Proceedings of the 33rd European Safety and Reliability Conference (ESREL 2023)
Edited by Mário P. Brito, Terje Aven, Piero Baraldi, Marko Čepin and Enrico Zio

Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-8071-1_P546-cd

Return Periods of Extreme Events in the Changing Climate: LEYP Model

M.D. Pandey

Department of Civil and Environmental Engineering, University of Waterloo, Waterloo, Canada.
E-mail: mdpandey@uwaterloo.ca

S. Mercier

Universite de Pau et des Pays de l Adour, E2S UPPA, CNRS, LMAP, Pau, France.
E-mail: sophie.mercier@univ-pau.fr

A rapid pace of climate change is now becoming evident by a marked increase in the frequency and intensity of weather extremes, and this trend is expected to continue with an increase in global warming in the coming decades. The paper presents the linear extension of the Yule process (LEYP) as a general stochastic model of environmental hazards induced by non-stationary climate conditions. The LEYP is a more versatile model than the Poisson process, as it can incorporate dependence among events occurring over time. In the paper, explicit expressions have been derived for the return period, a traditional measure of reliability that is commonly used in the design of infrastructure systems. Unlike the stationary climate, the return period between extreme events would continue to decrease as climate change effects would become more pronounced in the future. Examples presented in the paper demonstrate that a modest degree of statistical dependence among events leads to a significant reduction in the return period, i.e., a remarkable increase in the frequency of extreme events. Therefore, existing design codes would need to be revised to accommodate such non-stationary changes to ensure a high level of safety of infrastructure systems in the changing climate.

Keywords: Climate change, Non-stationary processes, Yule process, Non-homogeneous Poisson process, Return period, Extreme events.

1. Introduction

Climate is changing due to increase in the greenhouse gas emissions caused by human activities. The increasing pace of climate change is manifesting through increased frequency and intensity of extreme weather events, such as heat waves, droughts, rain storms, hurricanes, and wildfires in many parts of the world. Increasing severity of weather events is threatening the safety and functionality of existing infrastructure systems, namely, buildings, bridges, roads, transit systems, water supply, storm water and sewage systems, and adding the burden of costly repair of damaged systems.

Current infrastructure design codes are based on an implicit assumption that climate-induced stresses are stationary, i.e., their occurrence frequency and intensity do not change with chronological time. Under this assumption, historical time series of environmental extreme events are analyzed to estimate appropriate reliability measures used in the design of infrastructure systems. The return period is such a measure that is commonly used to specify a design scenario (event) in codes and standards.The return period can be defined as the average time between two consecutive extreme events. For example, 50-year wind speed is used in the building design, and 50 year rain event is used in the design of storm water systems. The flood protection structures are design for return periods ranging from 500 to 10,000 years, depending on the magnitude of losses caused by flooding.

The climate change causing temporal variations in the occurrence pattern of weather extremes raises a question about the validity of currently used stationary model of hazards in the future. This motivates the development of non-stationary stochastic models of environmental hazards, which can be used to revise existing codes to make them compatible with plausible scenarios of climate change.

Several studies have discussed methods for computing the return period in the non-stationary climate (Read and Vogel, 2015). However, almost all the previous studies primarily focused on the computation of the FIRST occurrence of the extreme event without recognizing that the subsequent return of extremes would not follow the same distribution of the inter-arrival time (i.e., return period) as the very first one. The point was further elaborated in a recent study which proposed the non-homogeneous Poisson process (NHPP) as a model of time-dependent increase in the frequency of hazards in the non-stationary climate and evaluated its impact on structural reliability (Pandey and Lounis, 2023). However, the "independence" property is a key limitation of the Poisson process, i.e., distributions of the number of events in any two disjoint intervals are independent of each other. The reason is that sustained global warming over a long period of time can introduce dependence among weather extremes. Phenomena like El Niño and La Nina are already known to introduce a short-term dependence among various weather events.

The limitation of NHPP can be overcome by the use of more general forms of the birth processes, such as the Linear Extension of the Yule Process (LEYP) proposed by Le Gat (2009) and its extensions proposed by Badía et al. (2019). In the LEYP, the occurrence frequency of events depends on time, similar to the NHPP, as well as on the number of events that occurred in the past, thereby introducing the dependence between the distributions of the number events in two disjoint intervals.

The main objective of this paper is to derive explicit expressions for the return period of extreme events modelled as LEYP, and quantify the dependence via the correlation coefficient between the number of events in two time intervals. In this manner, this study presents a more general approach to account for non-stationary nature of climate change effects in the design of infrastructure systems.

2. Stochastic Modelling of Extreme Events

2.1. *Background*

Extreme events arriving randomly over time can be modelled as a stochastic point process consisting of an increasing sequence of positive random variables, $S_0 = 0 < S_1 < S_2 < \cdots$, which represent arrival times of events (see Fig.1). A stochastic counting process, $N(t)$, associated with this sequence can be defined as, $N(t) = \sum_{i=1}^{\infty} \mathbf{1}_{\{S_i \leq t\}}$, where $\mathbf{1}_A$ is an indicator function associated with an event A. Note that $N(t)$ denotes the (random) number of events occurring in a time interval, $(0, t]$.

A class of continuous time, discrete state Markov processes, such as the Poisson process, have been widely used in the reliability analysis. Such processes can be defined using a rate function and a set of postulates to characterize the state transition in an infinitesimal time interval (Taylor and Karlin, 1998). First, the probability of state transition, i.e., occurrence of an event, in a small time interval, $(t, t+h]$, is postulated as

$$\begin{aligned} &\mathbb{P}\left[N(t+h) - N(t) = 1 | N(t) = k\right] \\ &= \lambda(k,t)h + o(h), \quad (h \to 0+) \end{aligned} \tag{1}$$

Here, $\lambda(k, t)$ denotes the rate of the process that can depend on time, t, and the number of events, k, that occurred in the past interval, $(0, t]$. Also, $o(h)$ denotes terms of order h. The second postulate defines the probability of no transition as

$$\begin{aligned} &\mathbb{P}\left[N(t+h) - N(t) = 0 | N(t) = k\right] \\ &= 1 - \lambda(k,t)h + o(h) \end{aligned} \tag{2}$$

Using these postulates, a differential equation for $P_n(t) = \mathbb{P}\left[N(t) = n\right]$ can be derived for $n \geq 1$ as

$$P_n'(t) = -\lambda(n,t)P_n(t) + \lambda(n-1,t)P_{n-1}(t) \tag{3}$$

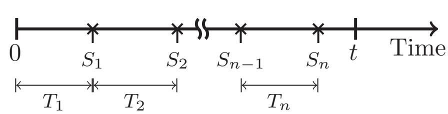

Fig. 1. A point process model of arrival of extreme events

and for $n = 0$, $P_0'(t) = -\lambda(0, t)$. This differential equation can be analytically solved for some specific forms of the rate function, $\lambda(k, t)$.

2.2. *Poisson Process*

When the rate function depends only on time, i.e., $\lambda(k, t) = \lambda(t)$, it leads to a well-known non-homogeneous Poisson process (NHPP) with the following expression for $P_n(t)$:

$$P_n(t) = \frac{(\Lambda(t))^n}{n!} e^{-\Lambda(t)}, \quad (0 \leq k < \infty) \quad (4)$$

where, $\Lambda(t) = \int_0^t \lambda(x)\mathrm{d}x$, is known as the mean value function, i.e., $\Lambda(t) = \mathbb{E}[N(t)]$. Note that a constant rate, i.e., $\lambda(k, t) = \lambda$, leads to the homogeneous Poisson process.

The "independence" is the most important property of the Poisson process, which implies that the number of events in non-overlapping time intervals are independent Poisson random variables. The independence property results in a considerable simplification of analysis of problems related to the Poisson process. On the other hand, this property is a main limitation of the model, as it precludes the modelling of a sequence of dependent events.

2.3. *Birth Processes*

When the rate of occurrence of an event at a given time depends on the number of events which have already occurred, it leads to another class of processes, known as the birth process. These processes are more general than the Poisson process, as the number of events in any two disjoint intervals are no longer independent.

The case of $\lambda(k, t) = k\lambda$ is known as the Yule process (Taylor and Karlin, 1998). A more general form of the birth process was introduced by Le Gat (2014) with the following rate function:

$$\lambda(k, t) = (ak + b)\lambda(t) \quad (5)$$

where k is the number of events occurred in $(0, t)$ with $a \geq 0$ and $b > 0$. The resulting process was named as the "Linear Extension of the Yule Process" (LEYP), and also referred to as the Generalized Polya Process (GPP) by Cha (2014). In the following, we assume that $a > 0$, which is not a restriction, as when $a = 0$, the LEYP becomes a NHPP for which all results can be found in Pandey and Lounis (2023).

In the LEYP model, the distribution of $P_n(t)$ follows the negative binomial distribution with the following standard form:

$$P_n(t) = \frac{\Gamma(\alpha + n)}{n!\Gamma(\alpha)} (\beta)^\alpha (1 - \beta)^n \quad (6)$$

for all $n \in \mathbb{N}^*$. Note that α and β are the distribution parameters, and $\Gamma(\cdot)$ is the gamma function. For the sake of brevity, this distribution is denoted as $\mathcal{NB}(\alpha, \beta)$. The mean ($\mu$) and variance ($\sigma^2$) of this distribution are given as

$$\mu = \alpha \frac{1 - \beta}{\beta}, \text{ and } \quad \sigma^2 = \frac{\mu}{\beta} \quad (7)$$

Other probabilistic properties of LEYP are discussed in more details in the next Section.

3. LEYP: Probabilistic Properties

3.1. *The Number of Events*

The marginal distribution of the number of events in a time interval, $(0, t]$, is given as Konno (2010)

$$\mathbb{P}(N(t) = n) = \frac{\Gamma\left(n + \frac{b}{a}\right)}{n!\Gamma\left(\frac{b}{a}\right)} \left(e^{-a\Lambda(t)}\right)^{\frac{b}{a}} \left(1 - e^{-a\Lambda(t)}\right)^n \quad (8)$$

This distribution is $\mathcal{NB}(\alpha, \beta)$ with $\alpha = b/a$ and $\beta = e^{-a\Lambda(t)}$. Its mean and variance can be computed using Eq.7.

Using the restarting property of the LEYP, as defined by Cha (2014), the distribution of an increment, $N(s, t) = N(t) - N(s)$, conditioned on $N(s^-) = k$, $0 < s < t$, is given as

$$\mathbb{P}\left(N(s, t) = n | N\left(s^-\right) = k\right) = \frac{\Gamma\left(n + k + \frac{b}{a}\right)}{n!\Gamma\left(\frac{b}{a} + k\right)} \left(e^{-a\Lambda(s,t)}\right)^{\frac{b}{a}+k} \left(1 - e^{-a\Lambda(s,t)}\right)^n \quad (9)$$

for all $k, n \in \mathbb{N}$. This distribution is $\mathcal{NB}(\frac{b}{a} + k, e^{-a\Lambda(s,t)})$, with $\Lambda(s, t) = \int_s^t \lambda(x)\mathrm{d}x$.

3.2. *Distribution of Arrival Times*

Using the fact that the events $[S_n > t]$ and $[N(t) \leq n]$ are equivalent, the marginal distribution of S_n can be written as,

$$\overline{F}_{S_n}(t) = \sum_{k=0}^{n-1} \mathbb{P}\left[N(t) = k\right] \tag{10}$$

Based on a result given by Johnson et al. (2005, Section 5.6), this can also be written as

$$\bar{F}_{S_n}(t) = I_{e^{-a\Lambda(t)}}\left(\frac{b}{a}, n\right) \tag{11}$$

where $I_x(\alpha, \beta)$ is the incomplete beta function. The pdf of S_n can now be derived as $f_{S_n}(t) = -\bar{F}'_{S_n}(t)$, which leads to

$$f_{S_n}(t) = \frac{a}{B\left(\frac{b}{a}, n\right)} \lambda(t)\, e^{-b\Lambda(t)} \left(1 - e^{-a\Lambda(t)}\right)^{n-1} \tag{12}$$

where $B(x, y)$ denotes the complete Beta function.

3.3. *A Measure of Dependence*

In case of a LEYP, the dependence between two successive increments, $N_{12} = N_{t_2} - N_{t_1}$ and $N_{23} = N_{t_3} - N_{t_2}, 0 \leq t_1 < t_2 < t_3$, can be evaluated in terms of the correlation coefficient that is defined in a usual manner as

$$\rho(N_{12}, N_{23}) = \frac{COV\left[N_{12}, N_{23}\right]}{\sigma\left[N_{12}\right]\sigma\left[N_{23}\right]} \tag{13}$$

where the numerator is the covariance between N_{12} and N_{23}. Here, the main task is the derivation of an expression for the product moment of increments, i.e., $\mathbb{E}\left[N_{12}N_{23}\right]$, which is presented in Appendix A. Expressions for the mean and standard deviation of an increment are given by Le Gat (2015). Using these expressions, the final expression for the correlation coefficient is derived as

$$\left[\rho\left(N_{12}, N_{23}\right)\right]^2 = \frac{\left(e^{a\Lambda(t_2)} - e^{a\Lambda(t_1)}\right)\left(e^{a\Lambda(t_3)} - e^{a\Lambda(t_2)}\right)}{\left(1 + e^{a\Lambda(t_2)} - e^{a\Lambda(t_1)}\right)\left(1 + e^{a\Lambda(t_3)} - e^{a\Lambda(t_2)}\right)} \tag{14}$$

4. Analysis of the Return Period

4.1. *Mean Inter-Arrival Time*

This Section derives the probability distribution and the mean of an n^{th} inter-arrival time, $T_n = S_n - S_{n-1}, n > 1, S_0 = 0$, which is also referred to as the n^{th} return period.

The event $[T_n > u]$ is equivalent to the event that no shocks occur in the interval, $(S_{n-1}, S_{n-1} + u]$. Thus,

$$\begin{aligned}\mathbb{P}\left[T_n > u\right] &= \mathbb{P}\left[N(S_{n-1}, S_{n-1} + u) = 0\right] \\ &= \int_0^\infty \mathbb{P}\left[N(s, s+u) = 0 | S_{n-1} = s\right] f_{S_{n-1}}(s)\mathrm{d}s\end{aligned} \tag{15}$$

The case of $n = 1$ is a especial case with $\Lambda(0) = 0$ for which the distribution of T_1 can be directly obtained using Eq.(8) as

$$\overline{F}_{T_1}(u) = \mathbb{P}\left[N(u) = 0\right] = e^{-b\Lambda(u)} \tag{16}$$

In the case of $n \geq 2$, the conditional probability in Eq.(15) can be simplified by noting that $[S_{n-1} = s] \equiv [N(0, s) = n - 1]$, i.e.,

$$\begin{aligned}&\mathbb{P}\left[N(s, s+u) = 0 | S_{n-1} = s\right] \\ &= \mathbb{P}\left[N(s, s+u) = 0 | N(0, s) = n - 1\right]\end{aligned}$$

Substituting for the conditional probability from Eq.(9) and for $f_{S_{n-1}}(s)$ from Eq.(12) into Eq.(15) leads to

$$\bar{F}_{T_n}(u) = \frac{a}{B\left(\frac{b}{a}, n-1\right)} \times \int_0^\infty \lambda(s)\, e^{-g(u,s)} \left(1 - e^{-a\Lambda(s)}\right)^{n-2} ds \tag{17}$$

where the function, $g(u, s)$, is defined as

$$g(u, s) = a(n-1)\Lambda(s, u+s) + b\Lambda(u+s) \tag{18}$$

The expected value of a positive random variable, $T_n > 0, N \geq 1$, can be obtained by integrating the complementary CDF function as

$$\mathbb{E}\left[T_n\right] = \int_0^\infty \overline{F}_{T_n}(u)\mathrm{d}u \tag{19}$$

Thus, any n^{th} return period can be computed using Eqs.(17) and (19).

In case of an NHPP with the rate function, $\lambda(t)$, the following expression for the return period was derived by Pandey and Lounis (2023):

$$\mathbb{E}\left[T_n\right] = \int_0^\infty e^{-\Lambda(s)} \frac{\left[\Lambda(s)\right]^{n-1}}{(n-1)!} \mathrm{d}s \tag{20}$$

4.2. *Waiting Time for the Next Event*

At any given time t, the waiting time, $W(t)$, to the occurrence of the next event can be defined, as shown in Figure 2, as $W(t) = S_{N(t)+1} - t$. The distribution of $W(t)$ can be obtained from the relation, $\mathbb{P}[W(t) > w] = \mathbb{P}[N(t, t + w) = 0]$, and noting that $N(s,t), s < t$, follows the negative binomial distribution with parameters, $\alpha = b/a$ and $1/\beta = 1 + e^{a\Lambda(t)} - e^{a\Lambda(s)}$ (Le Gat, 2015). Thus,

$$\overline{F}_W(w;t) = \left(\frac{1}{1 + e^{a\Lambda(t+w)} - e^{a\Lambda(t)}}\right)^{b/a} \tag{21}$$

The mean waiting time can then be obtained as

$$\mathbb{E}[W(t)] = \int_0^\infty \overline{F}_W(w;t)\mathrm{dw} \tag{22}$$

5. Numerical Examples

The impact of the non-stationary nature of a hypothetical hazard process is illustrated through several examples. The stationary climate condition is assumed as the base case in which the hazard is modelled as the homogeneous Poisson process (HPP) with the rate, $\lambda_o = 0.02$ events/year, which corresponds to a return period of 50 years. The time horizon of the analysis is taken as $t_e = 80$ years (from 2020 to 2100).

Under the non-stationary climate, the rate of occurrence of hazards is assumed to increase linearly from λ_o in 2020 to $k\lambda_o$ in 2100 as,

$$\lambda(k, s) = \lambda_o + (k-1)\lambda_o \frac{s}{t_e} \tag{23}$$

In this study, k is referred to as the frequency amplification factor. The rate function of LEYP is given by Eq.(5) with $b = 1$ and $\lambda(k,t)$ given by Eq.(23) for all numerical examples.

First the NHPP model is analyzed with the rate function given by Eq.(23) as a reference solution. Figure 1 plots the first six return periods for $k = 1, 2$, and 4, computed using Eq.(20). The case of $k = 1$ represents the stationary Poisson process with a constant return period of 50 years. In case of $k = 2$, the first return period decreases to 36 years and the 6th return period to 18 years. The decrease in the return period is further amplified with an increase in the non-stationary effect when $k = 4$.

In case of an LEYP model, first the dependence is quantified in terms of the correlation coefficient between the number of events in two consecutive intervals, i.e., $N(t-u, t)$ and $N(t, t+u)$ for $a = 0.5, 1$ and 2. Using Eq.(14), the Correlation

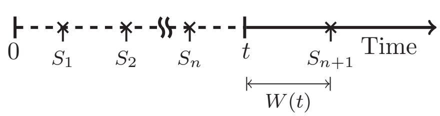

Fig. 2. Definition of the waiting time to the next event at time t

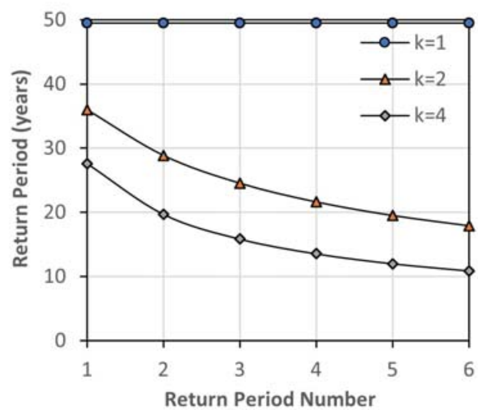

Fig. 3. Return periods for the NHPP model

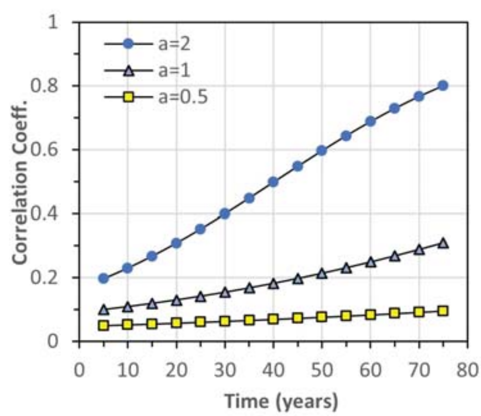

Fig. 4. Variation of the correlation coefficient with time in LEYP model

coefficient, $\rho(t, u)$ is computed as a function of time, t and $u = 5$ years, and plotted in Figure 4. It is clear that the correlation coefficient increases with increasing values of the parameter a and time t. For $a = 0.5$, $\rho(t, 5)$ increases rather modestly from 0.05 at $t = 5$ to 0.1 at $t = 75$. In case of $a = 1$, $\rho(t, 5)$ increases from 0.1 to 0.3, and for $a = 2$, it increases from 0.2 to 0.8 over the same period.

The effect of dependence can be further illustrated by comparing plots of the expected number of events versus time for LEYP and corresponding NHPP models, as shown in Figure 5. The expected number of events over an 80 year period, $\mathbb{E}[N(80)]$, is computed as 1.6. In case of an NHPP with $k = 2$, this expected value increases to 2.4. However, in the related LEYP model with $k = 2$ and $a = 0.5$, the value of $\mathbb{E}[N(80)]$ is almost doubled to 4.64. The increase in the expected number of events is much more dramatic for an LEYP with $k = 4$ and $a = 0.05$. In this case, $\mathbb{E}[N(80)]$ increases to 12.77, almost three times the value for a corresponding NHPP. It should be noted from Figure 4 that $a = 0.5$ implies a fairly modest degree of dependence with maximum correlation coefficient of 0.1. In spite of this, a significant increase is observed in the expected number of events. In summary, the dependence in the LEYP further amplifies the increase in the frequency of extreme events caused by a non-stationary rate of occurrence of the process.

The first six return periods were computed for the LEYP model with $a = 0.5$ and the three values of the frequency amplification factor, $k = 1, 2,$ and 4, and results are plotted in Figure 6. It should be noted that the first return period (RP) in the LEYP model is the same as that in the corresponding NHPP model. However, subsequent RPs of the LEYP model decrease significantly as compared to those of the corresponding NHPP models, as shown in Figure 3. Even in case of the stationary climate, i.e., $k = 1$, a large reduction is seen in return periods of the LEYP as compared to the 50-year reference value. For example, the second RP is reduced to 33.33 years and the sixth RP to 14.28 years when $k = 1$ and $a = 0.5$. This reduction is due entirely to the dependence property of the LEYP model.

The dependence combined with a non-stationary rate amplifies the reduction in RP values. For example, for $k = 2$ and $a = 0.5$, the second RP of LEYP model reduces to 20.5 years, whereas the corresponding NHPP model has this value as 28.9 years. If a is increased to 1, the second RP of the LEYP further reduces to 15.9 years, as shown in Figure 7.

Plots of the mean waiting times over a period of 0-80 years are presented in Figure 8 for the LEYP model with $a = 0.5$. The mean waiting time for

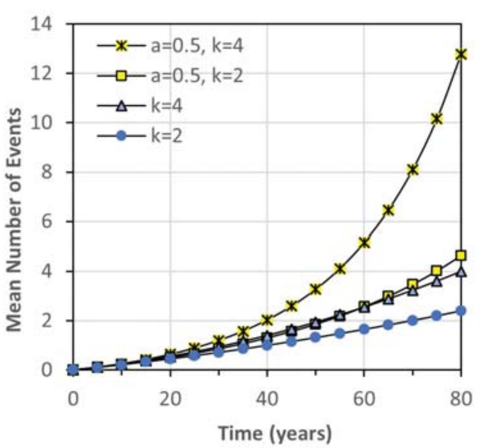

Fig. 5. A comparison of the expected number of events in NHPP and LEYP models

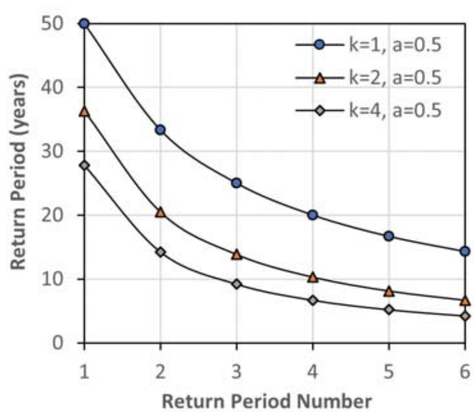

Fig. 6. A sequence of return periods for LEYP models with $a = 0.5$

the next event also decreases with the passage of time due to the combined effect of non-stationary rate and the dependence in the LEYP model. The mean waiting time could be a more useful measure, as it does not require any information about the number of events that occurred in the past.

6. Conclusions

The non-stationary nature of climate change effects is expected to increase the frequency of weather extremes over time and also introduce some (statistical) dependence among the events. To incorporate the time-dependent frequency and the dependence among events, the paper presents the Linear Extension of the Yule Process (LEYP) as a general stochastic model of environmental hazards. The non-homogeneous Poisson process (NHPP) is a especial case of the LEYP model, which retains the independence property.

To illustrate the proposed approach, the paper postulates a hypothetical example of a linear increase in the occurrence rate of a hazard over an 80-year period of global warming (2020-2100). The degree of climate change is defined by the ratio (k) of occurrence frequency in the 80th year to that in the base case of the stationary climate, and referred to as the climate amplification factor.

Examples presented in the paper show that a time-dependent increase in the occurrence rate, such as that in the NHPP model, leads to a significant reduction in return periods over time. This reduction is further amplified in case of the LEYP model, even with a fairly modest degree of correlation (or dependence) among events. Therefore, ignoring the dependence can result in a significant underestimation of return periods of weather extremes, which can have an adverse impact of the safety of infrastructure systems.

Since the return period would no longer be a constant in the changing climate, its use to specify design scenarios in codes standards would become questionable. The infrastructure design codes must be revised to address this issue.

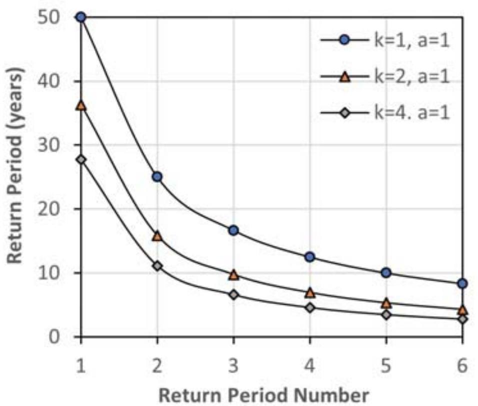

Fig. 7. A sequence of return periods for the LEYP model with $a = 1$

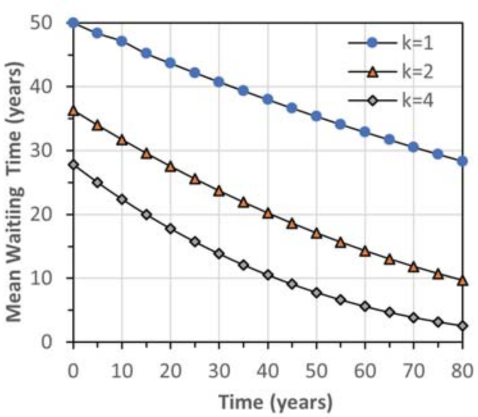

Fig. 8. The mean waiting time for the LEYP model with $a = 0.5$

Acknowledgement

The second authors acknowledges the financial support for this study provided by the Natural Science and Engineering Council of Canada (NSERC).

Appendix A.

To evaluate the correlation coefficient using Eq.(14), the main task is to compute the product moment, $\mathbb{E}\left[\left(N_{t_2} - N_{t_1}\right)\left(N_{t_3} - N_{t_2}\right)\right]$. This Section presents a brief outline of this derivation.

Conditionning by N_{t_1} and $N_{t_2} - N_{t_1}$, leads to

$$\mathbb{E}\left[\left(N_{t_2} - N_{t_1}\right)\left(N_{t_3} - N_{t_2}\right)\right]$$
$$= \mathbb{E}\left[\mathbb{E}\left[\left(N_{t_2} - N_{t_1}\right)\left(N_{t_3} - N_{t_2}\right)|N_{t_1}, N_{t_2} - N_{t_1}\right.\right.$$
$$= \mathbb{E}\left[\left(N_{t_2} - N_{t_1}\right)\mathbb{E}\left(N_{t_3} - N_{t_2}|N_{t_2}\right)\right] \quad \text{(A.1)}$$

as $N_{t_3} - N_{t_2}$ only depends on N_{t_1} and $N_{t_2} - N_{t_1}$ through $N_{t_2} = N_{t_1} + \left(N_{t_2} - N_{t_1}\right)$.

Based on the restarting property (Cha, 2014), it can be shown that

$$[N_{t_3} - N_{t_2}|N_{t_2} = k] \sim \mathcal{NB}\left(\frac{b}{a} + k, e^{-a\Lambda(t_2,t_3)}\right)$$

such that its expectation can be given as,

$$\mathbb{E}\left(N_{t_3} - N_{t_2}|N_{t_2}\right) = \left(\frac{b}{a} + N_{t_2}\right)\left(e^{a\Lambda(t_2,t_3)} - 1\right) \tag{A.2}$$

With this relation, the expectations given in Eq.(A.1) can be evaluated after a considerable analytical simplifications, which finally leads to an expression for the correlation coefficient given by Eq.(14).

References

Badía, F. G., S. Mercier, and C. Sangüesa (2019). Extensions of the generalized Pólya process. *Methodology and Computing in Applied Probability 21*, 1057–1085.

Cha, J. H. (2014, 12). Characterization of the generalized Pólya process and its applications. *Advances in Applied Probability 46*, 1148–1171.

Johnson, N. L., A. W. Kemp, and S. Kotz (2005). *Univariate discrete distributions*, Volume 444. New York, NY: John Wiley & Sons.

Konno, H. (2010). On the exact solution of a generalized Pólya process. *Advances in mathematical physics 2010, 504267*, 1–12.

Le Gat, Y. (2009). *Une extension du processus de Yule pour la modélisation stochastique des événements récurrents: application aux défaillances de canalisations d'eau sous pression*. Ph. D. thesis, AgroParisTech.

Le Gat, Y. (2014). Extending the yule process to model recurrent pipe failures in water supply networks. *Urban Water Journal 11*(8), 617–630.

Le Gat, Y. (2015). *Recurrent Event Modeling Based on the Yule Process: Application to Water Network Asset Management*. John Wiley & Sons.

Pandey, M. D. and Z. Lounis (2023). Stochastic modelling of non-stationary environmental loads for reliability analysis under the changing climate. *Structural Safety 103, 102348*, 1–11.

Read, L. K. & R. M. Vogel (2015). Reliability, return periods, and risk under nonstationarity. *Water Resources Research 51*(8), 6381–6398.

Taylor, H. M. and S. Karlin (1998). *An Introduction to Stochastic Modeling* (3rd ed.). San Diego, CA: Academic Press.