnerability. It is identified that the month of May in the pre-monsoon season and the months of October and November in the post-monsoon season are highly susceptible for the occurrence of severe tropical cyclones over Bay of Bengal and Arabian Sea.

Key words: Severe tropical cyclone, Bay of Bengal, Arabian Sea, Poisson distribution, Shannon entropy, vulnerability.

Introduction

Tropical Cyclone is the generic term for a non-frontal synoptic scale low-pressure system over tropical or sub-tropical water bodies with organized convection and definite cyclonic surface wind circulation (Holland, 1993; 1997). Tropical cyclones with maximum sustained surface wind speed of less than 17 m/s (34 knot, 39 mph) are called "tropical depressions". Once the tropical cyclone reaches winds of at least 17 m/s these are typically termed as "tropical storm" and assigned a name. If winds reach 33 m/s (64 knot, 74 mph), these are called "hurricane" (the North Atlantic Ocean, the North East Pacific Ocean east of the dateline, or the South Pacific Ocean east of 160° E); a "typhoon" (the Northwest Pacific Ocean west of the dateline); a "severe tropical cyclone" (the Southwest Pacific Ocean west of 160° E or the Southeast Indian Ocean east of 90° E); a "severe cyclonic storm" (the north Indian Ocean); and a "tropical cyclone" (the Southwest Indian Ocean) (Chen and Frank, 1993).

The tropical cyclone "basins", where storms occur on a regular basis can be classified as:

- Atlantic basin (including the North Atlantic Ocean, the Gulf of Mexico, and the Caribbean Sea)
- Northeast Pacific basin (from Mexico to about the dateline).
- Northwest Pacific basin (from the dateline to Asia including the South China Sea).
- North Indian basin (including the Bay of Bengal and the Arabian Sea).
- Southwest Indian basin (from Africa to about 100° E).
- Southeast Indian/Australian basin (100° E to 142° E) Australian/Southwest
- Pacific basin (142° E to about 120° W)

The basin considered in this study is North Indian Basin including Bay of Bengal and Arabian Sea. Severe

tropical cyclones over the water basin have immense tendency to harm human life and property. Each year thousands of people are killed at different segments of the coasts of Bay of Bengal and Arabian Sea. It is therefore, a responsibility of the atmospheric scientists to find possible measures to predict the occurrence of the damaging weather event with sufficient lead-time. In the present paper, the months when the water basins have maximum possibility of experiencing severe tropical cyclones are identified using statistical technique.

Methodology

The methodology adopted in this study consists of the following:

- Calculation of Auto Correlation Coefficients (ACC)
- Fitting a Poisson distribution to the monthly distribution of occurrence of Severe Tropical Cyclones (STC) over the North Indian Ocean using the method of Hypothesis Testing
- Calculation of actual and expected probabilities of occurrence of severe tropical cyclones over the said water basin
- Calculations of Shannon entropies for the probability distribution obtained through Poisson distribution to measure the degree of difficulties in predicting the occurrence of severe tropical cyclones in different months.

Auto Correlation Coefficients

Equi-spaced data points constitute a time series. A correlation coefficient measures the degree of association between two time series. In meteorology some time series appear very frequently that are persistent in nature. Persistence means the dependence of a variable on its own past values. The persistence is measured by Auto Correlation Coefficients (ACC). ACC is calculated by shifting the data series a few lags. For a data series with entries $x_1, x_2, \ldots x_n$, the ACC is computed as:

$$\text{ACC} = \frac{\sum_{i=1}^{n-1}(x_i - x_+)(x_{i+1} - x_+)}{\left(\sum_{i=1}^{n-1}(x_i - x_-)^2 \sum_{i=2}^{n}(x_i - x_+)^2\right)^{\frac{1}{2}}} \quad (1)$$

Hypothesis Testing

Hypothesis testing are procedures for making rational decisions about the reality of effects. All hypothesis tests confirm to similar principles and proceed with the same sequence of events.

- A model of the world is created in which there is no effect. The experiment is then repeated an infinite number of times.
- The results of the experiment are compared with the model of step one. If, given the model, the results are unlikely, then the model is rejected and the effects are accepted as real. If, the model could explain the results, the model must be retained. In the latter case no decision can be made about the reality effects.

Any hypothesis testing procedure consists of the following step:

1. Framing of a null hypothesis that assumes some characteristic about the population parameter.
2. Framing of an alternative hypothesis that negates the null hypothesis.
3. Construction of a test statistics on the basis of the null hypothesis.
4. Computation of the statistic and comparison with the corresponding tabular value.
5. Decision on acceptance/rejection of the null hypothesis. Null hypothesis is accepted if the tabular value exceeds the computed value of the statistics (Wilks, 1995).

Poisson Distribution

A random variable X is said to follow Poisson distribution if it satisfies the equation:

$$P(X = x) = \frac{\exp(-\lambda)\lambda^x}{x!} \quad (2)$$

where $x \in I$ and $\lambda > 0$.

Poisson distribution can present the events having significant degrees of randomness.

Shannon Entropy

The concept of *Shannon's entropy* (Chaudhuri et al., 2004) is the central role of the information theory sometimes referred as *measure of uncertainty*. The entropy of a random variable defined in terms of its probability distribution can be shown to be a good measure of randomness or uncertainty.

Let $X \rightarrow$ random variable and

$P \rightarrow$ space of all finite probability distributions:

$$P := \{f / f : dom(X) \mapsto [0,1], \; x \in dom(X) = \{s_1, \ldots, s_n\}, n = 1, 2, \ldots\} \quad (3)$$

The entropy associated with this probability distribution is defined as

$$H(X) = H(f(x) / x \in dom(X)) := -\sum_{x \in dom(X)} f(x) \log_2 f(x) bits \quad (4)$$

where *dom (X)* → value-set of variable X, $x \in X$ a specific value and $f \rightarrow$ probability distribution of X.

Chi-Square (χ^2) Test for Goodness of Fit of a Distribution

Fitting of a distribution to a dataset is done by the method of Chi-Square (Chaudhuri et al., 2004). In this method, a distribution is assumed to be a good fit to a given dataset. On the basis of this assumption, a set of expected frequencies are calculated for a set of observed frequencies. Using the observed and expected frequencies, a Chi-Square statistic is framed as follows:

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \quad (5)$$

where O_i = Observed frequency for the i^{th} cell and E_i = Expected frequency for the i^{th} class and n = Number of observations.

Poisson is assumed to be a good fit to the aforesaid frequency distribution when the computed chi-square is found to be expected by the tabular value.

Results and Discussions

The monthly distribution of severe tropical cyclone occurring over Bay of Bengal and Arabian Sea within the period from 1891 to 1989 (Mandal, 1991) is considered in this study.

First order ACC (Chaudhuri et al., 2004) is computed to check the persistence in the monthly frequency distributions of STC over Bay of Bengal and Arabian Sea. It is observed that all the ACC are very low in all the months. Thus, all the months have high degree of randomness with respect to occurrence of STC. The ACC are shown in Table 1. The ACC are also schematically presented in the Figure 1. It is thus, to be seen whether Poisson distribution can be a replica of the aforesaid weather event.

It is apparent from the result that all the computed chi-squares are less than the tabular values. It thus, leads to state that the null hypothesis is acceptable for the monthly frequency distributions of severe tropical cyclones over the water bodies. The results are shown in Figure 2 and data given in Table 2.

Poisson is, therefore established to be a good fit to the monthly frequency distribution of severe tropical cyclones over the water bodies under consideration. This implies that to predict the occurrences of severe tropical cyclones over Bay of Bengal and Arabian Sea, Poisson distribution can be an ideal model. Moreover, this

Table 1: A tabular presentation of the ACC computed from the frequency distribution of severe tropical cyclones over the North Indian Ocean between 1891 and 1989

Months	*Lag-1 ACC*
January	0.06
February	0.01
March	0.22
April	0.17
May	0.17
June	0.04
July	0.16
August	0.04
September	0.12
October	0.06
November	0.11
December	0.06

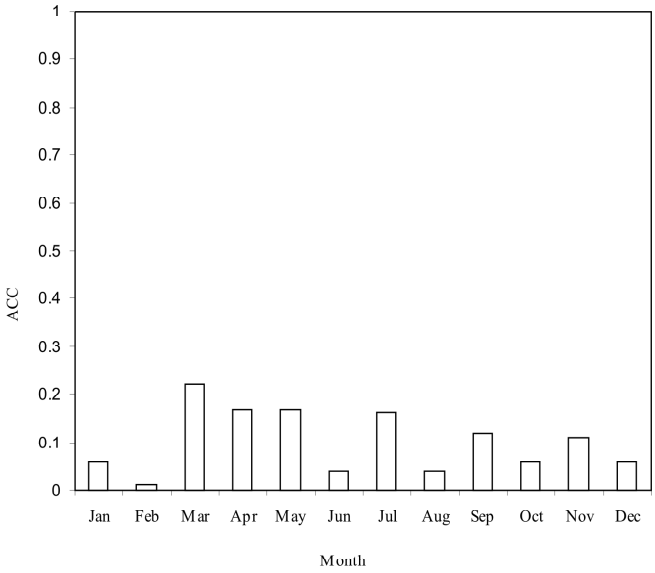

Figure 1: The ACC associated with the monthly frequencies of STC over Bay of Bengal and Arabian Sea.

goodness of fit of Poisson distribution to the frequencies of all months shows the randomness of occurrence of severe tropical cyclones over the water bodies.

The next job is to compare the degrees of randomness in the occurrence of the severe weather event in each month of a year. Shannon entropy is applied for the purpose. The Shannon entropies are computed separately for each month and presented in the Table 3.

It is apparent from the above table that October and November have the largest uncertainties in occurrence, thus in prediction. These two months are most vulnerable due to severe tropical cyclones. This is further supported

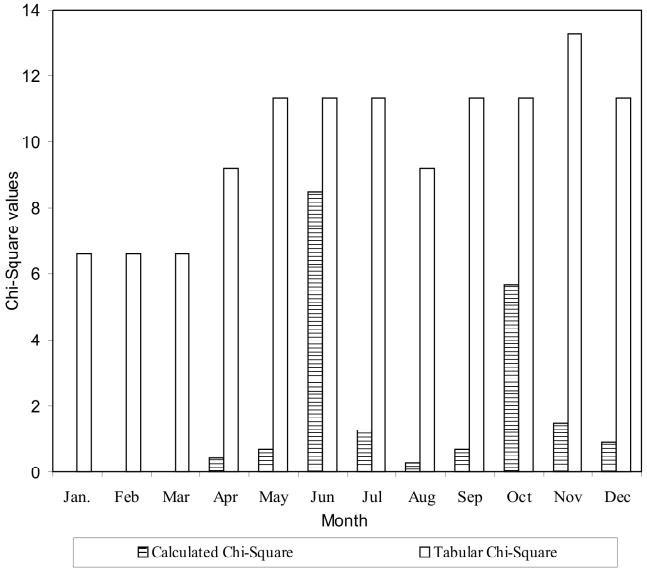

Figure 2: A schematic showing the comparison between the computed and tabular chi-square values.

Table 2: A tabular presentation of the computed and tabular values of Chi-square obtained on the basis of the assumption of good fit of Poisson distribution to the monthly frequency distribution of severe tropical cyclones

Months	*Computed* χ^2	*Tabular* χ^2
January	0.004	6.635
February	0.0001	6.635
March	0.007	6.635
April	0.4249	9.21
May	0.6934	11.345
June	8.467	11.345
July	1.279	11.345
August	0.295	9.21
September	0.6934	11.345
October	5.6843	11.345
November	1.479	13.277
December	0.911	11.345

by the fact that these two months have the lowest probabilities of occurrence of severe tropical cyclones (Table 4). Moreover, in these two months probabilities of other frequencies have significant values. Vulnerability of these two months for severe tropical cyclones is further supported by the versatility with respect to probability. The findings are shown in Figures 3 and 4.

It is also apparent from the results that in May there is significant possibility of occurrence of severe tropical cyclone over the water bodies.

Table 3: A tabular presentation of the Shannon entropies associated with the probability distributions of monthly frequency distribution of severe tropical cyclones over North Indian Ocean between 1891 and 1989

Month	*Shannon Entropy*
January	0.31796
February	0.08078
March	0.23786
April	0.98211
May	1.53453
June	1.45167
July	1.25081
August	1.02311
September	1.27647
October	1.76271
November	1.95773
December	1.27203

Table 4: A tabular presentation of the probabilities of monthly frequencies of STC over North Indian Ocean between 1891 and 1989

Month	*0STC*	*1STC*	*2STC*	*3STC*	*4STC*
January	0.95	0.05	0	0	0
February	0.99	0.01	0	0	0
March	0.96	0.04	0	0	0
April	0.76	0.2	0.04	0	0
May	0.5	0.34	0.12	0.04	0
June	0.56	0.31	0.1	0.02	0.01
July	0.64	0.29	0.06	0.01	0
August	0.72	0.24	0.04	0	0
September	0.62	0.3	0.07	0.01	0
October	0.36	0.37	0.18	0.09	0
November	0.31	0.36	0.21	0.08	0.04
December	0.63	0.29	0.07	0.01	0

Conclusion

The results obtained lead to conclude that Poisson is a good fit to the monthly frequency distribution of occurrence of severe tropical cyclones over Bay of Bengal and Arabian Sea. Thus, Poisson distribution can be used as a statistical predictive model for forecasting the occurrence of severe tropical cyclones over Bay of Bengal and Arabian Sea.

Poisson being a good fit to the data, it can be concluded that all months have some degree of vulnerability to severe tropical cyclones. But, October and November are most vulnerable to severe tropical cyclones. It is also concluded that in May there is significant possibility of occurrence of severe tropical cyclone over the water bodies.

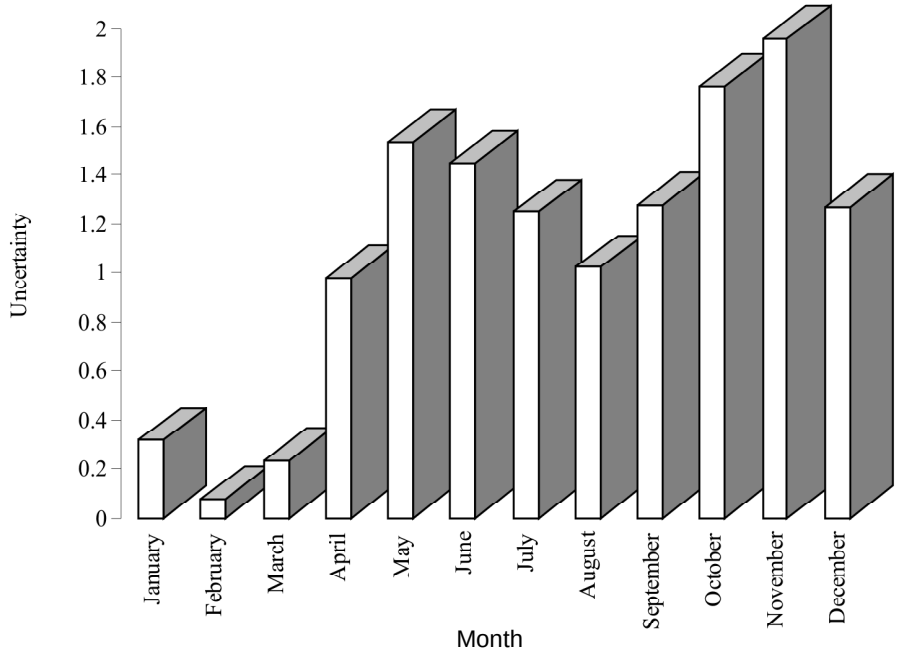

Figure 3: Uncertainties associated with probability distributions of severe tropical cyclones in different months.

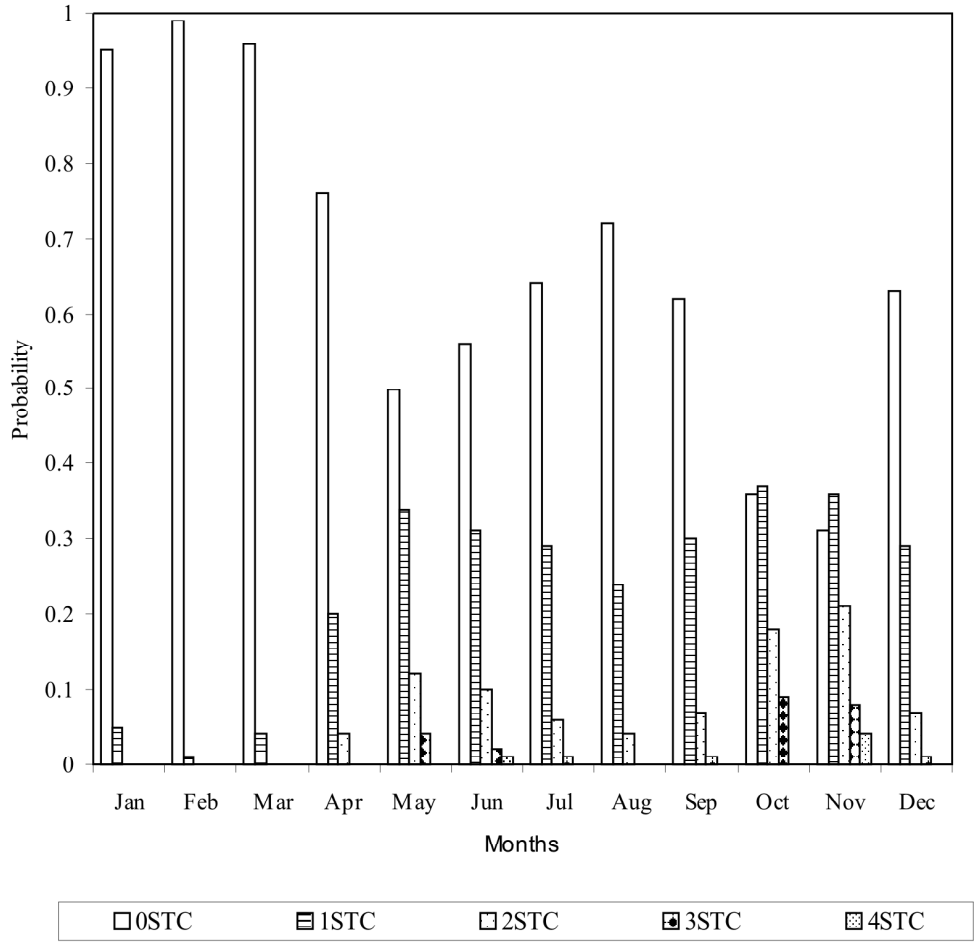

Figure 4: Probabilities of different frequencies of severe tropical cyclones in different months.

Thus, May in the pre-monsoon season and October and November in the post monsoon period is highly vulnerable for the occurrence of severe tropical cyclones.

References

Chaudhuri, Sutapa and S. Chattopadhyay (2004). "Tropical cyclones—Statistical Perspectives", *Mausam*, **55, 2:** 360-364.

Chaudhuri, Sutapa and S. Chattopadhyay (2004). "Identification of coasts vulnerable for severe tropical cyclones—Statistical elucidation.", *Mausam*, **55, 3:** 502-507.

Chaudhuri, Sutapa and S. Chattopadhyay (2004). "Predictability of severe tropical cyclones over North Indian Ocean—Statistical Exposition", *J. Science & Culture*, **70, 4 – 5:** 186-188.

Chen, S.A. and W.M. Frank (1993). "A numerical study of genesis of the extra tropical convective meso-vortices. Part I: Evolution and dynamics" *J. Atmos. Sci*. **50:** 2401-2426.

Holland, G.J. (1993). "Ready Reckoner"-, WMO/TC-No. 560, Report No. TCP-31, *World Meteorological Organization*; Geneva.

Holland, G.J. (1997). " The maximum potential intensity of tropical cyclones", *Journal Atmos. Sci.*, **54:** 2519-2541.

Mandal, G.S. (1991). Tropical Cyclone Programme, Report Number TCP-28, WMO/TD-No. 430, WMO, Geneva.

Wilks, D.S. (1995). Statistical Methods in Atmospheric Sciences, *Academic Press,* USA.