551.515.9 : 519.21

IDENTIFICATION OF COASTS VULNERABLE FOR SEVERE TROPICAL CYCLONES – STATISTICAL ELUCIDATION

1. Severe Tropical Cyclone (STC) is among the most destructive natural hazards, initiating catastrophic ravages to life, property and environment.

An accurate prediction of occurrence of STC with adequate lead time or forecasting a definite trend of occurrence of these cyclonic storms over different coasts is essential for mitigating the squandering due to such casualty.

The purpose of the present study is, thus, to identify those coasts over Bay of Bengal, which are vulnerable for the occurrence of STC by using method of statistical techniques.

2. It would suffice to acquire the intent of the study if the following could be ordained;

(*i*) Occurrence of STC is random over the concerned coasts.

(*ii*) Degree of uncertainty in forecasting the occurrence of STC is high over such coasts.

(*iii*) Veering of TC to STC represents a linear regression relationship over the coasts.

2.1. *Fitting of Poisson distribution* - Severe Tropical Cyclones being discrete events, a Poisson distribution is tried to fit to the data under consideration. The probability distribution function (pdf) for Poisson distribution is ;

$$P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!} \quad (1)$$

Where,

x varies from 0 to ∞

and, λ implies the average occurrence rate.

The recurrence relation is;

$$P(X = x+1) = \frac{\lambda}{x+1} P(X = x) \quad (2)$$

Different probability levels are computed for each of the coasts under consideration by using expression (1) and (2). Expected decadal frequency distributions have also been computed for each coast by multiplying these probability levels with total frequency of STC over the coasts.

To observe whether Poisson distribution fits well with the decadal frequency distribution of STC, a null hypothesis, H_0 is defined;

H_0 : Poisson distribution is a good fit.

This proposition is tested against an alternative hypothesis, H_1;

H_1 : Poisson distribution is not a good fit.

Under the null hypothesis, a Chi-square (χ^2) statistic is computed ;

$$\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i} \tag{3}$$

Where,

O_i = Observed frequency of the i^{th} class.

E_i = Expected frequency of the i^{th} class.

Computed values of Chi-square are compared with the standard tabular (Wilks, 1995) values with appropriate degree of freedom. The null hypothesis is accepted if the tabular value exceeds the computed value otherwise, rejected. Tests have been carried out at 5% level of significance. A false acceptance of null hypothesis generates type - II error. In any situation, if the null hypothesis is accepted but either skewness or kurtosis is observed to be negative, type - II error appears. To reach an error free conclusion, results with type - II error are to be identified.

2.2. *Measure of uncertainty* - Shannon (1948) introduced the concept of entropy that measures the degree of uncertainty. Shannon entropy is defined as;

$$\mathrm{H}[(p(x)/x \in X)] = -\sum_{x \in X} p(x) \log_2 p(x) \tag{4}$$

where,

$[p(x)/x \in X] \Rightarrow$ Probability distribution on a finite set X.

In the present study different probability levels are used to observe the uncertainty associated with the occurrence of STC or the anticipation of development of TC to STC using equation (4).

2.3. *Analysis of variance* - A linear regression equation of the form;

$$\hat{y} = \mathrm{a} + \mathrm{b}x \tag{5}$$

is tried to fit to the data set of decadal frequency of TC and STC over the coasts.

Frequency of TC is considered to be the predictor (x) and frequency of STC as predictand (y). $\hat{y}$ represents the predicted value of y.

In expression (5),

$$b = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} (x)^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \tag{6}$$

and

$$\mathrm{a} = \bar{y} - \mathrm{b}\bar{x} \tag{7}$$

where $n \rightarrow$ number of observations.

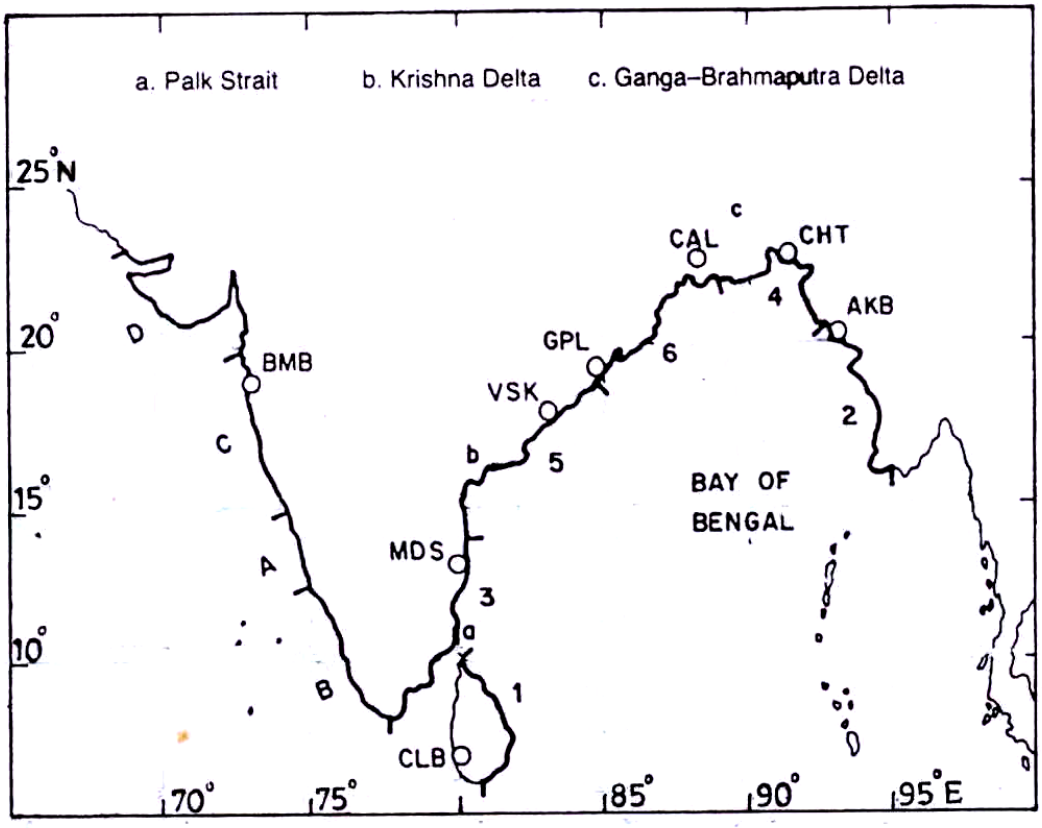
Analysis of variance (ANOVA) is executed to test the goodness of fit of the regression to the observed data. Total sum of squares (SST), regression sum of squares (SSR), sum of squared error (SSE), mean squared regression (MSR), mean squared errors (MSE) are computed for this purpose;

$$\mathrm{SST} = \sum_{i=1}^{n} {y_i}^2 - n\bar{y}^2 \tag{8a}$$

$$\mathrm{SSR} = b^2\left[\sum_{i=1}^{n} {x_i}^2 - n\bar{x}^2\right] \tag{8b}$$

$$\mathrm{SSE} = \sum_{i=1}^{n} [y_i - \hat{y}(x_i)]^2 \tag{8c}$$

$$\mathrm{MSE} = \frac{1}{n-2}(\mathrm{SST} - \mathrm{SSR}) \tag{8d}$$



1. East Sri Lanka, 2. Arakan, 3. Tamil Nadu, 4. Bangladesh, 5. Andhra Pradesh, 6. Orissa-West Bengal,
A – Karnataka, B – Kerala, C – Maharashtra-Goa, D – Gujarat

Fig. 1. The Bay of Bengal and Arabian Sea. The segments of the Bay coast are marked in the figure by numbers 1 to 6. Segments of west coast of India are marked by the letters A, B, C, D. The cities marked are : AKB, Akyab, BMB, Mumbai (Bombay); CAL, Calcutta (Kolkata); CHT, Chittagong; CLB, Colombo; GPL, Gopalpur; MDS, Madras (Chennai); VSK, Visakhapatnam

TABLE 1

Fitting of Poisson distribution and amount of uncertainty associated with the probability distribution

Coasts	Skewness	Kurtosis	Observed Chi-square	Degree of freedom	Tabular Chi- square	Decission regarding null hypothesis	Type of error	Amount of uncertainty
TN	0.96	1.23	0.001	3	7.815	accepted	No error	2.713
AP	0.46	-1.32	0.090	5	11.070	accepted	Type II	2.646
O/WB	-0.33	-1.07	1.020	4	9.488	accepted	Type II	2.530
BD	1.73	2.30	2.040	2	5.991	accepted	No error	2.766
AR	-0.54	-1.13	2.057	1	3.841	accepted	Type II	1.881
ESL	0.86	-2.51	0.699	1	3.841	accepted	Type II	0.897

TN - Tamilnadu coast, AP – Andhra Pradesh coast, O /WB - Orissa & West Bengal coast,
BD - Bangladesh coast, AR - Arakan coast, ESL - East SriLanka coast

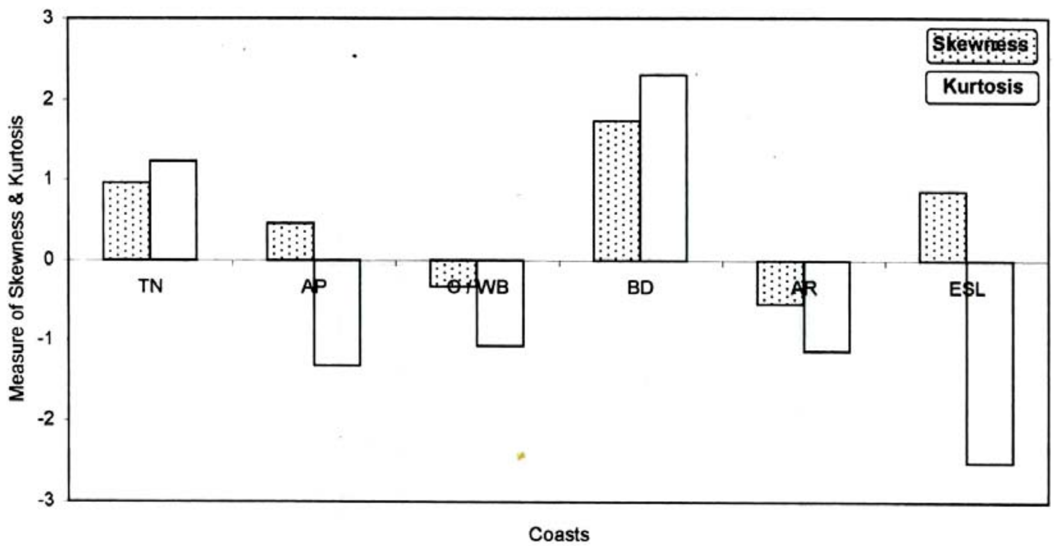


Fig. 2. Randomness of occurrence of severe tropical cyclones over different coasts

$$\text{MSR} = \frac{\text{SSR}}{1} = \text{SSR} \qquad (8e)$$

F-statistic is then constructed as ;

$$F = \frac{\text{MSR}}{\text{MSE}} \qquad (9)$$

Construction of F-statistic is based on the null hypothesis H_{00}

H_{00} : regression is not a good fit.

This is tested against an alternative hypothesis H_{11}.

H_{11} : regression is a good fit.

If the computed value of F exceeds the tabular value with suitable degree of freedom, then the null hypothesis will be rejected otherwise accepted.

Measure of goodness of fit of a regression equation is the coefficient of determination R^2 that is defined as:

$$R^2 = \frac{\text{SSR}}{\text{SST}} \qquad (10)$$

3. The present study is executed with 110 years data available in Report No. TCP-28, TD-No. 430, WMO, Geneva (Mondal, 1991).

The coasts considered for this study are the segments of Bay coasts comprising Tamilnadu (TN), Andhra

TABLE 2

ANOVA table with coefficient of determination for linear regression

Source	Degree of freedom	Sum of squares	Mean of squares	F- Statistic	Coefficient of determination
East Sri Lanka					
Total	10	2.574			
Regression	1	1.058	1.058	6.30	0.41
Residual	9	1.516	0.168		
Tamil Nadu					
Total	10	27.018			
Regression	1	20.104	20.104	26.18	0.74
Residual	9	6.917	0.768		
Andhra Pradesh					
Total	10	35.524			
Regression	1	7.503	7.503	2.41	0.21
Residual	9	28.021	3.110		
Orissa / West Bengal					
Total	10	41.272			
Regression	1	2.626	2.626	0.61	0.06
Residual	9	38.646	4.294		
Bangladesh					
Total	10	116.903			
Regression	1	86.570	86.570	26.68	0.74
Residual	9	30.333	3.370		
Arakan					
Total	10	11.644			
Regression	1	3.024	3.024	3.87	0.26
Residual	9	8.620	0.780		

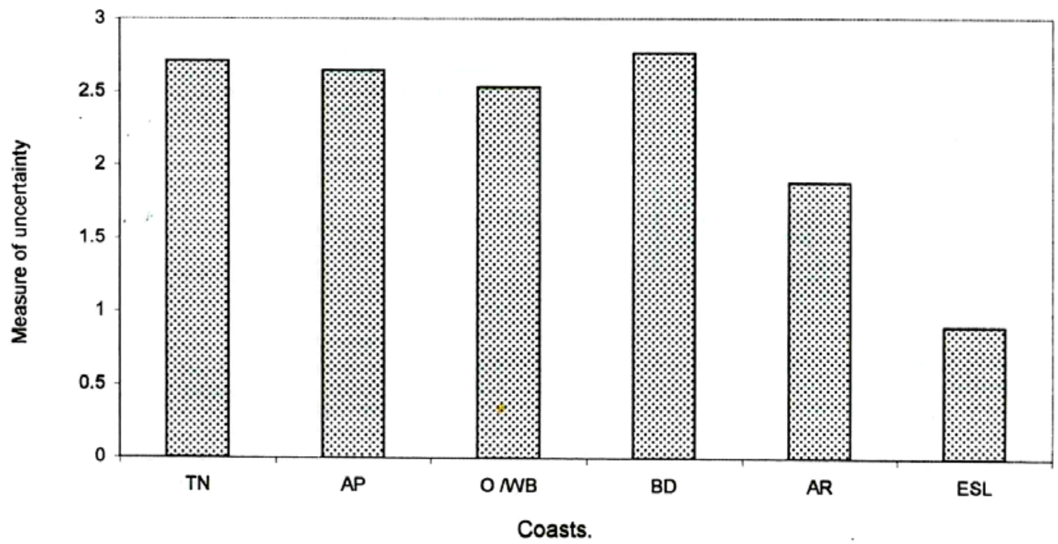
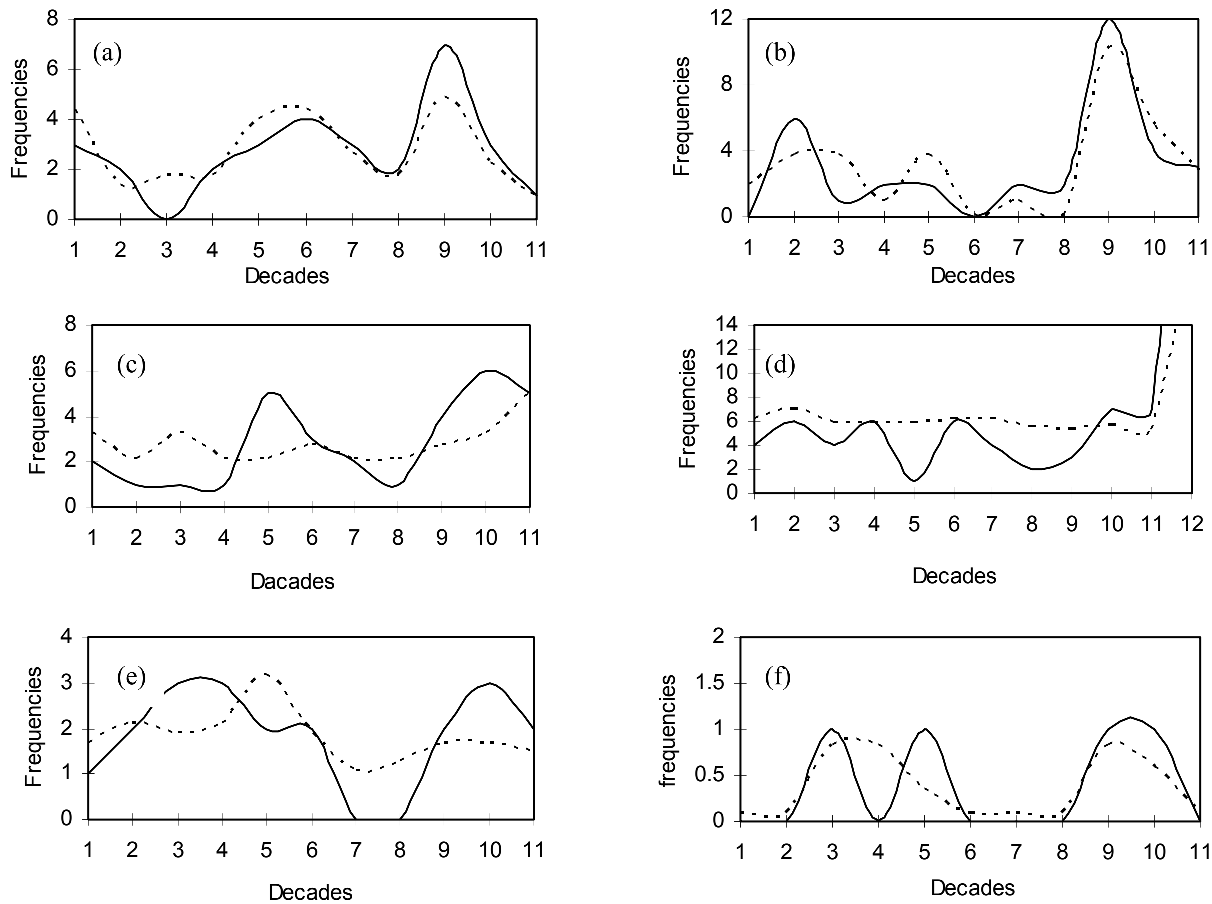


Fig. 3. Uncertainty of occurrence of severe tropical cyclones (STC) over different coasts



Figs. 4(a-f). Actual (—) and predicted (---) decadal frequencies of STC over different segments of Bay coast. (a) Tamilnadu coast (1881-1989), (b) Bangladesh coast (1881-1989), (c) Andhra Pradesh coast (1881-1989), (d) Orissa / West Bengal coast (1881-1989), (e) Arakan coast (1881-1989) and (f) East-Sri Lanka coast (1881-1989)

Pradesh (AP), Orissa/West Bengal (O/WB), Bangladesh (BD), Arakan (AR), and East Srilanka (ESL) within 5° N to 22° N latitude and 80° E to 95° E longitude (Fig. 1).

Values of Chi-square are computed and compared with the standard table for all these coasts at 5% level of significance with suitable degree of freedom (Table 1). The computed values of Chi-square are found to be less than the tabular values. Tamilnadu and Bangladesh coasts are observed to own both skewness and kurtosis to be positive (Table 1). It is, thus, perceived that the acceptance of null hypothesis for all the coasts under consideration except the coasts of TN and BD involve type-II error. It can, therefore, be affirmed that decadal frequency distribution of STC over the coasts of TN and BD imitate Poisson distribution. This indicates that the occurrence of STC over these two coasts has high degree of randomness. This is made schematically apparent in Fig. 2. Fig. 3 exhibits that the measure of uncertainty in the occurrence of STC is also high over the coasts of Tamilnadu and Bangladesh.

Figs. 4(a&b) reveal that the existence of linear regression relationship between the decadal frequencies of Tropical Cyclone (TC) and Severe Tropical Cyclone (STC) over the coasts of TN and BD. No such relationship is available for the other coasts under consideration [Figs. 4(c-f)]. The equations tested for the purpose are :

$$\hat{y}_{\text{ESL}} = 0.10 + 0.24x$$
$$\hat{y}_{\text{TN}} = 0.4 + 0.44x$$
$$\hat{y}_{\text{AP}} = -1.24 + 0.56x$$
$$\hat{y}_{\text{O/WB}} = 4.45 + 0.10x$$
$$\hat{y}_{\text{BD}} = -2.66 + 0.92x$$
$$\hat{y}_{\text{AR}} = 1.06 + 0.21x$$

Table 2 displays that the computed values of '*F*' exceeds the standard tabular values at, 1% level of significance, only for the coasts of TN and BD. For the coast of ESL, on the other hand, the null hypothesis is rejected at 5% level of significance.It is mentioned here that the tabular values of $F_{0.05;1,9}$ and $F_{0.01;1,9}$ are 5.12 and 10.56 respectively. It is also indicative from the table that the coasts of TN and BD hold high coefficient of determination. Coefficient of determination is neither high nor low over the coast of East Sri Lanka (ESL). No specific decision is, thus, possible for the randomness of veering of TC's to STc's over the coast of ESL, whereas, over the coasts of TN and BD frequency of the occurrence of STC is related linearly to the frequency of occurrence of TC.

4. Among the different segments of Bay coast under consideration, the coasts of Tamilnadu and Bangladesh are most vulnerable for Severe Tropical Cyclones as the anticipation of occurrence of such weather calamities are difficult over these two coasts.

References

Mandal, G. S., 1991, "Tropical cyclone program", Report No. TCP - 28, WMO/TD - No.**430**, WMO, Geneva.

Shannon, C. E., 1948, "The mathematical theory of communications", *The Bell System Technical Journal*, **27**, 379-423.

Wilks, Daniel, S., 1995, "Statistical methods in atmospheric sciences", *Academic Press*, USA.

SUTAPA CHAUDHURI
SURAJIT CHATTOPADHYAY

Department of Atmospheric Sciences,
University of Calcutta,
Kolkata - 700 009, India
(*5 October 2001, Modified 7 July 2003*)