An Approach for Modeling Inter-arrival Time of Floods

P. Stoynov, P. Zlateva, and D. Velev

***Abstract*—The paper proposes an approach for modeling inter-arrival time of floods. It is shown that the inter-arrival time between floods can be modeled by a newly proposed type of distribution called ST distributions. The corresponding stochastic point process registering the appearances of the floods is called ST process. ST distribution and process are defined and their application to modeling flood appearances is demonstrated. The proposed approach is applied for modeling inter-arrival time between two consecutive floods based on a specific type of stochastic distributions and processes.**

***Index Terms*—Natural disasters risk, flood, ST distribution, ST process.**

I. INTRODUCTION

United Nations International Strategy for Disaster Reduction defines risk of natural disaster as "a potentially damaging phenomenon that may lead to loss of life or injury, property damage, social and economic disruption or environmental degradation". Each hazard is characterized by location, intensity, frequency and probability. It is interesting to study inter-arrival time between two disasters in a vulnerable geographic area [1].

Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the assessment and management of flood risks (Floods Directive – FD) provides for drafting of Flood risk management plans (FRMPs) for each river basin management region in EU member countries. [2]. Floods have the potential to cause fatalities, displacement of people and damage to the environment, to severely compromise economic development and to undermine the economic activities of the Community.

Floods are natural phenomena which cannot be prevented. However, some human activities and climate change contribute to an increase in the likelihood and adverse impacts of flood events. It is feasible and desirable to reduce the risk of adverse consequences, especially for human health and life, the environment, cultural heritage, economic activity and infrastructure associated with floods [3].

It is interesting to study inter-arrival time between two disasters in a vulnerable geographic area.

Statistical data point out that floods are the most frequent natural disasters with 40% of all disasters. It is important to estimate the inter-arrival time between two consecutive floods.

The aim of this paper is to propose an approach for modeling inter-arrival time of floods. It is shown that the inter-arrival time between floods can be modeled by a newly proposed type of distribution called ST distributions [4]-[6].

The corresponding stochastic point process registering the appearances of the floods is called ST process. ST distribution and process are defined and their application to modeling flood appearances is demonstrated. The proposed approach is applied for modeling inter-arrival time between two consecutive floods based on a specific type of stochastic distributions and processes.

II. DEFINITION OF ST FAMILY OF DISTRIBUTIONS

In this study, a random variable ξ with probability mass function $f_\xi(x)$ has distribution of $ST(n,\beta)$ family and this fact is denoted $\xi \in ST(n,\beta)$, if the probability mass function of ξ is given by the formula:

$$f_\xi(x) = \begin{cases} \sum_{k=1}^{n+1} P(D^n = k)\, f_\xi(x \mid D^n = k) = \\ = \sum_{k=1}^{n+1} P(D^n = k)\, f_{G^k}(x), x \geq 0 \\ 0, x < 0. \end{cases},$$

where G^k are random variables with probability mass function $f_{G^k}(x) = f(k,\beta)$ and D^n are positive integer mixing random variables.

Here for G^k different families of distributions are adopted.

The case of ST family of first kind is considered when

$$G^k \in \Gamma(k, \frac{1}{\beta}) \equiv Erlang(k, \frac{1}{\beta}),$$

i.e.

$$\xi \mid D^n \equiv Erlang(D^n, \frac{1}{\beta}).$$

Correspondingly, D^n is a random variable, taking values $k = 1, \ldots, (n+1)$ with probabilities:

Manuscript received May 24, 2015; revised July 28, 2015. This work was supported in part by the Bulgarian National Science Fund for the support under the Grant No. DFNI-I02/ 15 from 2014.

P. Stoynov is with the Sofia University St. Kl. Ohridski, Bulgaria (e-mail: todorov@feb.uni-sofia.bg).

P. Zlateva is with the ISER, Bulgarian Academy of Sciences, Sofia 1113, Bl. 2, Bulgaria (e-mail: plamzlateva@abv.bg).

D. Velev is with the University of National and World Economy, Sofia 1700, Bulgaria (e-mail: dgvelev@unwe.bg).

DOI: 10.7763/IJIMT.2015.V6.613

$$P(D^n = k) = \frac{C(n,\beta)n!}{\beta^k (n-k+1)!}, k = 1,\ldots,(n+1),$$

where the coefficients $C(n,\beta)$ are given by the formulas:

$$C(n,\beta) = \frac{1}{I(n,\beta)},$$

$$I(0,\beta) = \frac{1}{\beta},$$

$$I(n,\beta) = \frac{1}{\beta} + \frac{n}{\beta} I(n-1,\beta), n = 1,2,\ldots$$

Also, variables $\widetilde{D}^n = D^n - 1$ can be introduced taking values $k = 0,\ldots,n$ with probabilities

$$P(\widetilde{D}^n = k) = \frac{C(n,\beta)n!}{\beta^{k+1}(n-k)!}, k = 0,\ldots,n.$$

Then the probability mass function $f_\xi(x)$ of ξ can be presented also by the formula

$$f_\xi(x) = \begin{cases} \sum_{k=0}^{n} P(\widetilde{D}^n = k) f_\xi(x \mid \widetilde{D}^n = k) = \\ = \sum_{k=0}^{n} P(\widetilde{D}^n = k) f_{G^{k+1}}(x), x \ge 0 \\ 0, x < 0. \end{cases}$$

If a random variable ξ has ST distribution of first kind, this fact is denoted as $\xi \in ST1(n,\beta)$.

In the case when $D^n = k$, i.e.

$$P(D^n = k) = 1, P(D^n = i) = 0,$$
$$1 \le i \le k-1 < k+1 \le i \le n+1,$$

which may be considered as degenerate ST distribution of first kind, Erlang distribution is actually obtained, i.e.

$$\xi \in Erlang(k, \frac{1}{\beta}).$$

The following theorem holds:

Theorem 1. Let $\xi \in ST1(n,\beta)$. Then its probability mass function can be presented as

$$f_\xi(x) = \begin{cases} C(n,\beta)e^{-\beta x}(1+x)^n, x \ge 0 \\ 0, x < 0 \end{cases}.$$

Proof: The probability density can be represented as:

$$C(n,\beta)e^{-\beta x}(1+x)^n = C(n,\beta)e^{-\beta x} \sum_{k=0}^{n} \binom{n}{k} x^k =$$

$$= \sum_{k=0}^{n} C(n,\beta) \binom{n}{k} x^k e^{-\beta x} =$$

$$= \sum_{k=0}^{n} \frac{C(n,\beta)n! x^k e^{-\beta x}}{k!(n-k)!} = \sum_{k=0}^{n} \frac{C(n,\beta)n! \beta^{k+1} x^k e^{-\beta x}}{\beta^{k+1}(n-k)!k!} =$$

$$= \sum_{k=0}^{n} \frac{C(n,\beta)n!}{\beta^{k+1}(n-k)!} \frac{\beta^{k+1} x^k e^{-\beta x}}{k!} =$$

$$= \sum_{k=0}^{n} P(D^n = k+1) f_{G^{k+1}}(x) =$$

$$= \sum_{k=1}^{n+1} \frac{C(n,\beta)n! \beta^k x^{k-1} e^{-\beta x}}{\beta^k (n-k+1)!(k-1)!} =$$

$$= \sum_{k=1}^{n+1} \frac{C(n,\beta)n!}{\beta^k (n-k+1)!} \frac{\beta^k x^{k-1} e^{-\beta x}}{\Gamma(k)} =$$

$$= \sum_{k=1}^{n+1} \frac{C(n,\beta)n!}{\beta^k (n-k+1)!} \frac{\beta^k x^{k-1} e^{-\beta x}}{(k-1)!} =$$

$$\sum_{k=1}^{n+1} P(D^n = k) f_{G^k}(x).$$

So, the proof is completed.

The distribution $ST1(n,\beta)$ can be considered as a special kind of generalized gamma distribution [4]. It is said that a random variable Λ has generalized gamma distribution if its probability density function is given by

$$f_\Lambda(x) = \frac{u^{s-\alpha}}{\Gamma(\alpha)U(\alpha, \alpha+1-s, u\beta)} \times$$
$$\times e^{-x\beta} x^{\alpha-1}(u+x)^{-s}, x > 0,$$

where $-\infty < s < +\infty, \alpha > 0, \beta > 0, u > 0$ and

$$U(a,b,z) = \frac{1}{\Gamma(a)} \int_0^\infty e^{-zt} t^{a-1}(1+t)^{b-a-1} dt,$$
$$a > 0, z > 0$$

is the integral representation of the hyper-geometric function of second kind [7].

In this case, for Λ can be written $\Lambda \in G\Gamma(\alpha,\beta,u,s)$.

It is necessary to remember that the random variable ξ has a gamma distribution and denote $\xi \in \Gamma(\alpha,\beta)$, if its probability mass function is given by

$$f_\xi(x) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, x \ge 0 \\ 0, x < 0 \end{cases}.$$

Here $\Gamma(\alpha)$ is defined by

$$\Gamma(\alpha)=\int_0^\infty x^{\alpha-1}e^{-x}dx.$$

So, $\Gamma(\alpha,\beta)=G\Gamma(\alpha,\beta,u,s=0).$

Actually, it is obtained that

$$ST1(n,\beta)=G\Gamma(\alpha=1,\beta,u=1,s=-n).$$

Another generalization of $ST1(n,\beta)$ distribution is presented in [5].

The exponential distribution is a special kind of $ST1(n,\beta)$ distribution and $ST1(0,\beta)\equiv Exp(\beta)$.

To recall that the random variable ξ has exponential distribution and this is denoted $\xi\in Exp(\beta)$, if its probability mass function is given by

$$f_\xi(x)=\begin{cases}\beta e^{-\beta x},\ x\ge 0\\ 0,\ x<0.\end{cases}$$

The case $ST1(1,\beta)\equiv Lindley(\beta)$ can be considered as a weighted version of exponential distribution with weighting function $w(x)=1+x.$ The random variable ξ has Lindley distribution and this is denoted $\xi\in Lin(\beta),$ if its probability mass function is given by

$$f_\xi(x)=\begin{cases}\dfrac{\beta^2}{\beta+1}(1+x)e^{-\beta x},\ x\ge 0\\ 0,\ x<0\end{cases}.$$

The case $n=2$ leads to $ST1(2,\beta)$ distribution which can be defined as weighted exponential distribution by the weight function $w(x)=(1+x)^2.$

The random variable ξ has $ST1(2,\beta)$ distribution and this fact is denoted $\xi\in ST1(2,\beta),$ if its probability mass function is given by

$$f_\xi(x)=\begin{cases}\dfrac{\beta^3}{\beta^2+2\beta+2}e^{-\beta x}(1+x)^2,\ x\ge 0\\ 0,\ x<0.\end{cases}$$

To check that the weight function in the definition of the distribution is the right one, it can be shown that

$$Ew(\xi)=\frac{\beta^2+2\beta+2}{\beta^2}.$$

III. Simulation of ST Distribution

The simulation of $ST1(n,\beta)$ distribution and corresponding PMF graphics for different values of the parameters n and β can be done by using R language.

The graphics for the case $n=100$ and $\beta=2$ is given in Fig. 1. The graphics for the case $n=40$ and $\beta=2$ is given in Fig. 2.

IV. Definition of ST Processes

The process $X(t)$ is a ST process of first kind or $ST1(n,\beta)$ process, and this fact is denoted as $X(t)\in ST1(t;n,\beta),$ if for it:

1) $X(0)=0.$

2) $X(t)$ is pure jump process with jumps at times $T_{i'}, i=1,2,\dots$ and jump sizes

$\Delta X(T_i)=1.$

3) The intervals between two jumps are

$\tau_i=T_i-T_{i-1}\in ST1(n-1,\beta), i=0,1,\dots,T_0=0.$

The process $X(t)$ is compound $ST1(n,\beta)$ process if condition 3) is replaced by condition:

3') $X(t)$ is pure jump process with jumps at times $T_i, i=1,2,\dots$ and jump sizes $\Delta X(T_i)=Y_i,$ where the variables $Y_i, i=1,2,\dots$ are independent and identically distributed random variables.

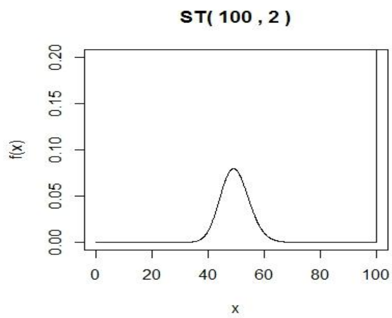

Fig. 1. The graphics of $ST1(n,\beta)$ distribution for the case $n=100$ and $\beta=2$.

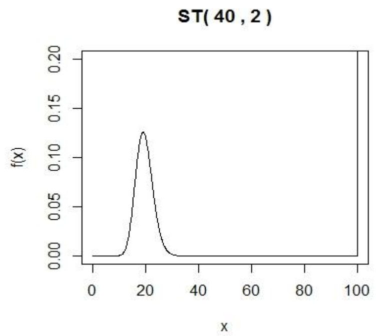

Fig. 2. The graphics of $ST1(n,\beta)$ distribution for the case $n=40$ and $\beta=2$.

V. Simulation of ST Processes

To make simulation of $ST(n,\beta)$ process, the following algorithm can be applied:

1. Define interval $[0,T]$ of the simulation.

2. Set $k = 0$.

3. While $\sum_{i=1}^{k} \tau_i < T$ do:

3.1. Set $k = k + 1$

3.2. Generate $\tau_k \in ST1(n-1, \beta)$
$(\tau_k \in ST2(n-1, \beta), \tau_k \in ST3(n-1, \beta))$.

3.3. Set $Y_k = 1$ for standard $ST(n, \beta)$ process or simulate Y based on a given distribution f for compound $ST(n, \beta)$ process.

Then the trajectory of $X(t) \in ST(t; n, \beta)$ is given by the formula

$$X(t) = \sum_{i=1}^{N(t)} Y_i,$$

where

$$N(t) = \sum_i 1_{\{T_i < T\}}.$$

This kind of processes are considered and simulated in [6] where also the R code for simulation is presented.

The graphics of a trajectory of $ST1(n, \beta)$ process for the case $n = 6$ and $\beta = 0.1$ is given in Fig. 3.

VI. A Proposal for Modeling Floods Arrivals by ST Processes

Floods are one of the biggest natural hazards in terms of fatalities. Probably the biggest natural disaster ever was the Henan Flood in China in 1887 with between 900 000 and 1.5 million victims [8]. The flood in summer 2002 in Central Europe was probably the costliest European natural disaster over. First estimation ranged from 20 to 100 billion euro.

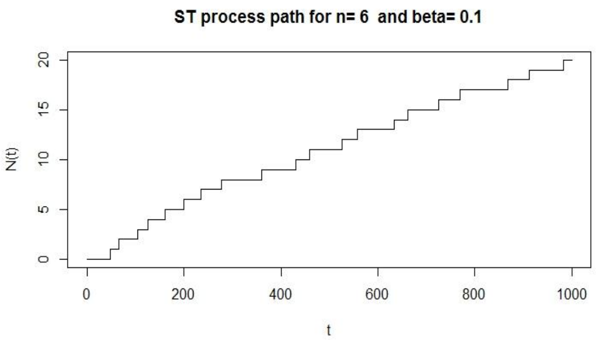

Fig. 3. The graphics of $ST1(n, \beta)$ distribution for the case $n = 6$ and $\beta = 0.1$.

Historical data of floods in Europe have shown that floods are correlated. In fact, so called "time clusters" could be found. There exist several theories to describe such effects, for example climate cycles or the fatigue of the water storage capacity of a landscape after a flood. Therefore, the next flood can occur even with a lower water supply.

The floods are two types.

The first type is "surface waves". They are sea floods caused by storm and tides, or river floods caused by the heavy release of water either through snow or ice melting or heavy rain. Another type of sea surface waves is "freak waves". They are explained by spatial and temporal interference, and are defined as being at least three times the size of an average storm wave.

In contrast to surface waves, tsunamis are deep-sea waves caused by under-sea earthquakes. They have a great wave length and wave speed, and usually are not noticeable at sea. However, at coastal regions, they can cause great damage.

In this section, it is demonstrated how ST processes can be used to model the point process of occurrence of surface wave floods.

It is supposed that there are two types of events related to floods. The first type is "condition for floods". The process which counts number of appearances of this kind of events is Poisson with parameter β. The second kind of events are called "real flood". It can be modeled by process $ST1(n = [\frac{t_{\max}}{t_{\min}}], \beta)$ where $t_{\max}$ is the maximal and $t_{\min}$ - the minimal observed interval between two major floods in a country (area, region) and $[.]$ is the integer part of a number. This approach can be illustrated with data about floods caused by river Elbe presented in the below table.

So, it can be estimated that $\frac{1}{\beta} = 15$.

Also, $n = [\frac{t_{\max}}{t_{\min}}] = \frac{154}{14} = 10$ and the realizations of the random variables D_i^n are as shown in Table I.

TABLE I: Historical Maximum Water Levels of the River Elbe at Dresden [9]

Date	Water level (m)	Volume (m3/s)	Time interval (years)	D_i^n
16 August 1501	8,57	5,0	154	10
7 February 1655	8,38	4,8	129	9
1 March 1784	8,57	5,2	15	1
24 February 1799	8,24	4,4	31	2
2 March 1830	7,96	3,95	15	1
31 March 1845	8,77	5,70	17	1
3 February 1862	8,24	4,493	14	1
20 February 1876	7,76	3,286	14	1
7 September 1890	8,37	4,35	50	3
17 March 1940	7,78	3,36	62	4
17 August 2002	9,4	4,7	-	

VII. Conclusion

ST distributions presented in the article posses some suitable properties for modeling process of arrival of floods.

The present work may be extended by studying other choices of G^k.

For example, it is possible to choose $G^k \in NB(k, e^{-\beta})$. In this case, $\xi \mid D^n \equiv NB(D^n, e^{-\beta})$ and $ST2(n,\beta)$ distribution is obtained.

As another example, the random variable $G^k \equiv \delta_k(x)$ can be considered which takes value k with probability one. Then $\xi \equiv D^n \in D^n(\beta)$. In this case, that random variable ξ has $ST3(n,\beta)$ distribution and this fact is denoted $\xi \in ST3(n,\beta)$.

These processes with suitable parameters can also be used to model times of occurrence of floods.

The model can also be used to model arrivals of other natural hazards.

Acknowledgment

The authors express their gratitude to the Bulgarian National Science Fund for financial support under the Grant No. DFNI-I02/ 15 from 2014, titled "Information System for Integrated Risk Assessment from Natural Hazards".

References

[1] United Nations International Strategy for Disaster Reduction. (2009). [Online]. Available: http://www.unisdr.org/files/ 7817_UNISDRTerminologyEnglish.pdf

[2] A European Flood Action Programme. European Commission. [Online]. Available: http://ec.europa.eu/environment/water/flood_risk/key_docs.htm

[3] Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the assessment and management of flood risks. [Online]. Available: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32 007L0060&from=EN

[4] P. Stoynov, "Mixed negative binomial distribution by weighted gamma mixing distribution," in *Proc. the Fortieth Spring conference of the Union of Bulgarian Mathematicians*, Borovets, Bulgaria, 5-9 April, 2011, pp. 327-331.

[5] P. Stoynov, "Some properties of a special class generalized Stoynov distributions," in *Proc. the International Conference Financial and Actuarial Mathematics*, Sofia, Bulgaria, 2012, pp. 102-106.

[6] P. Stoynov, "Stoynov process – Definition and simulations," in *Proc. the International Conference Financial and Actuarial Mathematics*, Sofia, Bulgaria, 2013, pp. 97-104.

[7] H. Zamani and N. Ismail, "Negative binomial-Lindley distribution and its applications," *Journal of Mathematics and Statistics*, vol. 6, no. 1, pp. 4-9, 2010.

[8] D. Proske, *Catalogue of Risks, Natural, Technical, Social and Health Risks*, Springer, 2008.

[9] D. Fischer, "Jahrhundert-Hochchwasser oder drastische Klimaveranderung," in *Als den Lowen das Wasser bis zum Rachen stand. Elbhang-Photo-Galerie*, F. D. Dresden and J. Frohse Eds., 2003, pp. 71-75.

Pavel Stoynov has graduated at the Sofia University St. Kl. Ohridski, Sofia, Bulgaria. Currently he is an assistant professor at the Faculty of Business and Administration of Sofia University St. Kl. Ohridski. His research interests are in the field of mathematics, economics, geosciences, information technologies and statistics. He is a member of AMS and Bulgarian Statistical Society.

Plamena Zlateva is currently an associate professor at the Institute of System Engineering and Robotics at the Bulgarian Academy of Sciences, Sofia, Bulgaria. She holds the M.Sc. degrees in applied mathematics from the Sofia Technical University and in economics from the Sofia University St. Kl. Ohridski, and the Ph.D. degree in manufacturing automation from the Institute of Conrol and System Research - BAS.

Her main areas of academic and research interest are control theory, mathematical modeling and system identification, risk management.

Dimiter Velev is a professor in the Department of Information Technologies and Communications at the University of National and World Economy, Sofia, Bulgaria. He holds the M.Sc. degree in electroengineering from the Sofia Technical University, Bulgaria and the Ph.D. degree in engineering sciences from the Institute of Modeling Problems in power engineering at the National Academy of Sciences of Ukraine, Kiev, Ukraine. His main areas of academic and research interest are internet-based business systems modeling and development, service oriented architectures, online social networks, cloud computing, web applications development and programming. His lectures cover such disciplines.