Natural Hazards manuscript No.
(will be inserted by the editor)

A parametric Markov renewal model for predicting tropical cyclones in Bangladesh

Md Asaduzzaman · A. H. M. Mahbub Latif

Received: date / Accepted: date

Abstract In this paper we consider a Markov renewal process to model tropical cyclones occurred in Bangladesh during 1877–2009. The model takes into account both the occurrence history and some physical constraints to capture the main physical characteristics of the storm surge process. We assume that the sequence of cyclones constitutes a Markov chain, and sojourn times follow a Weibull distribution. The parameters of the Weibull Markov renewal process jointly with transition probabilities are estimated using the maximum likelihood method. The model shows a good fit with the real events, and probabilities of occurrence of different types of cyclones are calculated for various lengths of time interval using the model. Stationary probabilities and mean recurrence times are also calculated. A brief comparison with a Poisson model and a marked Poisson model has also been demonstrated.

Keywords Cyclone prediction · marked Poisson process · Poisson process · semi-Markov process · statistical estimation of Markov renewal process

---

Md Asaduzzaman
Mathematics and Statistics Division
Faculty of Computing, Engineering and Sciences
Staffordshire University
Stoke-on-Trent ST4 2DE
United Kingdom
Tel.: +44 (0)1782 294707
E-mail: md.asaduzzaman@staffs.ac.uk

A. H. M. Mahbub Latif
Institute of Statistical Research and Training (ISRT)
University of Dhaka, Dhaka 1000
Bangladesh
E-mail: mlatif@isrt.ac.bd

1 Introduction

Bangladesh often suffers from devastating tropical cyclones due to its unique location. The Bay of Bengal in the southern part of Bangladesh is one of the world's most active areas for the development of tropical disturbances. The Bay of Bengal has a funnel-shaped northern part that causes tidal bores when cyclones make landfall, and thousands of people living in the coastal areas of Bangladesh are affected. In the past century, Bangladesh experienced two of the world's deadliest tropical cyclones in the years 1970 and 1991, which killed about 300 thousand and 140 thousand people, respectively. Recently, two devastating tropical cyclones *Sidr* (in 2007) and *Aila* (in 2009) caused about USD 2.3 million and USD 0.27 million property damage, respectively (Centre for Research on the Epidemiology of Disasters). Even storms with low intensity can be very deadly at landfall for Bangladesh compared with the developed countries because of its poor housing conditions and lack of effective early warning systems.

Although a number of studies have been carried out on cyclones that struck around the Bay of Bengal, a little attention has been paid to the Bangladesh coast (see Islam and Peterson, 2009, and references therein). Among the published works on the cyclones in Bangladesh, several studies have emphasized on the frequency, vulnerability, warning process, etc. (e.g., Khalil, 1992; Haque and Blair, 1992; Haque, 1995; Ali, 1996). However, Dube et al (1985, 1986) and Sinha et al (1986) proposed numerical simulation models for describing the storm surge process in Bangladesh. Among others, Mooley (1981) proposed a Poisson process for analyzing severe cyclonic storms that struck the coast around the Bay of Bengal during 1877–1977. However, there is no study found in the literature that focused on an extensive investigation of cyclone occurrences in the Bay of Bengal region through stochastic modelling except the one by Asaduzzaman and Latif (2014).

Mathematical models such as stochastic processes specially counting processes are often considered for describing different types of natural hazards, e.g. earthquakes (Ogata, 1998, 1988; Gospodinov and Rotondi, 2001), drought (Gupta and Duckstein, 1975), flood (Fiorentino et al, 1984), storm (Rumpf et al, 2007), etc., as they are able to capture physical properties of a natural phenomenon. The well-known form of counting process, the Poisson model, assumes that sojourn times are exponential random variables. The Poisson model or some of its variants are considered in several studies for describing cyclone occurrences, for instance, Lu and Garrido (2005) applied a doubly periodic non-homogeneous Poisson model, and Jagger et al (2002) proposed a space-time model for hurricane data. Among other modelling approaches that are available in the literature to model natural hazards, Markov renewal process (MRP) and semi-Markov process (SMP) have received considerable focus (e.g., Gregory et al, 1993; Lardet and Obled, 1994; Alvarez, 2005; Garavaglia and Pavani, 2011; Masala, 2012a).

An MRP is a semi-Markov type process that visits a finite number of states and spends a random amount of time in a particular state. In seismic hazard analysis Alvarez (2005) argued that Poisson model may not be appropriate for large earthquakes as the process is inherently memoryless, and proposed a Markov renewal model for studying earthquake occurrences in North Anatolian Fault Zone in Turkey. Using the same data Garavaglia and Pavani (2011) proposed an MRP where a mixture of exponential and Weibull distributions were used for the waiting time distribution of the large earthquakes. Then the so-called 'crossing state probabilities' are estimated. Masala (2012a) applied a discrete time semi-Markov model to estimate earthquake occurrences in Italy. To analyze storm risk Masala (2012b) proposed a discrete-time homogeneous semi-Markov model to describe the evolution of lifespan of hurricanes, where the model was validated using a hurricane database, and cross state predictions were performed. However, the parameters of transition probability matrices were first estimated assuming a sample from a multinomial distribution, and then using these estimates, sojourn time parameter estimation was performed in Garavaglia and Pavani (2011), Masala (2012a), Masala (2012b). Recently, a non-parametric approach of a semi-Markov model was adopted by Votsi et al (2012) for the estimation of earthquake occurrences in the Northern Aegean Sea in Greece.

In the present paper we propose a parametric MRP in continuous time that investigates occurrences of land-falling tropical cyclones in Bangladesh. It provides a detail of estimation of parameters for an MRP, and while estimating parameters, likelihood function is maximized both for parameters of the transition probability matrix and sojourn time distributions simultaneously unlike Garavaglia and Pavani (2011), Masala (2012a) and Masala (2012b). Model building strategies were performed by testing a series of nested hypotheses whether we can further reduce the sojourn time parameters to find an optimal model.

The paper is organized as follows. In Section 2, we define a general Markov renewal processes, together with some details in likelihood construction, cross state prediction and asymptotic behavior. Section 3 contains the analysis of the data containing 125 tropical cyclones occurred in Bangladesh during 1877–2009 using both Poisson and marked Poisson processes. Then a Markov renewal model is fitted, and cross state prediction and estimation of recurrence time are performed. Some final comments are given in Section 4.

2 Markov renewal process

Let $\{J_n : n \geq 0\}$ be a Markov chain that successively visits some states, and let $\{X_n, n \geq 0\}$ be a sequence of random variables that represents the successive sojourn times corresponding to the visits of such states. Then the two-dimensional process $\{(J_n, X_n) : n \geq 0\}$ is called a Markov renewal process (MRP), where $X_0 = 0$ is assumed. The process $\{Z_k; k \geq 0\}$ is said to be a semi-Markov process associated to the Markov renewal process $\{(J_n, X_n)\}$ if $\{Z_k = J_{N(k)}; k \geq 0\}$, where $N(k) = \max\{n \,|\, X_n \leq k\}$.

Suppose $\mathbf{E} = \{1, \ldots, M\}$ be the set of all possible states of J_n, where $M \in \mathcal{N}$. Let $\mathbf{H} = \big\{H_i(\cdot),\ i \in \mathbf{E}\big\}$ be a vector of distribution functions of size M on $\Re_+$, $\mathbf{F} = \big\{F_{ij}(\cdot),\ i, j \in \mathbf{E}\big\}$ be an $M \times M$ matrix of distribution functions on $\Re_+$, $\mathbf{P} = \{(p_{ij}),\ i, j \in \mathbf{E}\}$ be an $M \times M$ transition probability matrix, and $\mathbf{a} = (a_1, \ldots, a_M)$ be a probability distribution on $\mathbf{E}$ with $a_i \geq 0$ and $\sum_{i \in \mathbf{E}} a_i = 1$. The Markov renewal process is then defined by the quadruplet $(\mathbf{E}, \mathbf{a}, \mathbf{P}, \mathbf{F})$, and some important properties of a Markov renewal process (see Pyke, 1961; Limnios and Oprian, 2001; Janssen and Manca, 2007, for details) are

$$(i)\ \mathbb{P}(J_n = j_n \mid J_0 = j_0, \ldots, J_{n-1} = j_{n-1}) = p_{j_{n-1}, j_n}$$

$$(ii)\ \mathbb{P}(X_n \leq x \mid J_0 = j_0, \ldots, J_{n-1} = j_{n-1}) = H_{j_{n-1}}(x)$$

$$(iii)\ \mathbb{P}(X_n \leq x \mid J_0 = j_0, \ldots, J_{n-1} = j_{n-1}, J_n = j_n) = F_{j_{n-1}, j_n}(x)$$

$$(iv)\ \mathbb{P}(X_1 \leq x_1, \ldots, X_n \leq x_n \mid J_n = j_n, n \geq 0) = \prod_{i=1}^{n} F_{j_{i-1}, j_i}(x_i)$$

$$(v)\ \mathbb{P}(J_n = j_n, X_n \leq x \mid J_0 = j_0, \ldots, J_{n-1} = j_{n-1}, X_{n-1} = x_{n-1}) = p_{j_{n-1}, j_n} F_{j_{n-1}, j_n}(x).$$

The J-process is a homogeneous Markov chain with transition probability matrix $\mathbf{P}$, which is also known as the imbedded Markov chain of the Markov renewal process. The random variable J_n represents the state of the system just after the nth transition. Assume that the process is homogeneous in time and aperiodic (i.e. ergodic), then it follows that there exists one and only one stationary distribution $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_M)$, where

$$\pi_j = \lim_{n \to \infty} \mathbb{P}(J_n = j \mid J_o = i) = \lim_{n \to \infty} p_{ij}^{(n)}$$

is independent of i. Then π_j is the unique non-negative solution of

$$\pi_j = \sum_{i=1}^{M} \pi_i \, p_{ij}, \quad \sum_{j=1}^{M} \pi_j = 1, \quad i, j = 1, \ldots, M.$$

From the properties of the Markov renewal process (ii) and (iii), it is clear that two types of sojourn time distributions

$$F_{j_{n-1}, j_n}(x) = \mathbb{P}(X_n \leq x | J_{n-1} = j_{n-1}, J_n = j_n) \text{ and } H_{j_{n-1}}(x) = \mathbb{P}(X_n \leq x | J_{n-1} = j_{n-1})$$

are involved, which are called the conditional and unconditional distribution functions of the sojourn time X_n, respectively. Simplifying the subscripts we write $F_{ij}(x)$ and $H_i(x)$ instead of $F_{j_{n-1}, j_n}(x)$ and $H_{j_{n-1}}(x)$. We assume that the distribution F_{ij} have a density f_{ij}, for every $i, j = 1, \ldots, M$. Hence, from (iv) one can deduce, if an event is classified as of type $i \in M$ and the successive event is of type $j \in M$, the time between the two events is a positive random variable that follows a distribution with distribution function F_{ij}.

Now using the conditional expectation, we can express the unconditional distribution functions in terms of the conditional distribution functions (see Pyke, 1961; Limnios and Oprian, 2001; Janssen and Manca, 2007, for details) in simplified notation as

$$H_i(x) = \sum_{j=1}^{M} p_{ij} \, F_{ij}(x). \tag{1}$$

If exist, means of the conditional and unconditional distribution of sojourn times can be expressed as

$$\nu_{ij} = \int x \, dF_{ij}(x) \text{ and } \eta_i = \int x \, dH_i(x), \quad \forall \, i, j = 1, \ldots, M,$$

respectively. Then the equation (1) leads to the following relation

$$\eta_i = \sum_{j=1}^{M} p_{ij} \nu_{ij}. \tag{2}$$

2.1 Likelihood function and parameter estimation

Let $(j_0, j_1, x_1, \ldots, j_{\tau-1}, x_{\tau-1}, x_\tau)$ be a realization of a Markov renewal process on the time window $[0, T]$, where τ represents the number of states visited in $[0, T]$ and x_τ, the sojourn time between the last event J_τ and T, is considered as censored, i.e. $x_\tau > [T - (x_1 + \cdots + x_{\tau-1})]$. Then the conditional likelihood function given $J_0 = j_0$ can be expressed as

$$L(j_0) = \left[\prod_{i=1}^{\tau-1} p_{j_{i-1}, j_i} f_{j_{i-1}, j_i}(x_i) \right] \times \left[\sum_{k=1}^{M} p_{j_{\tau-1}, k} \left(1 - F_{j_{\tau-1}, k}(x_\tau) \right) \right]$$

and the corresponding log-likelihood function is

$$l(j_0) = \sum_{i=1}^{\tau-1} \ln p_{j_{i-1},j_i} + \sum_{i=1}^{\tau-1} \ln f_{j_{i-1},j_i}(x_i) + \ln \left[\sum_{k=1}^{M} p_{j_{\tau-1},k} \left(1 - F_{j_{\tau-1},k}(x_\tau)\right) \right]. \quad (3)$$

There are two types of parameters in the log-likelihood function $l(j_0)$, which are: (i) parameters of the transition probability matrix $\{(p_{ij}),\ i,j = 1,\ldots,M\}$ and (ii) parameters of the distribution of sojourn times corresponding to different transitions.

The maximum likelihood estimates are obtained by maximizing the log-likelihood over the time window $[0,T]$. To maximize the log-likelihood, initial estimator of the elements of the transition probability matrix are obtained using the following expression

$$\widehat{p}_{ij}^{0} = \frac{\#\text{ of transitions from state } i \text{ to state } j}{\#\text{ of transitions from state } i \text{ to all states}},$$

which can be derived by considering the data on different transitions over the period $[0,T]$ as a sample from a multinomial distribution. Then, initial estimator of parameters of the sojourn time distributions along with $\widehat{p}_{ij}^{0}$ are plugged into the conditional log-likelihood function $l(j_0)$ and Newton-Raphson method is used to obtain the maximum likelihood estimates.

2.2 Cross-state prediction

Once a Markov renewal model is fitted it is worth predicting a state (say, k) of the next event within t^* time period given the state corresponding to the last event is i and the time t_0 passed since the last event occurred. Under these assumptions the probability of the next event during $(t_0, t_0 + t^*]$ can be given by

$$\mathbb{P}(t^*, k \mid t_0, i) = \mathbb{P}(J_{n+1} = k, t_0 < X_{n+1} \le t_0 + t^* \mid J_n = i,\ X_{n+1} > t_0), \quad i,k = 1,\ldots,M, \quad (4)$$

where J_n is the state of the last event, J_{n+1} is the state of the next event, X_{n+1} is the time already passed by the last occurrence until the instant in which the prediction is made. For the given $F_{ij}(x)$, equation (4) takes the following form

$$\mathbb{P}(t^*, k \mid t_0, i) = \frac{[F_{ik}(t_0 + t^*) - F_{ik}(t_0)]\, p_{ik}}{\sum_{j=1}^{M} [1 - F_{ij}(t_0)] p_{ij}}, \quad (5)$$

which gives the probability that the next event of type k will be occurred after time t^* knowing that the last event was of type i and t_0 time has elapsed. In case of just one state, equation (5) becomes a renewal process, and takes the following form

$$\mathbb{P}(t^* | t_0) = \frac{[F(t_0 + t^*) - F(t_0)]}{[1 - F(t_0)]}.$$

2.3 Asymptotic behavior of an MRP: mean recurrence time

Parametric methods, in general, provide estimators with several attractive asymptotic properties. If a chain started from state j, the expected number of steps to return to the same state j for the first time is defined as the mean recurrence time for the state j. The mean return times for state j of a Markov renewal process denoted by ρ_j can be given as

$$\rho_j = \frac{1}{\pi_j} \sum_{i=1}^{M} \pi_i \eta_i, \tag{6}$$

where η_i, $i = 1, \ldots M$ is defined in equation (2), and $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_M)$ is the unique stationary distribution of the imbedded Markov chain $\{J_n, n \geq 0\}$.

3 Analysis of cyclone data

3.1 Data source and description

In this paper data on 125 tropical cyclones that hit the coast of Bangladesh over the period 1877–2009 are analyzed. The data are obtained from two separate sources, information on the first 115 cyclones that hit the coast during 1877–2003 are obtained from Global Tropical Cyclone Climatic Atlas (GTCCA) as reported in Islam and Peterson (2009), and the data on the last 10 cyclones of the most recent years 2003–2009 are obtained from the Centre for Research on the Epidemiology of Disasters (CRED). The sojourn times (inter-event) of the cyclones and their types are given in Table 1. The inter-event time for an event, for instance, No. 10 is 458 implies that the time interval between event No. 9 (which was a TD) and No. 10 (which was a TS) is 458 days. Tropical cyclones are classified into three categories depending on the wind speed in *knots*, that are *tropical depression* (TD), *tropical storm* (TS) and *hurricane* (HU). Out of 125 cyclones that hit the coast of Bangladesh over the period 1877–2009, 40 were classified as tropical depressions (TDs), 52 were tropical storms (TSs), and 33 were hurricanes (HUs) (Table 2). Beside the classification of cyclone, the information on the date of cyclone occurrence and the place of the landfall are also available in the data set.

We first fit a single Poisson model considering all three types of cyclone as a single type of events (e.g., Mooley, 1981), and then a marked Poisson model, i.e. three independent Poisson models for each type of cyclone (e.g., Gospodinov and Rotondi, 2001). These models are then compared with a Markov renewal model where we consider three type of cyclones and their sojourn times simultaneously.

Table 1 Inter-event times (in days) and type of tropical cyclones in Bangladesh during 1867–2009

No.	Inter-event time	Type	No.	Inter-event time	Type	No.	Inter-event time	Type
1	–	TD	46	92	TD	91	335	HU
2	2	TD	47	92	HU	92	212	TS
3	10	TD	48	303	TD	93	153	TS
4	89	TD	49	1	TD	94	335	TS
5	699	TD	50	608	HU	95	212	TS
6	578	TD	51	20	TS	96	1645	HU
7	1492	TS	52	11	TD	97	730	TS
8	153	TS	53	30	TD	98	547	TS
9	607	TD	54	31	TS	99	549	TS
10	458	TS	55	1035	TD	100	212	TS
11	608	TD	56	61	TD	101	488	TS
12	122	TS	57	2252	TD	102	31	HU
13	1096	HU	58	31	TS	103	760	TS
14	700	HU	59	731	TD	104	121	HU
15	91	HU	60	2007	TS	105	30	TS
16	1308	TD	61	153	TD	106	519	TS
17	1796	TD	62	273	TD	107	547	HU
18	153	TS	63	92	TD	108	579	HU
19	334	TS	64	358	HU	109	182	TS
20	273	TS	65	20	HU	110	153	TS
21	335	TS	66	189	TS	111	212	HU
22	92	TD	67	23	TS	112	123	HU
23	274	TS	68	141	TD	113	242	HU
24	548	HU	69	365	TS	114	184	HU
25	1339	TD	70	212	HU	115	700	TS
26	1035	TS	71	153	TD	116	168	HU
27	395	TD	72	366	TD	117	592	TD
28	304	TS	73	206	TS	118	401	HU
29	457	TD	74	17	HU	119	1003	HU
30	31	HU	75	142	TD	120	239	TS
31	335	TD	76	61	HU	121	184	HU
32	1003	HU	77	274	TS	122	159	TS
33	61	TD	78	91	TS	123	188	HU
34	336	TS	79	296	TS	124	174	HU
35	122	TD	80	12	TS	125	36	HU
36	638	TS	81	240	TD			
37	762	TD	82	365	TS			
38	304	TS	83	122	TS			
39	30	TD	84	212	TS			
40	365	TS	85	153	TS			
41	701	TD	86	31	HU			
42	153	TD	87	181	TS			
43	1673	TS	88	184	HU			
44	122	TS	89	731	TS			
45	212	TD	90	30	HU			

3.2 Poisson and marked Poisson models for the cyclone data

To consider a single sequence of events in continuous time the most common model is the homogeneous Poisson process $\{N(t), t \geq 0\}$ that assumes inter-event times are exponential random variables. The probability distribution of $N(t)$, the number of

Table 2 GTCCA classification of cyclones (Islam and Peterson, 2009) and number of land-falling cyclones occurred in Bangladesh during 1877–2009

Type	Category	Wind speed	Number of storms
Tropical depression	TD	< 34	40
Tropical storm	TS	$34 - 63$	52
Hurricane	HU	≥ 64	33
Total			125

events during the time interval $(0, t]$, can be defined as

$$\mathbb{P}(N(t) = n) = \begin{cases} \dfrac{(\lambda t)^n}{n!} \exp(-\lambda t), & n = 0, 1, \ldots \\ 0 & \text{otherwise,} \end{cases}$$

where λ is the mean number of events in unit time. Once the parameter λ is obtained, Poisson model is able to calculate the probability of occurrence of events ($n = 0, 1, \ldots$), i.e. the distribution of number of events for a time period t.

In this study we first fit a single Poisson process to the cyclone data obtained over the period 1877–2009 assuming all cyclones are of the same type. The estimated value of λ is 0.9398 per year (0.0783/month) and using the estimated value we can easily obtain the distribution of the number of cyclones. We plot the empirical cumulative distribution function (ECDF) and estimated cumulative distribution function (CDF) of the sojourn times to check the fit (Figure 1). The distribution of the number of cyclones ($N(t) = 0, 1, 2, 3, 4$) against time period is shown in Figure 2. Although the model seems reasonably accurate it gives only the probability distribution of number of cyclones without capturing their severity and their evolution. However, cyclones can be naturally categorized as tropical depressions (TD), tropical storms (TS) or hurricanes (HU) on the basis of their severity (mainly based on wind speed), and we are interested in predicting their types during a specified time period. Therefore, we fit a marked Poisson model for the Bangladesh cyclone data with the mark as the type of cyclones. We consider three different Poisson processes for three types of cyclones (TD, TS and HU), and the rates for the fitted processes are 0.0251, 0.0326 and 0.0207 per month for the types of cyclones TD, TS and HU, respectively. The fit of the sojourn time for each type of cyclones is shown in Figure 3. Using the rates of the processes probability distribution of occurrence of a TD, a TS and a HU against time is shown in Figure 4, which shows that although the probability of occurring a TD or a TS or a HU looks similar, a TS is more likely to occur in a shorter period than a TD or a HU, and a HU is more likely than a TS or a TD for a longer period.

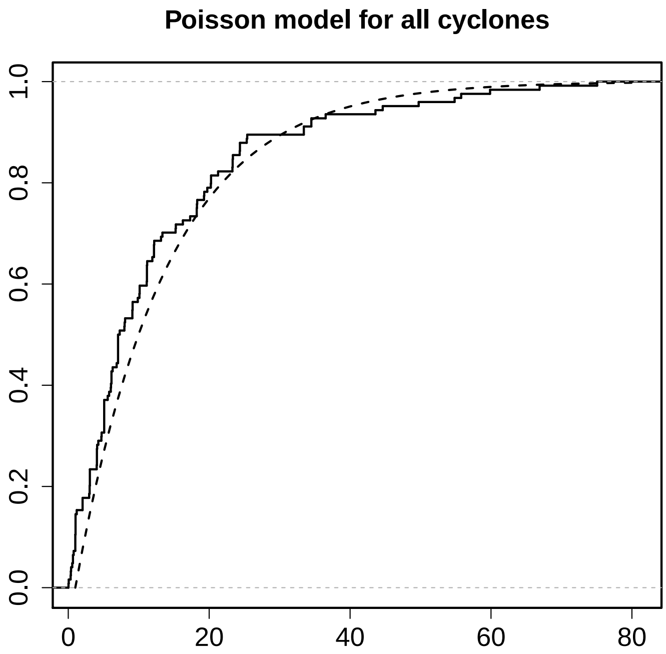

Fig. 1 Comparison between empirical (step-function) and estimated (dashed line) distributions of the sojourn time of cyclones

Although the Poisson and marked Poisson models give the probability distribution of the number of occurrences of cyclones, several issues restrict the use of these modelling approaches in cyclone prediction:

1. these models cannot capture the storm surge process TD, TS and HU simultaneously, i.e. they do not consider cross state transitions, and
2. these models are less interpretable as they are inherently memoryless. Therefore, to predict the probability of occurrence of next event they do not take into account the elapsed time after an event has occurred.

Hence, we consider a model that simultaneously captures the three types of cyclones, cross state transitions and the random amount of time the process spent in a particular state.

3.3 A Markov renewal model for the cyclone data

Poisson and marked Poisson models seem inadequate to model different categories of the storm surge process because these models are unable to capture different types of events simultaneously. It is, therefore, natural to consider a model which would be able to capture the three types of cyclones simultaneously. Hence, we propose a Markov

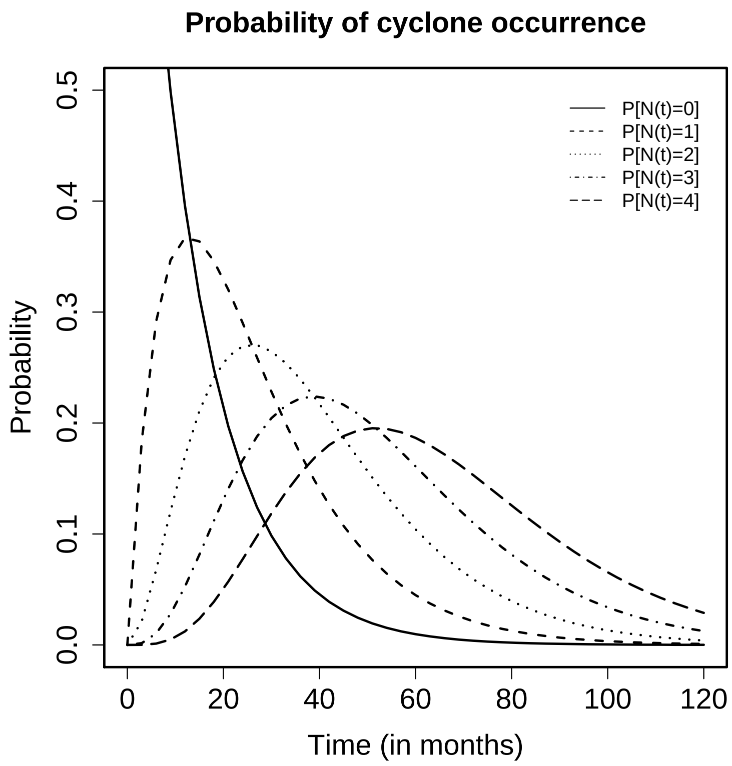

Fig. 2 Probability of occurrence of cyclones using a Poisson process

renewal model to consider the three categories of cyclones simultaneously having their random sojourn times. The following conjectures are assumed for fitting the Markov renewal process:

1. cyclones occur according to a renewal process,
2. type of cyclone is a discrete random variable, and constitutes a homogeneous Markov chain, and
3. the longer is the sojourn time for transition from one state to another the higher is the probability that the transition happens.

The model also requires the definition of the states visited by the process of events during its evolution. In our case three different states have been assumed according to the type of a cyclone, that are tropical depression (TD), tropical storm (TS), and hurricane (HU).

Fitting Markov renewal process requires specifying the distribution of sojourn times, which depends on the physics of the data under consideration. Weibull distribution is often used for modelling sojourn time (e.g., Alvarez, 2005; Garavaglia and Pavani, 2011; Masala, 2012a) because of its generalizability. The Weibull model for a

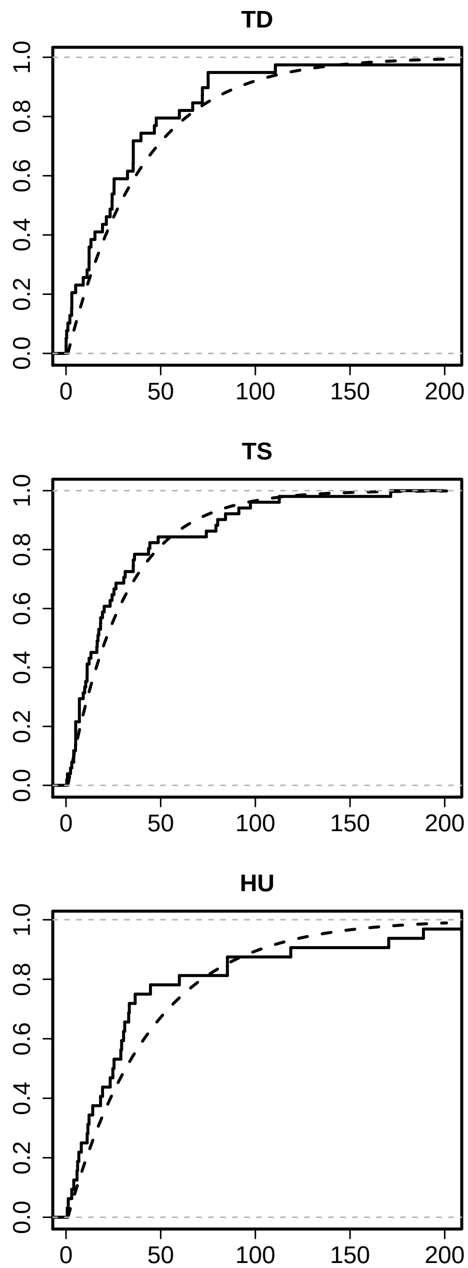

Fig. 3 Comparison between empirical (step-function) and estimated (dashed line) distributions of the sojourn times of TD, TS and HU separately

specific transition $i \to j$, $\forall\, i, j \in \{1, 2, 3\}$ ($1 \to$TD, $2 \to$TS, $3 \to$HU) takes the form

$$f_{ij}(x) = \frac{\alpha_{ij}}{\mu_{ij}} \left(\frac{x}{\mu_{ij}}\right)^{\alpha_{ij}-1} \exp\left[\left(-\frac{x}{\mu_{ij}}\right)^{\alpha_{ij}}\right], \quad \alpha_{ij}, \mu_{ij} > 0,$$

where α_{ij} and μ_{ij} are transition specific shape and scale parameters, respectively. Using this saturated Weibull distribution, where parameters corresponding to all possible

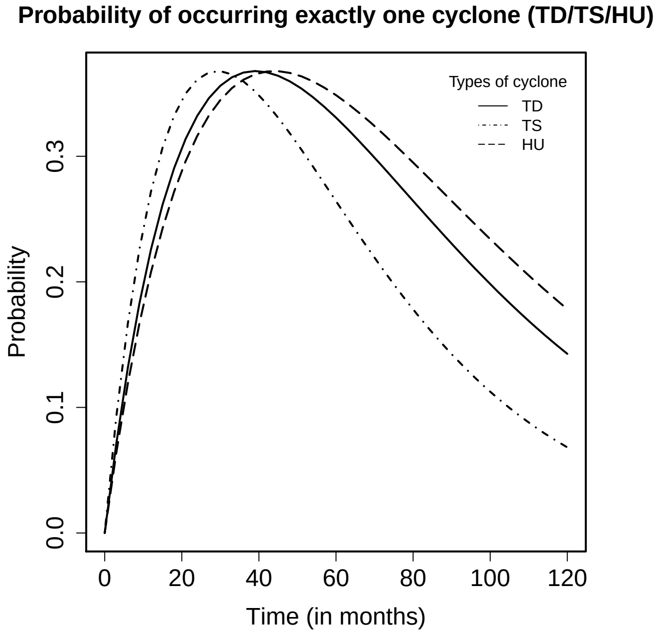

Fig. 4 Probability of occurrence of a TD, a TS and a HU using a marked Poisson process

transitions are considered, the log-likelihood (3) then becomes

$$l(j_0) = \sum_{i=0}^{\tau-1} \ln p_{j_i,j_{i+1}} - \sum_{i=0}^{\tau-1} \left\{ \ln \left[\frac{\alpha_{ij}}{\mu_{ij}} \left(\frac{x}{\mu_{ij}} \right)^{\alpha_{ij}-1} \right] - \left(\frac{x}{\mu_{ij}} \right)^{\alpha_{ij}} \right\} + \ln \left[\sum_{k=0}^{M} p_{j_\tau,k} \exp \left(- \frac{x_\tau}{\mu_{j_\tau k}} \right)^{\alpha_{j_\tau k}} \right]. \quad (7)$$

The log-likelihood function has two types of parameters, six parameters for the

Table 3 Tests for a sequence of nested models

Model	Description	p	$\log L$	p-value
Model I	Full Weibull	24	−561	–
Model II	$\alpha_{11} = \alpha_{23}$ and	18	−562	0.920
	$\alpha_{12} = \alpha_{13} = \alpha_{21} = \alpha_{3j}\ \forall j$			
Model III	$\alpha_{11} = \alpha_{23}$, $\alpha_{12} = \alpha_{13} = \alpha_{21} = \alpha_{3j}, \forall j$	11	−564	0.780
	$\mu_{12} = \mu_{31}$, $\mu_{11} = \mu_{13} = \mu_{2j} = \mu_{32} = \mu_{33},\ \forall j$			
Model IV	$\alpha_{ij} = \alpha_{kl}, \forall i, j, k, l$	9	−567	0.023
	$\mu_{12} = \mu_{31}$, $\mu_{11} = \mu_{13} = \mu_{2j} = \mu_{32} = \mu_{33}, \forall j$			

transition probability matrix (as the other three are determined by summing each row to one) and 18 parameters corresponding to nine transitions between the states {TD, TS, HU}. We have tried different models to reduce the number of parameters of the Weibull distribution, and four such important models are defined in Table 3.

The Model I is based on the full Weibull model with 24 parameters, Model II assumes three shape parameters and nine scale parameters, Model III assumes three shape and two scale parameters, and Model IV assumes one shape and two scale parameters. The models have nested structures and likelihood ratio test procedure is used to compare models, and the corresponding p-values are reported in Table 3. The analysis shows that the Model III is the best Weibull model for analyzing cyclone data, because there is no significant difference between models I, II and III, but Model III gives significantly better fit than the Model IV.

The estimated parameters of the best model Model III are reported in Table 4. The estimated transition probabilities show that the tropical storms (TS) are more likely to hit Bangladesh coast irrespective of the immediate last event.

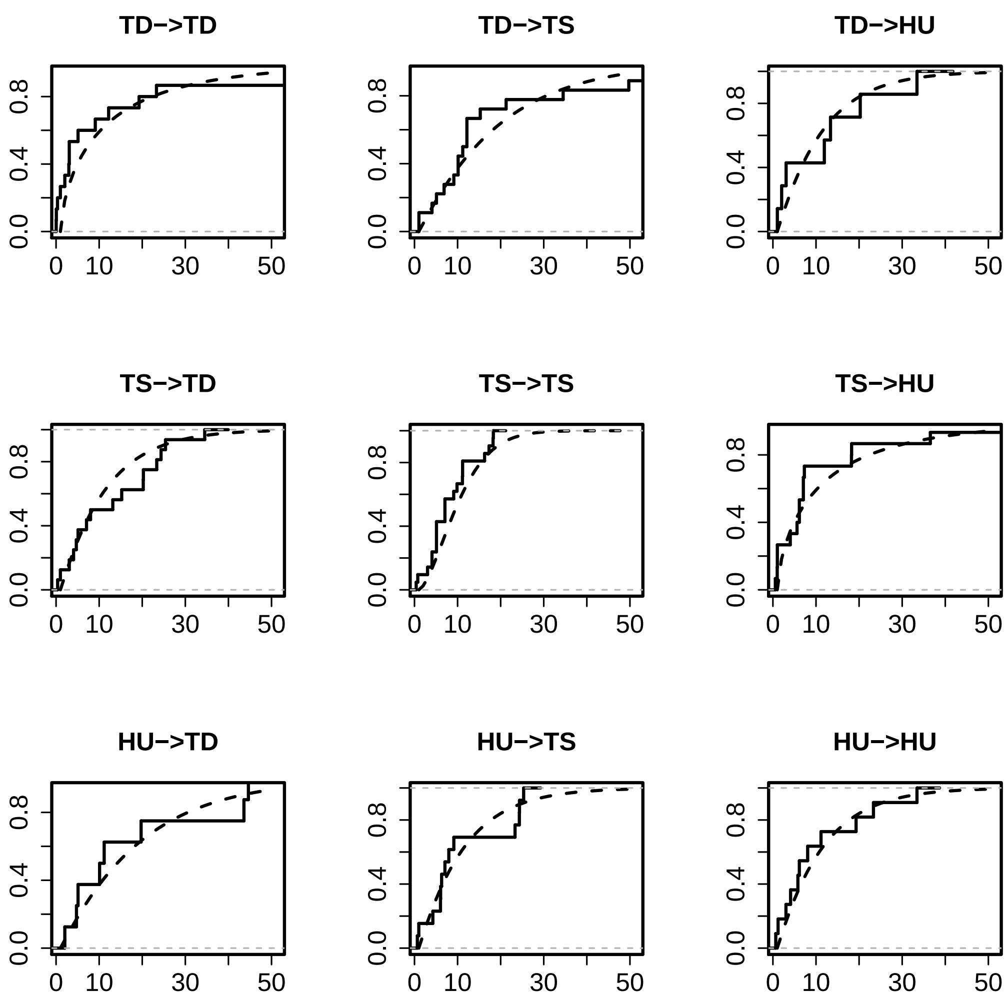

Fig. 5 Comparison between empirical (step function) and estimated (dashed line) distributions of the sojourn time of cyclone data

Table 4 Estimated transition probabilities, shape and scale parameters of the Weibull MRP for Model III.

	Shape ($\widehat{\boldsymbol{\alpha}}$)			Scale ($\widehat{\boldsymbol{\mu}}$)			Transition probs		
	TD	TS	HU	TD	TS	HU	TD	TS	HU
TD	$\widehat{\alpha}_{(1)}$	$\widehat{\alpha}_{(2)}$	$\widehat{\alpha}_{(2)}$	$\widehat{\mu}_{(1)}$	$\widehat{\mu}_{(2)}$	$\widehat{\mu}_{(1)}$	0.375	0.450	0.175
TS	$\widehat{\alpha}_{(2)}$	1.562	$\widehat{\alpha}_{(1)}$	$\widehat{\mu}_{(1)}$	$\widehat{\mu}_{(1)}$	$\widehat{\mu}_{(1)}$	0.308	0.404	0.289
HU	$\widehat{\alpha}_{(2)}$	$\widehat{\alpha}_{(2)}$	$\widehat{\alpha}_{(2)}$	$\widehat{\mu}_{(2)}$	$\widehat{\mu}_{(1)}$	$\widehat{\mu}_{(1)}$	0.249	0.407	0.344

$\widehat{\alpha}_{(1)} = 0.679$, $\widehat{\alpha}_{(2)} = 1.040$; $\widehat{\mu}_{(1)} = 10.593$, $\widehat{\mu}_{(2)} = 18.722$

To check the fits of the sojourn times the empirical CDF (step function) and the estimated CDF (dashed line) of Weibull distribution of sojourn times for each transition are plotted in Figure 5. As the fitted line is very close to the empirical line, it can be concluded that the Weibull distribution fits the data reasonably well.

Table 5 Probability of occurrence of next event of given a state, conditioned to a given state of last event occurred, evaluated for different t^* and t_0 years

		Predicted probability								
		Last event in state TD			Last event in state TS			Last event in state HU		
t_0	$(t_0, t_0 + t^*]$	TD→TD	TD→TS	TD→HU	TS→TD	TS→TS	TS→HU	HU→TD	HU→TS	HU→HU
0	(0,1]	0.2487	0.2103	0.1189	0.2091	0.2840	0.1913	0.1165	0.2764	0.2338
	(0,2]	0.3093	0.3267	0.1582	0.2781	0.3926	0.2380	0.1811	0.3675	0.3109
	(0,3]	0.3372	0.3875	0.1701	0.2990	0.4033	0.2594	0.2148	0.3951	0.3343
	(0,4]	0.3520	0.4186	0.1736	0.3052	0.4038	0.2708	0.2320	0.4033	0.3412
	(0,5]	0.3604	0.4343	0.1746	0.3070	0.4038	0.2773	0.2407	0.4057	0.3432
1	(0,1]	0.1800	0.2186	0.1224	0.2153	0.3661	0.1385	0.1211	0.2844	0.2406
	(0,2]	0.2628	0.3327	0.1596	0.2806	0.4022	0.2022	0.1844	0.3708	0.3137
	(0,3]	0.3066	0.3911	0.1706	0.2999	0.4038	0.2359	0.2167	0.3963	0.3353
	(0,4]	0.3317	0.4206	0.1737	0.3055	0.4038	0.2552	0.2331	0.4037	0.3415
	(0,5]	0.3467	0.4354	0.1746	0.3071	0.4038	0.2668	0.2413	0.4058	0.3433
2	(0,1]	0.1591	0.2218	0.1237	0.2176	0.3869	0.1224	0.1229	0.2875	0.2432
	(0,2]	0.2434	0.3354	0.1602	0.2817	0.4034	0.1873	0.1859	0.3723	0.315
	(0,3]	0.2916	0.3928	0.1708	0.3003	0.4038	0.2244	0.2177	0.3968	0.3357
	(0,4]	0.3207	0.4216	0.1738	0.3056	0.4038	0.2467	0.2337	0.4039	0.3417
	(0,5]	0.3388	0.4360	0.1747	0.3071	0.4038	0.2606	0.2416	0.4058	0.3434

Table 6 Probability of occurrence of next event of given a state, conditioned to a given state of last event occurred (HU), evaluated for different t^* with $t_0 = 3.25$ years

$(t_0, t_0 + t^*]$	Predicted probability		
(in years)	HU→TD	HU→TS	HU→HU
$(0, 1]$	0.1243	0.2899	0.2453
$(0, 2]$	0.1871	0.3735	0.316
$(0, 3]$	0.2185	0.3973	0.3361
$(0, 4]$	0.2342	0.4040	0.3418
$(0, 5]$	0.2419	0.4059	0.3434

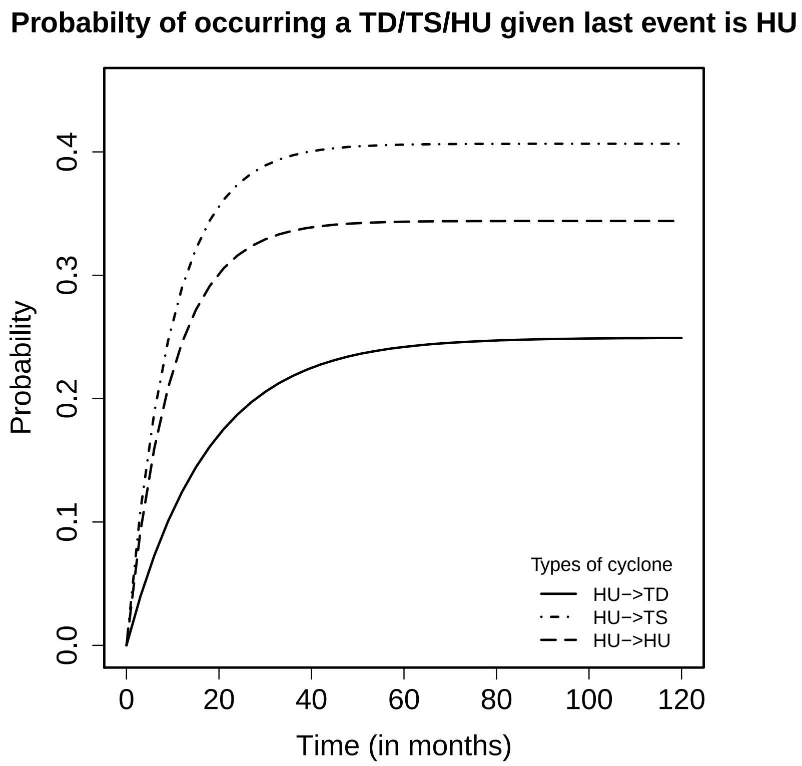

Fig. 6 Probability of occurring a TD, a TS and a HU given that last event is HU

Table 7 Recurrence periods for each type of cyclone

Types of cyclone	Marked Poisson process Average recurrence period (in year)	Markov renewal process Average recurrence period (in year)
TD	3.33	3.39
TS	2.56	2.54
HU	4.03	3.97

Average recurrence period of a cyclone (TD/TS/HU) using the Poisson process is 1.06 years

The probability of occurrence of a particular type of cyclone (TD/TS/HU) if last occurred event is a TD, TS or HU, is evaluated for varying elapsed time and time ahead to go to a particular state. The equation (5) gives the predicted probability of the next event TD, TS and HU, knowing that the last cyclone occurred is an event with any of the state TD, TS or HU for different values of t_0 and t^*, and the results are shown in Table 5. The predicted probability that within one year the transition concerns distinct cyclones TD→TD, TS→TS, HU→TS will occur are 0.2487, 0.2840 and 0.2764, respectively. This implies that immediately after the last event the chance of occurrence of a TS is the highest within the next 12 months period. However, as

t^* increases ($t^* = 2, 3, 4, 5$ year) the likelihood of TD→TS transition increases. If no time is elapsed the probability of occurrence of a TS is the highest for $t^* = 2, 3, 4, 5$. With $t_0 = 1$ year the likelihood of occurrence of cyclone increases in the same way with probabilities $0.2186, 0.3327, 0.3911, 0.4206, 0.4354$ for $t^* = 1, 2, 3, 4, 5$ given that the last state is a TS. The trend of predicted probabilities remains the same even if the last state is a HU. However, the probability of occurrence of a TD is higher than TS and HU if the last state is a TS when $t_0 = 1$ year and $t^* = 2, 3, 4, 5$ years. The probability of occurrence of events are similar for $t_0 = 1$ year and for $t_0 = 2$ year. As the occurrence of hurricanes (HU) is of interest the likelihood of occurrence of a HU is higher when the last state is either a TS or a HU.

Our model includes the last event occurred in the year 2009 that reached a hurricane level cyclone (HU), therefore, it may be interesting to see the prediction probability for the next event. Table 6 gives the predicted probability that the next cyclone will be a TD, TS or HU after $t^* = 1, 2, 3, 4, 5$ years given that about a period of 3 years and a quarter has elapsed ($t_0 = 3.25$). Within next one year period the probability of occurring a TS is the highest 0.2899 given that the last event occurred is a hurricane (HU). The corresponding probabilities of occurring a TD or a HU are 0.1243 and 0.2453. The chances of occurrence of a TD, a TS and a HU increases consistently as t^* increases (Figure 6). Hence, based on the estimated probabilities we see that the chance of the next cyclone to be a TS is the highest.

One major distinguishing feature of the counting process type model is that the mean recurrence time can be obtained for such processes. In storm risk analysis, the mean recurrence times plays an important role in terms of the probabilistic hazard assessment. The mean recurrence time of a Poisson and marked Poisson processes can be obtained by inverting the rate of occurrence of an event ($1/\lambda$), and for a Markov renewal process it is given by the equation (6). Estimated mean recurrence times using Poisson, marked Poisson and Markov renewal models are presented in Table 7. The mean recurrence time estimated by the Markov renewal model is slightly higher for TD than by the marked Poisson model, and lower for TS and HU. However, the results show that the average recurrence period for a TD is the lowest, and is two-year and a half, for a TD is three-year and a half, and for a HU is fours years.

4 Concluding remarks

Poisson, renewal and semi-Markov models have become a significant tool to model natural hazard processes as they are able to predict the probability of future events. In this study a stationary Weibull Markov renewal model for predicting tropical cyclone occurrences for Bangladesh is presented. The model formulation and parameter estimation procedure were validated using a data set from literature referring to occurrence of tropical cyclones in Bangladesh, which allowed us to make interesting comparisons. It has been found that the reduced Weibull distribution fits well for the distribution of sojourn times of cyclones. The predicted probabilities based on the model results have been presented, and a detail comparison with a Poisson process and a marked Poisson process was given. The model shows that the next event will be a tropical depression with a higher probability than a tropical storm or a hurricane. As no notable seasonality was observed in the occurrence of cyclones in Bangladesh we fitted a stationary renewal process. However, cyclones in Bangladesh tend to occur in May-June and October-November periods, and a slight seasonality observed in the data set could have been incorporated. In future this issue will be the focus of further research.

Acknowledgements The authors would like to thank the Referees for their constructive comments and helpful suggestions for the improvement of this paper.

References

Ali A (1996) Vulnerability of Bangladesh to climate change and sea level rise through tropical cyclones and storm surges. Water, Air, & Soil Pollution 92(1):171–179

Alvarez E (2005) Estimation in stationary Markov renewal processes, with application to earthquake forecasting in Turkey. Methodology and Computing in Applied Probability 7(1):119–130

Asaduzzaman M, Latif AHMM (2014) Computational intelligence techniques in earth and environmental sciences, Springer, chap 7

Dube S, Sinha P, Roy G (1985) The numerical simulation of storm surges along the Bangladesh coast. Dynamics of Atmospheres and Oceans 9(2):121–133

Dube S, Sinha P, Roy G (1986) Numerical simulation of storm surges in Bangladesh using a bay-river coupled model. Coastal Engineering 10(1):85–101

Fiorentino M, Versace P, Rossi F (1984) Two component extreme value distribution for flood frequency analysis. Water Resource Research 20(7)

Garavaglia E, Pavani R (2011) About earthquake forecasting by Markov renewal processes. Methodology and Computing in Applied Probability 13(1):155–169

Gospodinov D, Rotondi R (2001) Exploratory analysis of marked Poisson processes applied to Balkan earthquake sequences. Journal of the Balkan Geophysical Society 4(3):61–68

Gregory J, Wigley T, Jones P (1993) Application of Markov models to area-average daily precipitation series and interannual variability in seasonal totals. Climate Dynamics 8(6):299–310

Gupta V, Duckstein L (1975) A stochastic analysis of extreme droughts. Water Resources Research 11(2):221–228

Haque C (1995) Climatic hazards warning process in Bangladesh: experience of, and lessons from, the 1991 April cyclone. Environmental Management 19(5):719–734

Haque C, Blair D (1992) Vulnerability to tropical cyclones: evidence from the April 1991 cyclone in coastal Bangladesh. Disasters 16(3):217–229

Islam T, Peterson R (2009) Climatology of landfalling tropical cyclones in Bangladesh 1877-2003. Natural Hazards 48(1):115–135

Jagger T, Niu X, Elsner J (2002) A space–time model for seasonal hurricane prediction. International Journal of Climatology 22(4):451–465

Janssen J, Manca R (2007) Semi-Markov Risk Models for Finance, Insurance and Reliability. Springer-Verlag

Khalil G (1992) Cyclones and storm surges in Bangladesh: some mitigative measures. Natural Hazards 6(1):11–24

Lardet P, Obled C (1994) Real-time flood forecasting using a stochastic rainfall generator. Journal of Hydrology 162(3-4):391–408

Limnios N, Oprian G (2001) Semi-Markov processes and reliability. Springer

Lu Y, Garrido J (2005) Doubly periodic non-homogeneous Poisson models for hurricane data. Statistical Methodology 2(1):17–35

Masala G (2012a) Earthquakes occurrences estimation through a parametric semi-Markov approach. Journal of Applied Statistics 39(1):81–96

Masala G (2012b) Hurricane lifespan modeling through a semi-Markov parametric approach. Journal of Forecasting DOI 10.1002/for.2245

Mooley D (1981) Applicability of the Poisson probability model to the severe cyclonic storms striking the coast around the Bay of Bengal. Sankhyā: The Indian Journal of Statistics, Series B 43(2):187–197

Ogata Y (1988) Statistical models for earthquake occurrences and residual analysis for point processes. Journal of the American Statistical Association 83(401):9–27

Ogata Y (1998) Space-time point-process models for earthquake occurrences. Annals of the Institute of Statistical Mathematics 50(2):379–402

Pyke R (1961) Markov renewal processes: definitions and preliminary properties. The Annals of Mathematical Statistics 32(4):1231–1242

Rumpf J, Weindl H, Höppe P, Rauch E, Schmidt V (2007) Stochastic modelling of tropical cyclone tracks. Mathematical Methods of Operations Research 66(3):475–490

Sinha P, Dube S, Roy G, Jaggi S (1986) Numerical simulation of storm surges in Bangladesh using a multi-level model. International Journal for Numerical Methods in Fluids 6(5):305–311

Votsi I, Limnios N, Tsaklidis G, Papadimitriou E (2012) Estimation of the expected number of earthquake occurrences based on semi-Markov models. Methodology and Computing in Applied Probability 14(3):685–703