Stochastic modelling of non-stationary and dependent weather extremes for structural reliability analysis in the changing climate

Mahesh D Pandey, Sophie Mercier

Structural Safety 114 (2025) 102569

Contents lists available at ScienceDirect

Structural Safety

journal homepage: www.elsevier.com/locate/strusafe

Stochastic modelling of non-stationary and dependent weather extremes for structural reliability analysis in the changing climate

Mahesh D. Pandey [a,*], Sophie Mercier [b]

[a] Department of Civil and Environmental Engineering, University of Waterloo, Waterloo, ON, Canada, N2L 3G1
[b] Universite de Pau et des Pays de l Adour, E2S UPPA, CNRS, LMAP, Pau, France

ARTICLE INFO

Keywords:
Climate change
Global climate models
Weather extremes
Stochastic processes
Non-stationary loads
Birth processes
Linear Extension of Yule process
Non-homogeneous Poisson process
Return period
Extreme value distribution
Structural reliability

ABSTRACT

In recent times, the safety of infrastructure systems has been challenged by the increasing severity of extreme weather events caused by the effects of climate change . This trend is expected to continue, as shown by the simulations of future climate conditions under high-emission scenarios. The paper presents a general stochastic process, known as the Linear Extension of the Yule Process (LEYP), to model the non-stationary frequency and intensity of extremes. The LEYP model overcomes a major limitation of the classical Poisson process by including the statistical dependence among extreme events.

The paper presents a probabilistic framework for non-stationary structural reliability analysis, which includes new results for the return period, waiting time for the next event, correlation coefficient, and the distribution of the maximum load in a given time interval. The examples provided in the paper demonstrate that even a modest degree of dependence can significantly reduce the interval between events and increase the probability of failure with time. Furthermore, the paper illustrates the non-stationary modelling of future precipitation data, as simulated by the Canadian Earth Systems Model (CanESM5). The results of this study are expected to be useful for revising current "stationary" design codes and ensuring structural safety in the changing climate.

1. Introduction

1.1. Background

The increase in greenhouse gas emissions is contributing to climate change, as seen by long-term shifts in the temperature and weather patterns across the globe. These changes are leading to the increased frequency and intensity of extreme weather events, such as heat waves, droughts, precipitation, and wildfires in many parts of the world. The increasing severity of weather extremes is challenging the safety of existing infrastructure and increasing the repair cost of damaged systems. Therefore, the prediction of weather extremes has become an emerging area of research to ensure the structural safety in the changing climate.

Global Climate Models (GCMs) have become important tools for investigating the impact of different emission scenarios on the future climate. The GCM is essentially a complex mathematical representation of major components of the atmosphere, land surface, ocean, and sea ice systems and interactions among them.

Future climate data simulated by GCMs are increasingly used in predicting the trends in extreme weather events. Kharin and Zwiers [1] pioneered the statistical analysis of simulated climate extremes under different scenarios of anthropogenic forcing. The non-stationary GEV/Gumbel distribution was used to model extremes of temperature and precipitation, and the annual maxima method was used to estimate the time-dependent distribution parameters [2]. In the climate literature, this method has become a standard approach for detecting and modelling trends in future extremes [3]. The non-stationary Gumbel distribution was used in the structural reliability analysis under changing climate [4,5]. The effect of climate change on the tsunami hazard was investigated by Alhamid et al. [6].

The GEV/Gumbel distribution provides a high-level approach to incorporate the overall effect of non-stationary changes in extreme values generated by an environmental process. However, it cannot attribute the overall change to the frequency and intensity of the process. Suppose, for example, that the GEV model predicts an overall increase of $x\%$ in the 95th percentile of the annual maximum precipitation. However, it cannot separate the individual contributions made by the changes in frequency and intensity to this overall change. Furthermore, the GEV model does not allow to evaluate non-stationary changes in the distributions of inter-arrival times of extreme events. Such limitations

* Corresponding author.
E-mail address: mdpandey@uwaterloo.ca (M.D. Pandey).

https://doi.org/10.1016/j.strusafe.2024.102569
Received 4 June 2024; Received in revised form 19 October 2024; Accepted 16 December 2024
Available online 24 December 2024

of the GEV approach can be overcome through a more formal modelling of extreme events as a stochastic process, such as the non-homogeneous Poisson process (NHPP) proposed by Pandey and Lounis [7].

A main drawbacks of the non-stationary GEV and NHPP models is the assumption of "independence". The NHPP model assumes that the distributions of the number of events in any two disjoint intervals are independent. Similarly, the GEV model assumes that the annual maximum distribution for a given year is independent of all other years. However, the assumption of independence might be tenuous in the changing climate, since a gradual buildup of greenhouse gases over time can introduce the dependence among extreme events. Natural phenomena such as El Niño and La Niña are known to introduce short-term dependence among various weather events. To address this limitation, the paper proposes a more general model based on the non-homogeneous birth process, which can model non-stationary frequency, similar to the NHPP, and also includes the dependence on the history of the process [8].

1.2. Research objectives

The main objective of this paper is to develop a general stochastic process model that can explicitly consider non-stationary changes in the frequency and intensity of extreme events, as well as statistical dependence among them. The Linear Extension of the Yule Process (LEYP) is selected for modelling purposes. Another objective is to illustrate the statistical estimation of non-stationary effects using future precipitation data simulated by the Canadian Earth System Model (CanESM5).

The proposed LEYP model leads to a precise, finite-time solution for the extreme value distribution, which is an important input for structural reliability analysis. This model also shows that traditional concepts, such as the return period, are no longer applicable in the non-stationary climate, highlighting the need for revising existing codes and standards.

1.3. Organization

The next Section introduces the basic terminology of the stochastic point process, and Section 3 presents probabilistic properties of the proposed LEYP model. In Section 4, analytical solutions are presented for the return period, defined as the average inter-arrival time, and the mean waiting time for the next event. The distribution of LEYP extremes and its application to reliability analysis are presented in Section 5. For a homogeneous birth process, closed-form analytical results are presented in Section 6 for the nth return period, correlation coefficient, and maximum value distribution. These results serve as reference solutions for parametric studies and verification of numerical results. Additionally, two sets of numerical examples are presented in Section 7 for the homogeneous and non-homogeneous forms of the birth process. Section 8 presents a practical example of modelling of heavy precipitation events using the future climate data (2020–2100). The conclusions of this study are summarized in the final Section. Appendices A to C include mathematical derivations of several analytical results presented in the paper.

2. Stochastic point process

2.1. Definitions

A sequence of events arriving randomly over time can be modelled as a point process in which arrival times are represented by an increasing sequence of positive random variables, $S_0(=0) < S_1 < S_2 < \cdots$, (see Fig. 1). A counting process, $N(t)$, is associated with this sequence, which denotes the number of events occurred in an interval $(0, t]$.

A point process can be specified using the Conditional Intensity (CI) function, $\lambda_C(\cdot)$, which represents the rate of occurrence of an event at time t, conditioned on the history of the process over the interval, $[0, t)$. It is formally defined as [9, p.232],

$$\lambda_C(t)\mathrm{d}t \approx \mathbb{E}\left[\mathrm{d}N(t)|\mathcal{H}(t^-)\right] \tag{1}$$

where $\mathcal{H}(t^-)$ denotes the history, i.e., the number of events, $N(t^-)$, occurred up to just before time t.

This paper considers a specific form of the birth point processes with the CI function,

$$\lambda_C(t) = h(N(t^-))\,\lambda(t) \tag{2}$$

which includes two distinct components:

1. A stochastic function, $h(N(t^-))$, which depends on the history of the process, and
2. A time-dependent rate function, $\lambda(t)$.

Such a processes is also known as the self-exiting process, since its stochastic intensity at time t depends on the number of events occurred up to time t^-.

The probability of occurrence of an event in an infinitesimal interval, $(t, t+\mathrm{d}t]$, can be written in terms of the CI function as [10],

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

Here, $o(\mathrm{d}t)$ denotes the terms of the order of $\mathrm{d}t$. Using this relation, a differential equation for $P_n(t) = \mathbb{P}\left[N(t) = n\right]$ can be derived as

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

with $P_0'(t) = -h(0)\lambda(t)$ for $n = 0$. This differential equation can be analytically solved for some specific forms of the CI function, as discussed next.

2.1.1. Poisson process

When the conditional intensity is independent of the history, i.e., $\lambda_C(t) = \lambda(t)$, it leads to a well known Non-Homogeneous Poisson Process (NHPP) with an explicit solution of Eq. (3):

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

where $\Lambda(t)$ is known as the mean value function, i.e., $\Lambda(t) = \mathbb{E}\left[N(t)\right] = \int_0^t \lambda(x)\mathrm{d}x$. A process with a constant rate, $\lambda(t) = \lambda$, is known as the homogeneous Poisson process (HPP).

The "independence" is the most important property of this process, implying that the number of events in non-overlapping intervals are independent Poisson random variables. Although the independence property leads to significant simplifications in the analysis, it does not allow the NHPP to model a sequence of dependent events, which is a major limitation the NHPP model.

2.1.2. Pólya and Yule processes, and extensions

Some simple forms of the history function, $h(\cdot)$, in Eq. (2) lead to analytical solutions of Eq. (3). For example, $h(n) = n$ leads to the classical Yule process [10].

A process with a linear function, $h(n) = an + b$, is referred to as the generalized Pólya process (GPP) [11]. Le Gat [12] refers the GPP with a fixed parameter, $b = 1$, as the "Linear Extension of the Yule Process" (LEYP).

Badía et al. [8] analysed more general forms of the history function, such as $h(n) = (a+n)^{\rho_1}$ and $h(n) = \rho_2^n$ for $\rho_1 \in \mathbb{R}$ and $\rho_2 > 0$. Such processes are referred as the "Extended Pólya Process" (EPP).

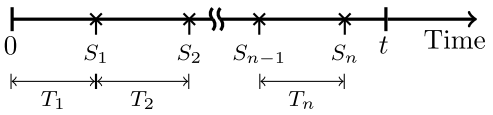

Fig. 1. A point process model of random arrivals of events.

Since the conditional intensity of the LEYP depends on its history, the number of events in any two disjoint intervals are no longer independent. In this manner, the LEYP model overcomes the main limitation of the Poisson process.

2.2. Negative binomial distribution

Konno [13] presented an analytical solution of Eq. (3) for an LEYP and proved that $N(t)$ follows the negative binomial (NB) distribution with the following probability mass function (PMF):

$$P_n(t) = \frac{\Gamma(\alpha+n)}{\Gamma(n+1)\Gamma(\alpha)}(\beta)^{\alpha}(1-\beta)^{n} \quad (n \geqslant 0) \tag{5}$$

Here, $\alpha \geqslant 0$ and $\beta \geqslant 0$, are the distribution parameters, and $\Gamma(\cdot)$ denotes the gamma function. Note that $\Gamma(n+1) = n!$ for an integer n. For the sake of brevity, this distribution is denoted as $\mathcal{NB}(\alpha, \beta)$ in the subsequent discussions.

Applying a normalizing constraint, $\sum_{n=0}^{\infty} P_n(t) = 1$, in conjunction with Eq. (5) leads to a useful relation:

$$\sum_{n=0}^{\infty} \frac{\Gamma(\alpha+n)}{\Gamma(n+1)\Gamma(\alpha)}(1-\beta)^{n} = (\beta)^{-\alpha} \tag{6}$$

The mean (μ_N), mean square (μ_{2N}), and variance (σ_N^2) of this distribution are given as [14]:

$$\mu_N = \alpha\frac{1-\beta}{\beta},\ \mu_{2N} = \mu_N\left(\frac{1}{\beta}+\mu_N\right), \text{ and } \quad \sigma_N^2 = \frac{\mu_N}{\beta} \tag{7}$$

3. LEYP: Probabilistic properties

This Section summarizes basic probabilistic properties of the non-homogeneous LEYP, which will be required to derive expressions for the inter-arrival time and the maximum load distribution. This process is defined by a linear history function, $h(n) = an+b$, and a positive time-dependent rate function, $\lambda(t)$. The stochastic CI function is therefore given as,

$$\lambda_C(n,t) = h(n)\lambda(t) = (an+b)\lambda(t), \quad (a \geq 0, b > 0) \tag{8}$$

Note that the statistical identifiability of the parameters of $h(n)$ can be ensured by fixing $b = 1$, as discussed by Badía et al. [8]. The LEYP is equivalent to the NHPP when $a = 0$.

For the sake of brevity, the time interval associated with a variable is specified by its subscript. For example, N_t denotes the number of events in an interval, $(0,t]$, and N_{st} the number of events in $(s,t]$ with $s < t$.

3.1. Number of events

The marginal distribution of the number of events in $(0,t]$ follows $\mathcal{NB}(\alpha,\beta_t)$ [13]:

$$\mathbb{P}\left[N_t = n\right] = \frac{\Gamma(\alpha+n)}{\Gamma(n+1)\Gamma(\alpha)}\left(\beta_t\right)^{\alpha}\left(1-\beta_t\right)^{n} \tag{9}$$

where the distribution parameters are defined as

$$\alpha = \frac{b}{a}, \quad \beta_t = e^{-a\Lambda(t)}, \text{ and } \Lambda(t) = \int_0^t \lambda(u)\mathrm{d}u, \quad (\alpha > 0, \lambda(t) \geqslant 0) \tag{10}$$

It can be shown that β_t is bounded between 0 and 1 as follows. As $t \to \infty$, $\Lambda(t) \to \infty$, such that $\beta_t \to 0$. Similarly, $t \to 0$ leads to $\Lambda(t) \to 0$ and $\beta_t \to 1$. The bounded nature of β_t ensures the numerical stability of computational algorithms. The mean and standard deviation of N_t are obtained using Eq. (7) as

$$\mu_t = \alpha\left(\frac{1}{\beta_t}-1\right), \quad \sigma_t = \sqrt{\frac{\mu_t}{\beta_t}} \tag{11}$$

The "restarting property" of the LEYP is highly useful in deriving various analytical results presented in this paper. Let $N(t)$ be an LEYP with parameters, $(\lambda(t), a, b)$. At an arbitrary time, s and given the history, $N(s^-) = k$, the conditional future process, $N(s+t), t \geqslant 0$, is also an LEYP with parameters, $(\lambda(s+t), a, b+ak)$ [12].

The restarting property implies that the increment, $N(s,t) = N(t) - N(s)$, conditioned on $N(s) = k$, $0 < s < t$, follows the distribution, $\mathcal{NB}(\alpha+k, \beta_{st})$,

$$\mathbb{P}\left[N_{st} = n | N_s = k\right] = \frac{\Gamma(\alpha+k+n)}{\Gamma(n+1)\Gamma(\alpha+k)}\left(\beta_{st}\right)^{\alpha+k}\left(1-\beta_{st}\right)^{n} \tag{12}$$

where

$$\beta_{st} = e^{-a\Lambda(s,t)} \quad \text{and} \quad \Lambda(s,t) = \int_s^t \lambda(u)\mathrm{d}u = \Lambda(t) - \Lambda(s), \quad \text{such that}$$

$$\beta_{st} = \frac{\beta_t}{\beta_s}, \quad (s \leqslant t).$$

Thus, the marginal distribution of N_{st} can be derived using the total probability theorem as

$$\mathbb{P}\left[N_{st} = n\right] = \sum_{k=0}^{\infty}\mathbb{P}\left[N_{st} = n, N_s = k\right] = \sum_{k=0}^{\infty}\mathbb{P}\left[N_{st} = n | N_s = k\right]\mathbb{P}\left[N_s = k\right] \tag{13}$$

Substituting from Eqs. (9) and (12) and using the property of the NB series given by Eq. (6) leads to the following result:

$$\mathbb{P}\left[N_{st} = n\right] = \frac{\Gamma(\alpha+n)}{\Gamma(n+1)\Gamma(\alpha)}(\beta_{st}^*)^{\alpha}(1-\beta_{st}^*)^{n} \tag{14}$$

This distribution is $\mathcal{NB}(\alpha,\beta_{st}^*)$ with

$$\beta_{st}^* = \frac{\beta_t}{1+\beta_t-\beta_{st}}, \quad \text{or} \quad \frac{1}{\beta_{st}^*} = 1 + \frac{1}{\beta_t} - \frac{1}{\beta_s} \tag{15}$$

The parameter β_{st}^* is also bounded between 0 and 1, as $t \to \infty$ and $t \to 0$, respectively.

Using Eqs. (7) and (15), the expected number of events in $(s,t]$ can be given as

$$\mathbb{E}\left[N_{st}\right] = \mu_{st} = \alpha\frac{1-\beta_{st}^*}{\beta_{st}^*} = \alpha\frac{1-\beta_{st}}{\beta_t}, \quad (s \leqslant t)$$
$$\text{or} \quad \frac{\mu_{st}}{\alpha} = \frac{1}{\beta_t} - \frac{1}{\beta_s} \tag{16}$$

The variance of N_{st} is given by

$$\mathrm{VAR}[N_{st}] = \sigma_{st}^2 = \frac{\mu_{st}}{\beta_{st}^*} \tag{17}$$

3.2. Correlation coefficient: A measure of dependence

In an LEYP, the number of events in the two adjacent intervals, $(t_1,t_2]$ and $(t_2,t_3]$ with $0 \leq t_1 < t_2 < t_3$, are statistically dependent. The degree of dependence can be quantified in terms of the coefficient of correlation, as discussed here.

For the sake of brevity, the number of events in an interval, $(t_i,t_j], t_i < t_j$, is denoted as N_{ij}. Similarly, μ_{ij} and σ_{ij} denote the mean and standard deviation of N_{ij}, which can be computed using Eqs. (16) and (17), respectively.

The correlation coefficient between N_{12} and N_{23} is defined in the usual manner as

$$\rho[N_{12}, N_{23}] = \frac{\mathrm{COV}\left[N_{12}, N_{23}\right]}{\sigma_{12}\,\sigma_{23}} \tag{18}$$

where $\mathrm{COV}[\cdot,\cdot]$ denotes the covariance between the variables inside the bracket. Based on the derivation presented in Appendix A, the following concise expression is derived:

$$\left[\rho\left(N_{12}, N_{23}\right)\right]^2 = \left(\frac{\mu_{12}}{\alpha+\mu_{12}}\right)\left(\frac{\mu_{23}}{\alpha+\mu_{23}}\right) = (1-\beta_{12}^*)(1-\beta_{23}^*) \tag{19}$$

Sensitivity of the correlation coefficient to the LEYP parameters is explored through numerical examples presented in Section 7.

4. Analysis of inter-arrival times

4.1. Definitions

In general, the return period can be defined as the mean of the inter-arrival time (T_n) between any two consecutive events [7]. In Fig. 1, the nth inter-arrival time is defined as, $T_n = S_n - S_{n-1}$ for $n \geq 1$ and $S_0 = 0$, i.e., the time for the nth event to occur after the $(n-1)$th event. Therefore, $\mathbb{E}[T_n]$ denotes the nth mean inter-arrival time, which can also be referred to as the nth return period.

In the HPP, all inter-arrival times are independent and identically distributed exponential random variables with a constant mean. This results in a single value of the return period, i.e., $\mathbb{E}[T_n] = (1/\lambda)$ for all $n \geqslant 1$. Since the HPP is commonly used to model stationary environmental loads, the return period has become a convenient way to define design values in the existing codes. However, this is not the case for a non-stationary process.

Since it is not feasible to keep track of the number of events occurred in the past, the nth return period is not practical to use in the non-stationary climate. Instead, the concept of mean waiting time for the next event is more useful. At any time t, the waiting time, $W(t)$, is the time to occurrence of the next event regardless of the history of the process, as shown in Fig. 2.

Analytical expressions for the mean inter-arrival and mean waiting time are presented in the following subsections.

4.2. Marginal distribution of the arrival time, S_n

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

$$\mathbb{P}[S_n > t] = \overline{F}_{S_n}(t) = \sum_{k=0}^{n-1} \mathbb{P}[N(t) = k] \tag{20}$$

Based on the derivation presented in Appendix B.1, it can be written in terms of the incomplete beta function ratio as

$$\overline{F}_{S_n}(t) = I_{\beta_t}(\alpha, n) = \frac{B_{\beta_t}(\alpha, n)}{B(\alpha, n)} \tag{21}$$

where $B(\alpha, n)$ denotes the complete Beta function defined as

$$B(\alpha, n) = \int_0^1 x^{\alpha-1}(1-x)^{n-1}\mathrm{d}x = \frac{\Gamma(\alpha)\Gamma(n)}{\Gamma(\alpha+n)} \tag{22}$$

Similarly, $B_{\beta_t}(\alpha, n)$ denotes an incomplete beta function given as

$$B_{\beta_t}(\alpha, n) = \int_0^{\beta_t} x^{\alpha-1}(1-x)^{n-1}\mathrm{d}x \tag{23}$$

Based on the derivation presented in Appendix B.2, the probability density function (PDF) of S_n can be written as

$$f_{S_n}(t) = \frac{a\lambda(t)}{B(\alpha, n)}(\beta_t)^{\alpha}(1-\beta_t)^{n-1} \tag{24}$$

4.3. Distribution of the inter-arrival time, T_n

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

$$\mathbb{P}[T_n > u] = \overline{F}_{T_n}(u) = \mathbb{P}[N(S_{n-1}, S_{n-1}+u) = 0]$$
$$= \int_0^{\infty} \mathbb{P}[N(s, s+u) = 0 | S_{n-1} = s] f_{S_{n-1}}(s)\mathrm{d}s \tag{25}$$

The case of $n = 1$, i.e., the time of first arrival, is a special case for which

$$\overline{F}_{T_1}(u) = \mathbb{P}[N(S_0, S_0+u) = 0] \tag{26}$$

Since $S_0 = 0$, this probability can be directly obtained from Eq. (9) as

$$\overline{F}_{T_1}(u) = \mathbb{P}[N_u = 0] = (\beta_u)^{\alpha} \tag{27}$$

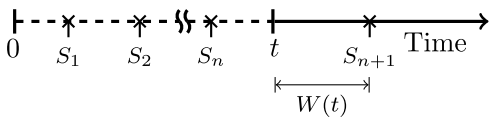

Fig. 2. Definition of the waiting time to the next event at time t.

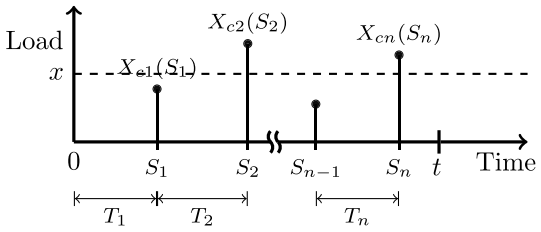

Fig. 3. The LEYP process of time-dependent loads.

For $n \geq 2$, the conditional probability term inside the integration in Eq. (25) can be simplified by noting that events, $[S_{n-1} = s] \equiv [N(0, s) = n-1]$, i.e.,

$$\mathbb{P}[N(s, s+u) = 0 | S_{n-1} = s] = \mathbb{P}[N(s, s+u) = 0 | N(0, s) = n-1]$$

Using the restarting property of the LEYP, Eq. (12), it can be concluded that

$$\mathbb{P}[N(s, s+u) = 0 | N(0, s) = n-1] = (\beta_{s,s+u})^{\alpha+n-1} \tag{28}$$

Substituting this result and the expression for $f_{S_{n-1}}(s)$ from Eq. (24) into Eq. (25) leads to

$$\overline{F}_{T_n}(u) = \frac{a}{B(\alpha, n-1)} \int_0^{\infty} (\beta_{s,s+u})^{\alpha+n-1} \lambda(s) (\beta_s)^{\alpha} (1-\beta_s)^{n-2}\mathrm{d}s \tag{29}$$

Noting that $(\beta_{s,s+u})^{\alpha}(\beta_s)^{\alpha} = (\beta_{s+u})^{\alpha}$, the terms inside the integral can be simplified to

$$\overline{F}_{T_n}(u) = \frac{a}{B(\alpha, n-1)} \int_0^{\infty} (\beta_{s+u})^{\alpha} (\beta_{s,s+u})^{n-1} (1-\beta_s)^{n-2} \lambda(s)\,\mathrm{d}s \tag{30}$$

The expected value of the nth inter-arrival time, also referred to the nth return period, can be computed by integrating the complementary CDF as,

$$\mathbb{E}[T_n] = \int_0^{\infty} \overline{F}_{T_n}(u)\mathrm{d}u \tag{31}$$

4.4. Waiting time for the next event

At a given time, t, the waiting time, $W(t)$, for the occurrence of the next event is defined as, $W(t) = S_{N(t)+1} - t$, as shown in Fig. 2. The distribution of $W(t)$ can be obtained from the equivalence:

$$\mathbb{P}[W(t) > w] = \overline{F}_W(w; t) = \mathbb{P}[N(t, t+w) = 0]$$

The probability term in the RHS can be evaluated from the marginal distribution of $N(s, t)$ given by Eq. (14) with $n = 0$, which leads to

$$\overline{F}_W(w; t) = (\beta^*_{t,t+w})^{\alpha} = \left(\frac{\beta_{t+w}}{1 + \beta_{t+w} - \beta_{t,t+w}}\right)^{\alpha} \tag{32}$$

The mean waiting time can then be computed by integrating the complementary CDF function, as given by Eq. (31).

5. Distribution of maximum load and reliability analysis

5.1. Proposed non-stationary model

In Fig. 3, the LEYP shock process is presented as the model of a sequence of loads occurring over the service life of a structure.

The non-homogeneous LEYP with the conditional intensity function, $\lambda_C(n,t) = h(n)\lambda(t)$, and time-dependent loads incorporates the following non-stationary effects of climate change:

(1) Frequency: The frequency of loading events is modelled through the time-dependent rate, $\lambda(t)$, which captures the non-stationary changes in the frequency of loading events.

(2) Dependence: The history function, $h(n) = (an + b)$, models the statistical dependence among the loading events, which can be quantified by the time-dependent correlation coefficient given by Eq. (19).

(3) Intensity: The non-stationary intensity is modelled by expressing the load magnitude, $X_{ck}(S_k)$, as a time-dependent function,

$$X_{ck}(s_k) = \phi_1(s_k) + \phi_2(s_k)X_k\,, \quad (k = 1, 2, \ldots) \tag{33}$$

It is important to note that the load depends on its time of arrival ($S_K = s_k$) to account for the prevailing conditions, such as temperature and concentration of greenhouse gases in the atmosphere.

Here, $\phi_1(s_k)$ and $\phi_2(s_k)$ are real positive (at least for ϕ_2) additive and multiplicative functions of time, respectively, which can take any general form. Further, X_i's are i.i.d. random variables with a common CDF, $F_X(x)$, and independent of the arrival process. By choosing different forms of $\phi_1(s_k)$ and $\phi_2(s_k)$, a wide variety of time-dependent loads can be modelled.

The CDF of the time-dependent load can be written as

$$\mathbb{P}\left[X_{ck}(s) \leqslant x\right] = \mathbb{P}\left[\phi_1(s) + \phi_2(s)X_k \leqslant x\right] = F_X\left(\psi(x,s)\right)$$
$$\text{where} \quad \psi(x,s) = \frac{x - \phi_1(s)}{\phi_2(s)} \tag{34}$$

The distribution of maximum load, $X_{max}(t_1,t_2)$, generated by the LEYP shock process over a specified time interval, $(t_1,t_2]$, is a key input to the structural reliability assessment. It is defined as [7]:

$$F_{max}(x,t_1,t_2) = \begin{cases} \mathbb{P}\left[X_{cN_1+1}(S_{N_1+1}) \leqslant x, \ldots, X_{cN_2}(S_{N_2}) \leqslant x\right], & \text{if } N_{12} \geqslant 1, \\ 0 & \text{if } N_{12} = 0 \end{cases} \tag{35}$$

where $N_i = N(t_i), i = 1,2$ and $N_{12} = N_2 - N_1$. Based on the derivation given in Appendix C, the final expression for this distribution is given as

$$F_{max}(x,t_1,t_2) = \left[1 + \frac{\mu_{12}}{\alpha} - aq^*(x,t_1,t_2)\right]^{-\alpha} \tag{36}$$

where

$$q^*(x,t_1,t_2) = \int_{t_1}^{t_2} \frac{\lambda(s)}{\beta_s} F_X(\psi(x,s))\mathrm{d}s$$
$$\text{with} \quad \psi(x,s) = \frac{x - \phi_1(s)}{\phi_2(s)}, \quad \beta_s = e^{-a\Lambda(s)}, \quad \text{and} \quad \Lambda(s) = \int_0^s \lambda(u)\,\mathrm{d}u \tag{37}$$

5.2. Reliability analysis

The time-dependent probability of failure in an interval, $(t_1,t_2]$, can be defined as $P_f(t_1,t_2) = \mathbb{P}\left[R - X_{max}(t_1,t_2) \leqslant 0\right]$, where R denotes the strength of a structural component [7]. It can be evaluated as,

$$P_f(t_1,t_2) = \int_0^\infty \overline{F}_{max}(x,t_1,t_2) f_R(x)\mathrm{d}x \tag{38}$$

The Monte Carlo integration method is used in the computation. A sample of strength values, $x_1, \ldots, x_n$, is simulated from the distribution, $f_R(x)$. For a given x_i, $\overline{F}_{max}(x_i,t_1,t_2)$ is computed using Eq. (36). Finally, the failure probability is estimated as

$$P_f(t_1,t_2) \approx \frac{1}{n}\sum_{i=1}^{n} \overline{F}_{max}(x_i,t_1,t_2) \tag{39}$$

Other efficient algorithms can be developed for solving more complex reliability problems, such the importance sampling method.

6. Analytical solutions

In the previous Sections, general expressions have been derived for several variables, such as the return period or mean inter-arrival time, correlation coefficient, and maximum load distribution. For a non-homogeneous LEYP, these expressions involve complex time integrals, which can be solved only by numerical integration. Therefore, it is desirable to have closed-form analytical results for these quantities to gain an intuitive understanding of the impact of different model parameters. With this motivation, several analytical results are presented for a homogeneous birth process (HBP) with a constant rate, $\lambda(t) = \lambda$, such that $\Lambda(t) = \lambda t$, and $\beta_t = e^{-a\lambda t}$.

6.1. nth inter-arrival time

For the HBP, inter-arrival times are independent but non-identical exponential random variables. To prove this result, the following substitutions are made in the complementary CDF (CCDF) of the nth inter-arrival time defined by Eq. (29),

$$\beta_s = e^{-a\lambda s}, \beta_u = e^{-a\lambda u}, \beta_{s+u} = e^{-a\lambda(s+u)} = \beta_s\beta_u, \quad \text{and} \quad \beta_{s,s+u} = \frac{\beta_{s+u}}{\beta_s} = \beta_u, \tag{40}$$

which leads to

$$\overline{F}_{T_n}(u) = \frac{a(\beta_u)^{\alpha+n-1}}{B(\alpha,n-1)} \int_0^\infty (\beta_s)^\alpha (1-\beta_s)^{n-2}\lambda\mathrm{d}s \tag{41}$$

Recognizing that $\mathrm{d}\beta_s = -a\lambda\beta_s\mathrm{d}s$, the above integral term can be transformed in terms of β_s as

$$\int_0^\infty (\beta_s)^\alpha (1-\beta_s)^{n-2}\,\mathrm{d}s = \frac{1}{a\lambda}\int_0^1 \beta_s^{\alpha-1}(1-\beta_s)^{n-2}\,\mathrm{d}\beta_s = \frac{1}{a\lambda}B(\alpha,n-1)$$

Substituting this result in Eq. (41) leads to the following result,

$$\overline{F}_{T_n}(u) = (\beta_u)^{\alpha+n-1} = e^{-a\lambda(\alpha+n-1)u} \tag{42}$$

This is the CCDF of an exponential distribution with parameter, $\lambda_n = a\lambda(\alpha+n-1)$. Therefore, the mean of an nth-inter arrival time (or return period) of this distribution is equivalent to

$$\mathbb{E}\left[T_n\right] = \frac{1}{\lambda_n} = \frac{1}{a\lambda(\alpha+n-1)} = \frac{\mathbb{E}\left[T_s\right]}{a(\alpha+n-1)}, \quad (n \geqslant 1) \tag{43}$$

where, $\mathbb{E}\left[T_s\right] = 1/\lambda$, is the constant return period in the HPP representing a stationary climate process. It is clear from this expression that all mean inter-arrival times are unequal and they will decrease with increasing n.

6.2. Mean waiting time

The mean waiting time can be evaluated using Eq. (32) as

$$\mathbb{E}\left[W(t)\right] = \int_0^\infty \overline{F}_W(w;t)\,\mathrm{d}w = \int_0^\infty \left(\frac{\beta_{t+w}}{1+\beta_{t+w}-\beta_{t,t+w}}\right)^\alpha \mathrm{d}w \tag{44}$$

Using relations given by Eq. (40), the above equation can be rewritten as

$$\mathbb{E}\left[W(t)\right] = (\beta_t)^\alpha \int_0^\infty \left(\frac{\beta_w}{1+(\beta_t-1)\beta_w}\right)^\alpha \mathrm{d}w \tag{45}$$

For an integer value of α, this integral can be solved by substituting $x = \beta_w = e^{-a\lambda w}$ and $\mathrm{d}x = -a\lambda\beta_w\mathrm{d}w$, which lead to

$$\mathbb{E}\left[W(t)\right] = \frac{(\beta_t)^\alpha}{a\lambda}\int_0^1 \frac{x^{\alpha-1}}{\left(1+(\beta_t-1)x\right)^\alpha}\,\mathrm{d}x \tag{46}$$

A solution of this form of integral can be written as [15, p.69, formula 2.111(3)]:

$$\int \frac{x^{m-1}\,\mathrm{d}x}{(c_1+c_2x)^m} = \frac{-x^{m-1}}{c_2(m-1)(c_1+c_2x)^{m-1}} + \frac{1}{c_2}\int \frac{x^{m-2}\,\mathrm{d}x}{(c_1+c_2x)^{m-1}} \tag{47}$$

Using this solution, an explicit result is obtained for $\alpha = 2$ as

$$\mathbb{E}[W(t)] = \left(\frac{-t(\beta_t)^2}{(\beta_t - 1)^2} - \frac{\beta_t}{a\lambda(\beta_t - 1)} \right) \tag{48}$$

Results for other integer values α can be similarly derived.

6.3. Correlation coefficient

The correlation coefficient given by Eq. (19) can be analytically evaluated for the HBP.

Consider the correlation between the number of events in two adjacent intervals, $t \pm \delta$, i.e., $t_1 = t - \delta, t_2 = t$, and $t_3 = t + \delta$, $\delta > 0$. Since in this case, $\beta_{12} = \beta_{23} = e^{-a\lambda\delta}$, the term, β^*_{12}, in Eq. (19) can be evaluated as

$$\beta^*_{12} = \frac{e^{-a\lambda t}}{1 + e^{-a\lambda t} - e^{-a\lambda\delta}}$$

Using this relation, an explicit expression for the correlation coefficient is obtained as

$$\left[\rho\left(N_{12}, N_{23}\right)\right]^2 = \frac{1 - e^{-a\lambda\delta}}{1 + e^{-a\lambda t} - e^{-a\lambda\delta}} \times \frac{1 - e^{-a\lambda\delta}}{1 + e^{-a\lambda(t+\delta)} - e^{-a\lambda\delta}} \tag{49}$$

It is interesting to note that the correlation coefficient depends on both the length of the interval, δ, and its location, t, in time.

Since $a = 0$ for the HPP, Eq. (49) would lead to $\rho\left(N_{12}, N_{23}\right) = 0$, in accordance with the independence property of the Poisson process.

6.4. Maximum load distribution

An analytical result for the maximum load distribution, given by Eq. (36), can be derived assuming that the load magnitude is independent of time, i.e., $\phi_1(s_k) = 0$ and $\phi_2(s_k) = 1$, such that $X_{ck}(s_k) = X_k$ and $\psi(x, s) = x$. Using these results and Eq. (16) for μ_{12}/α, the integral term, $q^*(x, t_1, t_2)$, can be rewritten as

$$\begin{aligned} q^*(x, t_1, t_2) &= \int_{t_1}^{t_2} \lambda e^{-a\lambda s} F_X(x) \mathrm{d}s = \lambda F_X(x) \frac{1}{a\lambda} \left(\frac{1}{\beta_2} - \frac{1}{\beta_1} \right) \\ &= \frac{1}{a} F_X(x) \frac{\mu_{12}}{\alpha} \end{aligned} \tag{50}$$

Substituting this result in Eq. (36) leads to the following simple expression:

$$F_{max}(x, t_1, t_2) = \left[1 + \frac{\mu_{12}}{\alpha}\left(1 - F_X(x)\right)\right]^{-\alpha} \tag{51}$$

A pth percentile of the maximum load can be computed by inverting the above CDF as,

$$x_p(t_1, t_2) = F_X^{-1}\left[1 - \frac{\alpha}{\mu_{12}}\left(\frac{1}{p^{1/\alpha}} - 1\right)\right] \tag{52}$$

When $\alpha \to \infty$, a limiting case corresponding to the homogeneous Poisson process is obtained as

$$F_{max}(x, t_1, t_2) \to \exp\left(-\mu_{12}\left(1 - F_X(x)\right)\right), \quad (\alpha \to \infty) \tag{53}$$

where $\mu_{12} = \lambda b(t_2 - t_1)$ with a fixed value of b, e.g., $b = 1$.

7. Illustrative examples

7.1. Homogeneous birth process (HBP)

Using the analytical results derived in Section 6, the effect of dependence is investigated on the probabilistic properties of the HBP. In numerical calculation, the following values are chosen: $\lambda = 0.02$ per year corresponding to a 50-year return period in the HPP model which represents a reference case of the stationary climate. For a fixed λ, the correlation coefficient depends only on a and is independent of b, as shown by Eq. (49). Therefore, the parameter b is fixed as $b = 1$ without the loss of generality. The parameter, a, also referred to as the Yule scale parameter, is primarily varied to examine its effect on the correlation coefficient, waiting time, and reliability index.

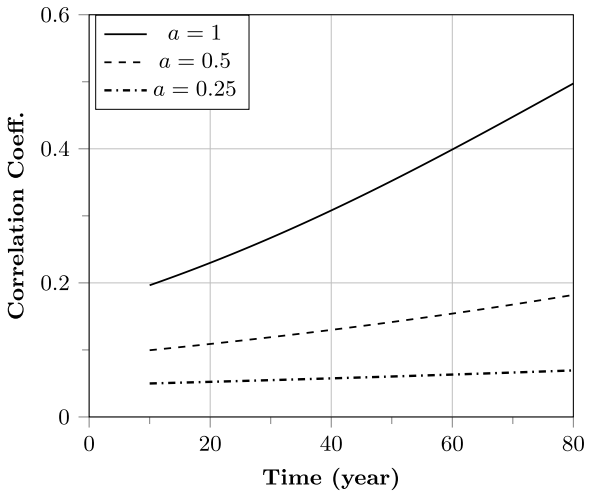

Fig. 4. The decadal correlation coefficient between the number of events (HBP).

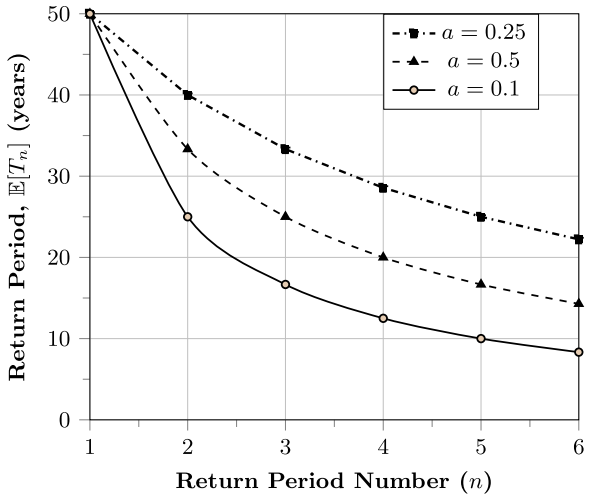

Fig. 5. The trend in the first six return periods (HBP).

The decadal correlation coefficient, $\rho\left[N_{t-10}, N_{t+10}\right]$ is calculated as a function of t using Eq. (49). The three values of a are chosen as 0.25, 0.5, and 1 in order to cover a wide range of the correlation coefficient. Results presented in Fig. 4 show two interesting trends. First, the correlation coefficient increases continuously as t is increased from 10 to 80 years. Second, the rate and the magnitude of this increase are controlled by a.

The first six mean inter-arrival times (or return periods) are calculated using Eq. (43) and plotted in Fig. 5. Note that the mean of the $n$thinter-arrival time is defined as, $\mathbb{E}\left[T_n\right] = \mathbb{E}\left[S_n - S_{n-1}\right]$, $n \geq 1, S_0 = 0$. The mean inter-arrival time decreases rapidly as the number of occurrences, n, increases. For example, for $a = 0.5$, the third return period reduces to half ($\approx$ 25 years) of that in the stationary climate (=50 years). It is also evident that the magnitude of this decrease is highly sensitive to the Yule scale parameter (a).

The mean waiting time for the next event is calculated as a function of time using Eq. (46) and plotted in Fig. 6. For $a = 0.5$, it decreases from 50 to 30 years over an 80-year period indicating an increase in the frequency of events. The mean waiting time also decreases with increasing dependence, as reflected by increasing values of a.

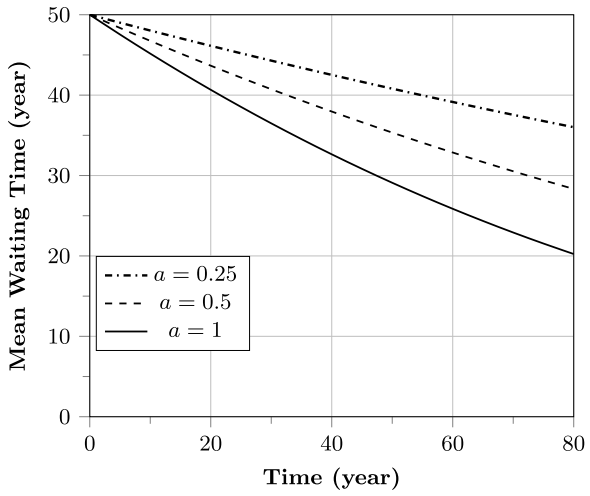

Fig. 6. The mean waiting time for the next event (HBP).

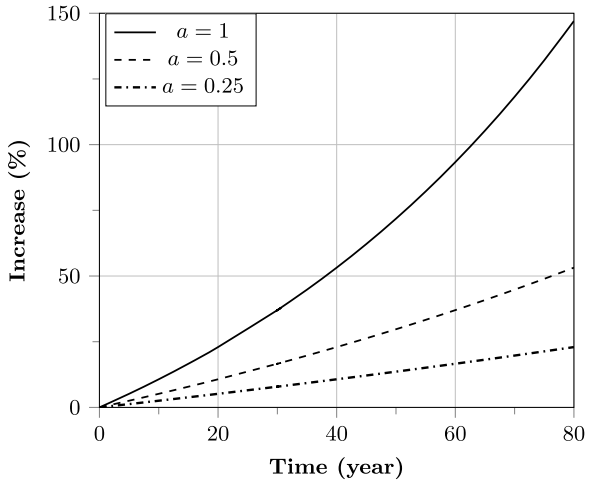

Fig. 7. Increase in the expected number of events, $\mathbb{E}[N_t]$, over time (HBP).

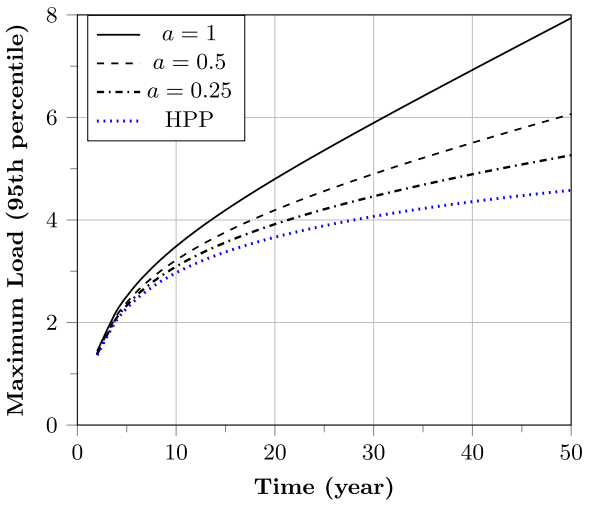

Fig. 8. The 95th percentile of the maximum load vs. time (HBP).

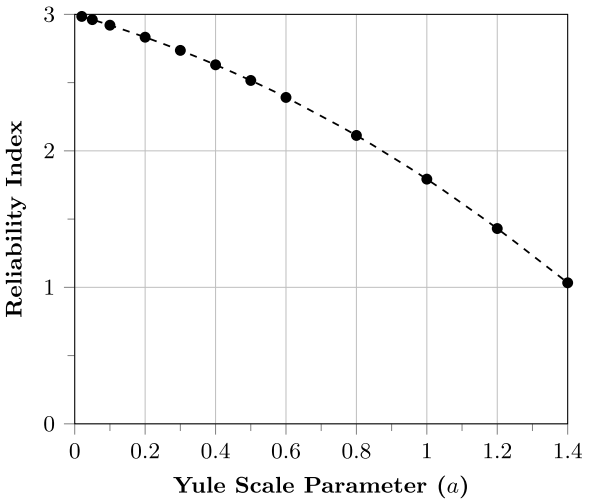

Fig. 9. The 50-year reliability index vs. the Yule scale parameter (HBP).

The expected value of the cumulative number of events in $(0,t]$ are computed as $\mu_t = (1/a)(e^{a\lambda t} - 1)$ and compared with the expected number of events in the stationary climate ($= \lambda t$). Fig. 7 plots the percentage increases in the expected number of events in the HBP as compared to that in the HPP model in the same period, calculated as $(\mu_t - \lambda t)/\lambda t$. It is clear that the expected number of events increases rapidly with an increase in the value of a. The increasing number of events is a reason for a decreasing trend in the inter-arrival time.

Figs. 5–7, illustrate that even a modest change in the correlation coefficient has a significant impact on probabilistic properties of the HBP. For example, when $a = 0.5$, the correlation coefficient increases modestly from 0.1 to 0.2 as t increases from 10 to 80 years. Over this period, the mean number of events increases by over 40% as shown by Fig. 7. In case of $a = 1$, as the correlation increases from 0.2 to 0.5, the mean number of events increases almost by 150%, and the mean waiting time decreases from 50 years to 20 years, as shown by Fig. 6.

Further, the effect of a on the maximum load generated by the HBP is examined by assuming that the load magnitude (X) follows an exponential distribution with a mean, $\mu_X = 1$ unit, and the loading occurs at a rate of $\lambda = 0.1$ events per year. Fig. 8 shows an increase in the 95th percentile of the maximum load with time for different values of a. For a comparison, results are also plotted for the HPP with the same λ and X. It is clear that dependence between the number of events in the HBP would lead to higher extreme loads as compared to the HPP. The higher the correlation coefficient, as implied by increasing values of a, the higher is the load magnitude. For example for $a = 1$, the 95th load percentile almost doubles in 50 years as compared to that of the HPP.

To evaluate the effect of dependence on reliability, consider a base case of the HPP load process as used before ($\lambda = 0.1$ event/year and $X \sim Exp(1)$. The structural strength of a component (R) is assumed to be normally distributed with a COV of 0.10. The mean strength was determined as, $\mu_R = 8.55$, corresponding to a 50-year reliability index of 3, or the probability of failure, $P_f(50) = 1.35\times10^{-3}$, under the stationary climate.

Since the percentile function of the maximum load distribution is given in an explicit form by Eq. (52), the First-Order Reliability Method (FORM) was used to compute the reliability index. Results plotted in Fig. 9 show that the 50-year reliability index, as expected, decreases with increasing values of a.

In summary, results presented in this section demonstrate that the statistical dependence alone has a profound effect on the mean number of events, return/waiting period, extreme percentiles, and the probability of failure. This effect is scaled up by the Yule scale parameter (a) by increasing the correlation coefficient between the number of events.

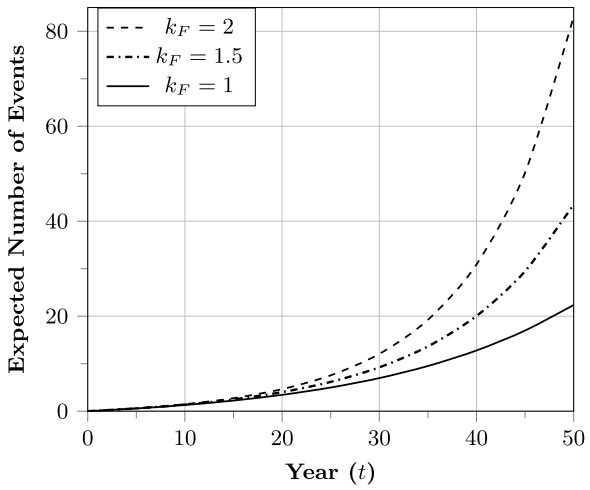

Fig. 10. Expected number of events in the LEYP load model ($a = 0.5$).

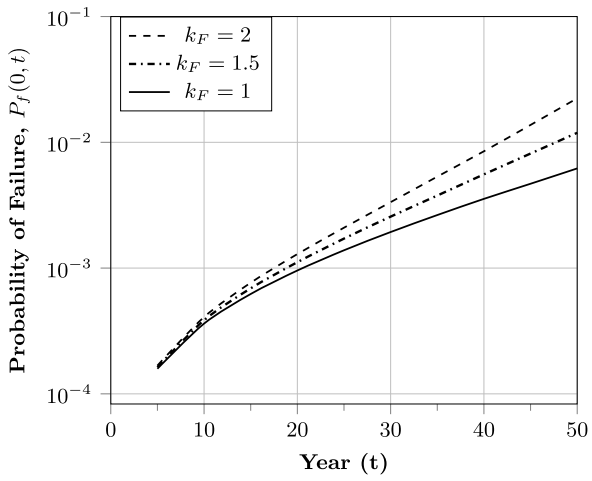

Fig. 11. Effect of non-stationary frequency on the probability of failure ($a = 0.5, k_L = 1$).

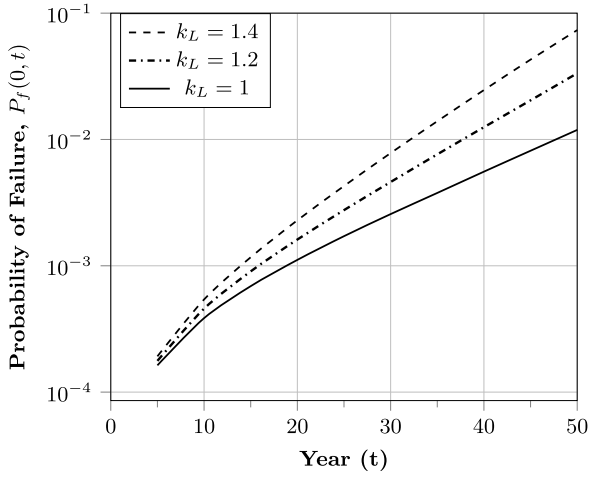

Fig. 12. Effect of non-stationary load intensity on the probability of failure ($a = 0.5, k_F = 1.5$).

These effects will be further amplified in cases where the rate ($= \lambda(t)$) is non-stationary.

7.2. Non-homogeneous LEYP

This Section illustrates the effect of a non-stationary load process on the probability of failure of a structural component. Recall that the conditional intensity of the non-homogeneous LEYP is given as $\lambda_C(n,t) = h(n)\lambda(t)$. In the present example, the occurrence rate of events is assumed to increase linearly from λ_o to $k_F\lambda_o$ over the time horizon of t_e years:

$$\lambda(t) = \lambda_o\left(1 + (k_F - 1)\frac{t}{t_e}\right) \tag{54}$$

The parameters of the LEYP's conditional intensity represent the following three stochastic effects associated with the load arrival process: 5

1. Pure stationary effect represented by a constant base rate, λ_o.
2. Temporal effect of the non-stationary climate represented by $\lambda(t)$, as given by Eq. (54).
3. Dependence effect represented by $h(n) = (an + 1)$, given by Eq. (8).

Similarly, a linearly increasing function is chosen to model the load magnitude occurring at time s_k as,

$$X_{ck}(s_k) = \phi_2(s_k)X_k = \left(1 + (k_L - 1)\frac{s_k}{t_e}\right)X_k, \quad (k = 1, 2, \ldots) \tag{55}$$

A linear variation is assumed in both the occurrence rate and the load magnitude for the sake of simplicity. An advantage of the linear function is that only one factor is required (k_L or k_F) to model the non-stationary effect. Note that the amplification factors, k_F and k_L, denote the order of magnitude of increases in the loading frequency and magnitude, respectively, over their stationary values. Thus, these factors provide a quick and simple way to estimate the order of increase in the probability of failure under some bounding cases of non-stationary changes. For example, if the climate change causes the loading frequency to double over the next 80 years, what will be the order of increase in P_f. Such scenarios are illustrated next through examples. The following parameters are fixed in all the examples presented here, $\lambda_o = 0.1$ events/year, and the time horizon, $t_e = 50$ years (from 2050 to 2100).

The expected number of events over a 50 year period were computed for $a = 0.5$, which represents a modest increase in correlation coefficient (0.1–0.2). Results are plotted in Fig. 10 for three different values of the frequency amplification factor, $k_F = 1, 1.5$ and 2. Results show that a non-stationary increase in the rate leads to a significant increase in the expected number of events, especially in the latter part of the time horizon.

The impact of the non-stationary rate on the probability of failure is presented in Fig. 11. The load and strength data for this example are the same as those used in computing results shown in Fig. 9. The cumulative probability of failure, $P_f(t)$ in a interval, $(0,t]$ was computed using 250,000 simulations in Eq. (39). Fig. 11 shows that the increase in k_F further amplifies the increase in $P_f(t)$ in a more modest manner than the increase in the expected number of events, as shown by Fig. 10.

In addition to the non-stationary frequency (with $k_F = 1.5$), a time-dependent increase in the load magnitude would further increase the probability of failure, as shown by Fig. 12. For example, a 20% increase in the load ($k_L = 1.2$) over 50 years leads to a significant increase in the probability of failure.

In summary, the dependence effect of the LEYP is further amplified by the non-stationary increases in the frequency and magnitude of load events, which causes a marked increase in the probability of failure over time. A non-stationary increase in the load magnitude has a more pronounced effect on P_f than a comparable increase in the frequency of loading events.

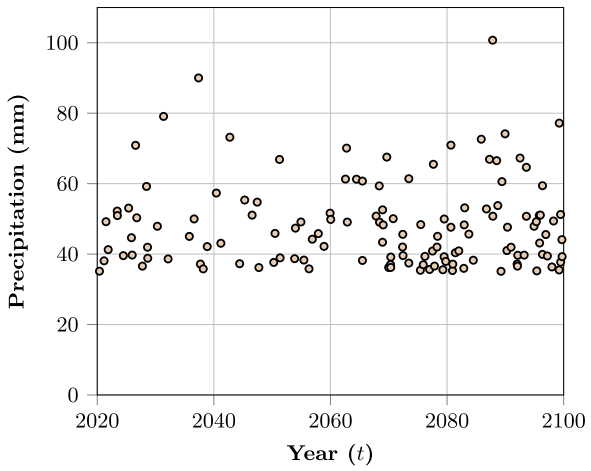

Fig. 13. Heavy daily-precipitation events in the future climate (CanESM5 model, 2020–2100).

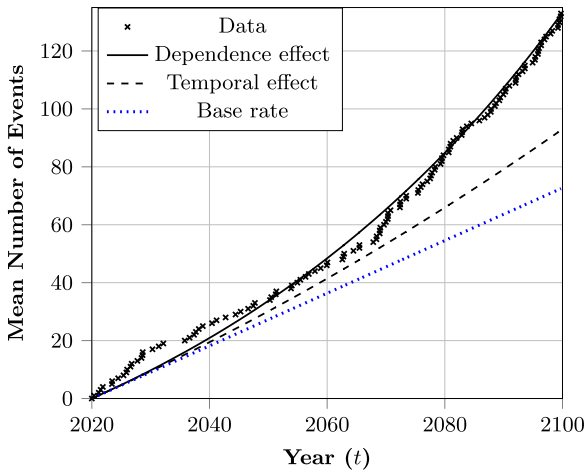

Fig. 14. Impact of different effects included in the LEYP intensity function (CanESM5 data, 2020–2100).

8. Practical example: Extreme precipitation in the future climate

8.1. Data

The current generation of climate models provides a comprehensive representation of various atmospheric systems and processes affecting the climate. These models have played a crucial role in studying the effects of different emission scenarios on future climate conditions, including temperature and precipitation [3]. This Section presents a practical example of statistical fitting of the LEYP model to the precipitation data simulated under a high-emission scenario.

Heavy rainfall events vary significantly from region to region in terms of amounts, spatial and temporal scales, occurrence rates, and seasonality. Atmospheric processes, such as synoptic, convective, and tropical cyclones/hurricanes, are typically linked to extreme rainfall events [16]. Given a time series of simulated precipitation data, it is challenging to determine which specific process is generating the data in a given time interval. Therefore, this analysis adopts a simple approach in which the events with total daily (24-hour) precipitation (rain and snow) exceeding 35 mm are separated and treated as being generated by a stochastic process that is distinct from the day-to-day precipitation of small amounts. This straightforward approach is considered reasonable for illustrating the statistical fitting of the LEYP parameters.

The analysis considers the future daily precipitation data (2020–2100) simulated by the Canadian Earth System Model version 5 (CanESM5.0.3), which is also a part of the phase 6 of the Coupled Model Intercomparison Project (CMIP6) [17]. The dataset was downloaded from https://climatedata.ca for a grid box of $6 \times 10\,\text{km}^2$ size (0.0833° latitude × 0.0833° longitude) in the city of Toronto (43.6532°N, 79.3832°W), Ontario, Canada. The dataset corresponds to the Shared Socioeconomic Pathway, SSP5-8.5 (Fossil-fuelled development) in which the mean global temperature is expected to increase by 4.4 °C by the end of the 21st century. The heavy precipitation data, 133 events over 80 years, are plotted in Fig. 13. These events are taken at least 10 days apart to ensure that multiple extremes from a single storm event are not included in the analysis.

8.2. Statistical analysis and results

The frequency of precipitation events is assumed to follow the linear function given by Eqs. (54). Using the maximum likelihood method, as described by Badía et al. [8], the model parameters were estimated as $\lambda_o = 0.91$ events/year, $k_F = 1.55$, and $a = 0.00734$.

The three stochastic effects included in the frequency function, as discussed below Eq. (54), are illustrated by Fig. 14, which includes a plot of the mean (cumulative) number of events versus the arrival time. A constant base rate of $\lambda_o = 0.91$ implies that 73 events are expected to occur over an 80-year period under stationary conditions. The factor, $k_F = 1.55$, implies an overall 55% increase in the frequency over 80 years, leading to the expected number of 92 events over this period. This is referred to as the "temporal effect" of non-stationary climate. The dependence effect, represented by $a = 0.00734$, is relatively small. To quantify its impact, the decadal correlation coefficient, $\rho\left[N_{t-10}, N_{t+10}\right]$, was computed using Eq. (19). Between 2020–30 and 2030–40 decades, the correlation coefficient was computed as 0.07, which increased to 0.16 for the last two decades of the time horizon (2060–70 and 2070–80). Even this modest correlation has a discernible effect, as shown by an increase in the expected number of events to 132 from 92 over 80 years. Finally, Fig. 14 confirms that the LEYP model including all the three effects provides a high degree of goodness-of-fit to precipitation data.

The precipitation amount was modelled as an exponential random variable with a linear increase given by Eq. (55) as,

$$\mathbb{P}\left[X_c(s) < x\right] = 1 - \exp\left(\frac{-\lambda_X(x - x_0)}{\phi_2(s)}\right) \tag{56}$$

where $x_0 = 35$ mm and $\phi_2(s) = 1 + (k_L - 1)s/80$. Using the maximum likelihood method, the model parameters were estimated as $\lambda_X = 0.0767$/mm and $k_L = 1.009$. Since $k_L \approx 1$, the precipitation magnitude is modelled as a time-invariant random variable.

Intense short-term precipitation events can cause flash flooding and damage to various infrastructure systems. Suppose the design of a water infrastructure system requires the 95th percentile of the maximum precipitation over a 50-year service life (2050–2100). Using Eq. (36), this value is estimated as 134 mm.

To determine a comparable value in the stationary climate, an additional analysis of climate data simulated for a preindustrial period (e.g., 1950–1980), is required [3]. This task is however beyond the scope of the current study.

8.3. Remarks

Some aspects of the example of precipitation data analysis are discussed as follows.

The previous precipitation example is based on a single dataset taken from a large repository of climate data, which may not fully represent the qualitative and quantitative effects of climate change.

For instance, a low correlation coefficient in this data does not necessarily imply that the dependence effect is completely absent from non-stationary climate conditions.

An evaluation of the climate change effect in a specific region requires a more comprehensive analysis. Typically, data generated by several climate models (more than 20 models) are analysed, and the results are averaged over multiple grid boxes to minimize the effect of internal variability [3]. Finally, the multi-model average or median value is used to assess the overall impact of the climate change. To benchmark changes in the future precipitation with reference to the stationary climate, a similar detailed analysis data simulated for a preindustrial period, such as 1950–1980, is required [3]. Such detailed statistical analyses are beyond the scope of this paper.

To conclude, the present example is only an illustration of the statistical fitting of the LEYP model. An additional utility of this model is that it provides a more scientific basis to evaluate the degree of dependence in any given climate dataset.

9. Conclusions

The climate change effects are responsible for the increasing frequency and intensity of weather extremes in recent times. Simulations of climate conditions under high emission scenarios also confirm the intensification of weather extremes in the future. To ensure the safety of existing and to-be-built infrastructure systems under the changing climate, the probabilistic modelling of climate extremes is emerging as an important area of research.

The paper presents a general stochastic process model to incorporate the non-stationary nature of extreme weather events at the three levels: frequency of occurrence of events, intensity or magnitude of events, and statistical dependence among them. The paper proposes a class of non-stationary birth processes, namely, the Linear Extension of the Yule Process (LEYP), for this purpose. Although the mathematical details of this model have been discussed in the literature, the paper presents several new results to support the time-dependent reliability analysis of infrastructure systems, such as

- Derivation of the correlation coefficient between the number of events in two adjacent time intervals as a measure of dependence in the LEYP model.
- Accurate formulae for computing the nth return period, or the mean inter-arrival time between the $(n-1)$ and nth events.
- Evaluation of the mean waiting time for the next event, which is a more appropriate measure for structural design under non-stationary conditions.
- Derivation of the probability distribution of the maximum value generated by an LEYP in a given time interval. In contrast with traditional methods of annual maxima and Peaks-over-Threshold, the proposed approach does not invoke any asymptotic argument.
- In the case of the homogeneous birth process (HBP), the derivation of explicit formulae for the correlation coefficient, nth return period, and maximum value distribution.
- An exposition of the reliability analysis under non-stationary load processes.

The paper presents several numerical examples including both the homogeneous and non-homogeneous LEYP. Unlike in a stationary climate, the mean inter-arrival time between events varies in the changing climate. These examples illustrate that increasing dependence has a significant impact on the mean inter-arrival time and waiting time for the next event. Even a modest degree of dependence among the events results in a considerable decrease in both of these quantities. The mean waiting time for the next event is proposed as a more practical measure for designing future infrastructure systems, since it does not rely on the process history. The statistical fitting of the LEYP model is demonstrated using a precipitation dataset simulated by the Canadian Earth Systems Model (CanESM5) under a high-emission scenario (SSP5-8.5).

The LEYP model overcomes a major limitation of the classical Poisson process by including the statistical dependence among extreme events. The correlation coefficient between the number of events in two adjacent intervals also increases with time, and the rate of increase is controlled by the Yule scale parameter (a). The upper tail of the extreme load distribution increases, and as a result, the probability of failure increases with the increasing degree of dependence. This effect is further amplified by non-stationary increases in the frequency and intensity of load events. Therefore, ignoring the dependence among extreme events can result in a significant underestimation of the probability of failure, which will have an adverse impact on the safety of infrastructure systems in the changing climate.

In the current literature, the non-stationary generalized extreme value (GEV) distribution is utilized as a high-level approach to model the annual maxima data. This model cannot account for the dependence effect, and is unable to account for individual changes in the frequency and intensity of the process. In contrast, the proposed approach explicitly models all non-stationary effects (frequency, intensity, and dependence) as basic components of the LEYP model, which are then aggregated into the distribution of maximum value. In this regard, the paper introduces a more structured approach to the probabilistic modelling of extremes in the changing climate.

CRediT authorship contribution statement

Mahesh D. Pandey: Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Sophie Mercier:** Writing – review & editing, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors gratefully acknowledge the financial support for this study provided by the Natural Sciences and Engineering Research Council of Canada (NSERC).

Appendix A. Derivation of the correlation coefficient

A.1. Approach

The statistical correlation coefficient between N_{12} and N_{23} is defined as

$$\rho[N_{12}, N_{23}] = \frac{\mathrm{COV}\left[N_{12}, N_{23}\right]}{\sigma_{12}\,\sigma_{23}} \tag{A.1}$$

The covariance function in the numerator has the standard definition:

$$\mathrm{COV}\left[N_{12}, N_{23}\right] = \mathbb{E}\left[(N_{12}-\mu_{12})(N_{23}-\mu_{23})\right] = \mathbb{E}\left[N_{12}N_{23}\right] - \mu_{12}\,\mu_{23} \tag{A.2}$$

Thus, the main task is to obtain an expression for the product moment, $\mathbb{E}\left[N_{12}N_{23}\right]$, which is derived in Appendix A.2, and the final result is as follows:

$$\mathbb{E}\left[N_{12}N_{23}\right] = \frac{\alpha+1}{\alpha}\mu_{12}\mu_{23} \tag{A.3}$$

Substituting this relation in Eq. (A.2) leads to

$$\mathrm{COV}\left[N_{12}, N_{23}\right] = \frac{1}{\alpha}\mu_{12}\mu_{23} \tag{A.4}$$

Recall that the variance is given by Eq. (17) as, $\sigma_{st}^2 = \mu_{st}/\beta_{st}^*$, and Eq. (16) gives β_{st}^* as

$$\beta_{st}^* = \frac{\alpha}{\mu_{st}+\alpha}$$

Therefore, the variance can be expressed in terms of the mean value and α as

$$\sigma_{st}^2 = \frac{1}{\alpha}\mu_{st}(\alpha+\mu_{st})$$

Substituting relations for the covariance and variance in Eq. (A.1) leads to the following expression:

$$\left(\rho[N_{12},N_{23}]\right)^2 = \frac{\mu_{12}}{\alpha+\mu_{12}}\,\frac{\mu_{23}}{\alpha+\mu_{23}} \tag{A.5}$$

Using Eq. (16), this relation can also be written in terms of β_{st}^* as

$$\left(\rho[N_{12},N_{23}]\right)^2 = (1-\beta_{12}^*)(1-\beta_{23}^*) \tag{A.6}$$

This relation shows that the correlation coefficient does not depend on the parameter b of the LEYP.

A.2. Derivation of the product moment, $\mathbb{E}\left[N_{12}N_{23}\right]$

The product moment of the number of events in the two consecutive time intervals is derived using a fundamental property of the expectation,

$$\mathbb{E}\left[N_{12}N_{23}\right] = \mathbb{E}_{N_{12}}\left[\mathbb{E}\left[N_{12}N_{23}|N_{12}\right]\right] \tag{A.7}$$

this can be simplified as

$$\mathbb{E}\left[N_{12}N_{23}\right] = \sum_{m=0}^{\infty} m\mathbb{E}\left[N_{23}|N_{12}=m\right]\,\mathbb{P}\left[N_{12}=m\right] \tag{A.8}$$

The conditional expectation, $\mathbb{E}\left[N_{23}|N_{12}=m\right]$, is derived in the next Appendix A.3. Substituting the final expression given by Eq. (A.14) in the above equation leads to

$$\begin{aligned}\mathbb{E}\left[N_{12}N_{23}\right] &= \frac{\mu_{23}}{\alpha}\beta_{12}^*\sum_{m=0}^{\infty} m(\alpha+m)\mathbb{P}\left[N_{12}=m\right]\\ &= \frac{\mu_{23}}{\alpha}\beta_{12}^*\Big(\sum_{m=0}^{\infty}\alpha m\mathbb{P}\left[N_{12}=m\right]+\sum_{m=0}^{\infty}m^2\mathbb{P}\left[N_{12}=m\right]\Big)\\ &= \frac{\mu_{23}}{\alpha}\beta_{12}^*\left(\alpha\mathbb{E}\left[N_{12}\right]+\mathbb{E}\left[N_{12}^2\right]\right)\end{aligned} \tag{A.9}$$

Since $N_{12} \overset{d}{=} \mathcal{NB}(\alpha,\beta_{12}^*)$, its mean-square is obtained from Eq. (7) as

$$\mathbb{E}\left[N_{12}^2\right] = \mu_{12}^2 + \frac{\mu_{12}}{\beta_{12}^*}$$

Substituting this relation in Eq. (A.9) leads to

$$\mathbb{E}\left[N_{12}N_{23}\right] = \frac{\mu_{23}}{\alpha}\beta_{12}^*\left(\alpha\mu_{12}+\mu_{12}^2+\frac{\mu_{12}}{\beta_{12}^*}\right) \tag{A.10}$$

Noting that Eq. (16) implies that $\beta_{12}^*\left(\mu_{12}+\alpha\right)=\alpha$, the final result is obtained:

$$\mathbb{E}\left[N_{12}N_{23}\right] = \frac{\alpha+1}{\alpha}\mu_{12}\mu_{23} \tag{A.11}$$

A.3. Derivation of the conditional expectation, $\mathbb{E}\left[N_{23}|N_{12}=m\right]$

The conditional random variable, $N_{23}|(N_{12}=m)$, follows a negative binomial distribution, as shown by [18, Sec.4.2.5, p.38], i.e., $[N_{23}|N_{12}=m] \overset{d}{=} \mathcal{NB}(\alpha+m,\beta_{123}^*)$ where

$$\beta_{123}^* = \frac{\beta_{23}+\beta_3-\beta_{13}}{1-\beta_{13}+\beta_3} \tag{A.12}$$

The expected value of this distribution is obtained using Eq. (7) as

$$\begin{aligned}\mathbb{E}\left[N_{23}|N_{12}=m\right] &= (\alpha+m)\frac{1-\beta_{123}^*}{\beta_{123}^*} = (\alpha+m)\frac{1-\beta_{23}}{\beta_{23}+\beta_3-\beta 13}\\ &= (\alpha+m)\frac{1-\beta_{23}}{\beta_3}\,\frac{\beta_3}{\beta_{23}+\beta_3-\beta_{13}}\end{aligned} \tag{A.13}$$

Using the relation $\beta_{st}=\beta_t/\beta_s$ for $s\leqslant t$, the following simplification is achieved

$$\frac{\beta_3}{\beta_{23}+\beta_3-\beta_{13}} = \frac{\beta_2}{1+\beta_2-\beta_{12}} = \beta_{12}^*$$

It can be shown using Eq. (16) that

$$\frac{1-\beta_{23}}{\beta_3} = \frac{\mu_{23}}{\alpha}$$

Using these simplifications, the final expression is obtained as

$$\mathbb{E}\left[N_{23}|N_{12}=m\right] = (\alpha+m)\frac{\mu_{23}}{\alpha}\beta_{12}^* \tag{A.14}$$

Appendix B. Marginal distribution of the nth arrival time, S_n

B.1. Cumulative distribution of S_n

The CDF of S_n can be written using Eq. (20) as

$$F_{S_n}(t) = 1-\sum_{k=0}^{n-1}\mathbb{P}\left[N(t)=k\right] = \sum_{k=n}^{\infty}\mathbb{P}\left[N(t)=k\right] \tag{B.1}$$

Substituting for $\mathbb{P}\left[N(t)=k\right]$ from Eq. (9) leads to

$$F_{S_n}(t) = \sum_{k=n}^{\infty}\frac{\Gamma(\alpha+k)}{\Gamma(\alpha)\Gamma(k+1)}(\beta_t)^{\alpha}\,(1-\beta_t)^k \tag{B.2}$$

Next, $F_{S_n}(t)$ is differentiated with respect to β_t, which leads to

$$\frac{\mathrm{d}F_{S_n}(t)}{\mathrm{d}\beta_t} = \sum_{k=n}^{\infty}\frac{\Gamma(\alpha+k)}{\Gamma(\alpha)\Gamma(k+1)}\left(\alpha\beta_t^{\alpha-1}(1-\beta_t)^k - k(\beta_t)^{\alpha}(1-\beta_t)^{k-1}\right) \tag{B.3}$$

Now add and subtract a term, $k\beta_t^{\alpha-1}(1-\beta_t)^k$, as suggested by Rider [19], and collect the common terms,

$$\frac{\mathrm{d}F_{S_n}(t)}{\mathrm{d}\beta_t} = \frac{\beta_t^{\alpha-1}}{1-\beta_t}\sum_{k=n}^{\infty}\frac{\Gamma(\alpha+k)}{\Gamma(\alpha)\Gamma(k+1)}\left((\alpha+k)(1-\beta_t)^{k+1}-k(1-\beta_t)^k\right) \tag{B.4}$$

To simplify the above equation, the following relations are used

$$(\alpha+k)\Gamma(\alpha+k) = \Gamma(\alpha+k+1), \quad \text{and} \quad \frac{k}{\Gamma(k+1)} = \frac{1}{\Gamma(k)},$$

Substituting these results in (B.4) leads to

$$\frac{\mathrm{d}F_{S_n}(t)}{\mathrm{d}\beta_t}\,\frac{1-\beta_t}{\beta_t^{\alpha-1}}\,\Gamma(\alpha) = \sum_{k=n}^{\infty}\frac{\Gamma(\alpha+k+1)}{\Gamma(k+1)}(1-\beta_t)^{k+1} - \sum_{k=n}^{\infty}\frac{\Gamma(\alpha+k)}{\Gamma(k)}(1-\beta_t)^k \tag{B.5}$$

A typical term in the series can be denoted as

$$d_k = \frac{\Gamma(\alpha+k)}{\Gamma(k)}(1-\beta_t)^k,$$

Thus, Eq. (B.5) involves a difference of sums of $(k+1)$th and kth terms, given as

$$\frac{\mathrm{d}F_{S_n}(t)}{\mathrm{d}\beta_t}\,\frac{(1-\beta_t)}{\beta_t^{\alpha-1}}\,\Gamma(\alpha) = \sum_{k=n+1}^{\infty}d_k - \sum_{k=n}^{\infty}d_k \tag{B.6}$$

Here, all the terms cancel out except the term $-d_n$, which leads to

$$\frac{\mathrm{d}F_{S_n}(t)}{\mathrm{d}\beta_t}\,\frac{(1-\beta_t)}{\beta_t^{\alpha-1}}\,\Gamma(\alpha) = -d_n = \frac{-\Gamma(\alpha+n)}{\Gamma(n)}(1-\beta_t)^n$$

Using the definition of the Beta function given by Eq. (22), the above equation is rewritten as

$$\frac{-\mathrm{d}F_{S_n}(t)}{\mathrm{d}\beta_t} = \frac{\beta_t^{\alpha-1}(1-\beta_t)^{n-1}}{B(\alpha,n)} \tag{B.7}$$

Next, the incomplete beta function ratio, given in the RHS of Eq. (21), is differentiated with respect to β_t,

$$\frac{\mathrm{d}}{\mathrm{d}\beta_t}\left(I_{\beta_t}(\alpha,n)\right) = \frac{1}{B(\alpha,n)}\,\frac{\mathrm{d}}{\mathrm{d}\beta_t}\left(\int_0^{\beta_t}x^{\alpha-1}(1-x)^{n-1}\mathrm{d}x\right)$$

Since β_t is included in the limit of the integral, the Leibnitz's rule of differentiating an integral is used to obtain,

$$\frac{\mathrm{d}}{\mathrm{d}\beta_t}\left(I_{\beta_t}(\alpha,n)\right) = \frac{\beta_t^{\alpha-1}(1-\beta_t)^{n-1}}{B(\alpha,n)} \tag{B.8}$$

Comparing Eqs. (B.7) and (B.8), it is concluded that

$$\frac{-\mathrm{d}F_{S_n}(t)}{\mathrm{d}\beta_t} = \frac{\mathrm{d}\overline{F}_{S_n}(t)}{\mathrm{d}\beta_t} = \frac{\mathrm{d}}{\mathrm{d}\beta_t}\left(I_{\beta_t}(\alpha,n)\right) \tag{B.9}$$

By integrating both sides w.r.t. β_t leads to

$$\overline{F}_{S_n}(t) = I_{\beta_t}(\alpha,n) + C$$

where C is a constant of integration. In the limit as $t \to \infty$, $\beta_t \to 0$, it results in $\overline{F}_{S_n}(t) = 0$ and $I_0(\alpha,n) = 0$. Therefore, $C = 0$, which leads to the final result:

$$\overline{F}_{S_n}(t) = I_{\beta_t}(\alpha,n) \tag{B.10}$$

B.2. Probability density function of S_n

The PDF of S_n is derived by differentiating the CDF, $F_{S_n}(t)$, w.r.t. time and using the chain rule as follows:

$$f_{S_n}(t) = \frac{\mathrm{d}F_{S_n}(t)}{\mathrm{d}t} = \frac{\mathrm{d}F_{S_n}(t)}{\mathrm{d}\beta_t}\frac{\mathrm{d}\beta_t}{\mathrm{d}t} \tag{B.11}$$

In the RHS, the first differential term is given by Eq. (B.8) and the second term can be evaluated using the definition of β_t given by Eq. (10) as

$$\frac{\mathrm{d}\beta_t}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}(e^{-a\Lambda(t)}) = -a\lambda(t)\,e^{-a\Lambda(t)} = -a\lambda(t)\,\beta_t$$

Using this result, the final expression is obtained as

$$f_{S_n}(t) = \frac{a\lambda(t)}{B(\alpha,n)}(\beta_t)^{\alpha}(1-\beta_t)^{n-1} \tag{B.12}$$

Appendix C. Derivation of the maximum load distribution

C.1. Approach

This Section presents the derivation of the distribution of maximum load generated by a LEYP process with time-dependent loads, as shown in Fig. 3.

The probabilistic structure of a point process in a time interval, $(s,t], s < t$, is characterized by the joint distribution of the number of events, N_{st}, and their arrivals times, $\mathbf{S}_{st} = (S_{N(s)+1}, S_{N(s)+2}, \ldots, S_{N(t)})$. For processes like LEYP and NHPP, it is convenient to write the joint distribution in a conditional form,

$$f_{\mathbf{S}_{st},N_{st}}(s_1,\ldots,s_n,n) = f_{\mathbf{S}_{st}|N_{st}}(s_1,\ldots,s_n|N_{st}=n)f_{N_{st}}(n) \tag{C.1}$$

The reason is that the conditional distribution can be expressed explicitly using the order statistics property of the arrival times, as shown in this section.

Suppose the number of events in $(0,t_1]$ is N_1 and N_2 in $(0,t_2]$. Then, the interval $(t_1,t_2]$ contains $N_{12} = N_2 - N_1$ events, numbered as $N_1+1,\ldots,N_2$. For the sake of brevity, loads and their arrival times in this interval are numbered as $X_{ci}, S_i, i = 1,2,\ldots,N_{12}$.

The maximum load CDF is defined as

$$\begin{aligned} F_{max}(x,t_1,t_2) &= \mathbb{P}\left[\max\left(X_{c1}(S_1),\ldots,X_{cN_{12}}(S_{N_{12}})\right) \leqslant x\right] \\ &= \mathbb{P}\Big[\bigcap_{i=1}^{N_{12}} X_{ci}(S_i) \leqslant x\Big] = \mathbb{P}\Big[\bigcap_{i=1}^{N_{12}} X_i \leqslant \psi(x,S_i)\Big] \end{aligned} \tag{C.2}$$

where $\psi(x,s)$ is defined by Eq. (34).

C.2. Derivation

Using the total probability theorem, Eq. (C.2) is simplified as,

$$\begin{aligned} F_{max}(x,t_1,t_2) &= \sum_{n=0}^{\infty}\mathbb{P}\Big[\bigcap_{i=1}^{n} X_i \leqslant \psi(x,S_i), N_{12}=n\Big] \\ &= \sum_{n=0}^{\infty}\mathbb{P}\Big[\bigcap_{i=1}^{n} X_i \leqslant \psi(x,S_i)|N_{12}=n\Big]\,\mathbb{P}\left[N_{12}=n\right] \end{aligned} \tag{C.3}$$

The conditional probability term in the above expression is the function of a vector of RVs, $\mathbf{S} \equiv (S_1,\ldots,S_n)$. Therefore, this term is evaluated as an expectation with respect to the joint distribution of arrival times, $S_1,\ldots,S_n$, conditioned on $(N_{12}=n)$, as

$$\begin{aligned} &\mathbb{P}\Big[\bigcap_{i=1}^{n} X_i \leqslant \psi(x,S_i)|N_{12}=n\Big] \\ &= \int\cdots\int_{t_1<s_1<\cdots<s_n\leqslant t_2} \mathbb{P}\left[X_1 \leqslant \psi(x,s_1),\ldots,X_n \leqslant \psi(x,s_n)\right] \\ &f(s_1,\ldots,s_n|N_{12}=n)\,\mathrm{d}s_1\ldots\mathrm{d}s_n \\ &= \int\cdots\int_{t_1<s_1<\cdots<s_n\leqslant t_2} F_X(\psi(x,s_1))\cdots F_X(\psi(x,s_n)) \\ &f(s_1,\ldots,s_n|N_{12}=n)\,\mathrm{d}s_1\ldots\mathrm{d}s_n \end{aligned} \tag{C.4}$$

The above equation uses the following two results. First, random variables, $X_1,\ldots,X_n$, are independent of $S_1,\ldots,S_n$. Second, $X_1,\ldots,X_n$ are themselves independent and identically distributed, which results in the product of univariate CDFs in the third line of Eq. (C.4).

To evaluate the multivariate integral, the conditional distribution of arrival times is required, which is equivalent to the joint distribution of the n order statistics, $(V^{(1)},\ldots,V^{(n)})$, of an i.i.d. sample of a random variable V with the following PDF [11, Theorem 3]:

$$f_V(v,t_1,t_2) = \frac{a\lambda(v)e^{a\Lambda(v)}}{e^{a\Lambda(t_2)}-e^{a\Lambda(t_1)}}, \qquad (t_1 < v \leq t_2) \tag{C.5}$$

This distribution can be rewritten in the present notation as

$$f_V(v,t_1,t_2) = \frac{a\lambda(v)}{\beta_v}\frac{\beta_2}{1-\beta_{12}} = \frac{a\alpha}{\mu_{12}}\frac{\lambda(v)}{\beta_v} \qquad (t_1 < s \leq t_2) \tag{C.6}$$

Therefore, the conditional PDF of $S_1,\ldots,S_n$ given $(N_{12}=n)$ in Eq. (C.4) can be replaced by the PDF of the order statistics, $(V^{(1)},\ldots,V^{(n)})$ [20, p. 317].

A further simplification is achieved by recognizing that the integrand, the product of n probabilities, is a symmetric function, meaning that the function value is the same for any permutation of $v_{(1)},\ldots,v_{(n)}$. Consequently, the order statistics can be replaced by (unordered) *iid* random variables, $v_1,\ldots,v_n$, with an identical distribution given by Eq. (C.6) [21, Section 3.4]. This leads to

$$\begin{aligned} &\mathbb{P}\Big[\bigcap_{i=1}^{n} X_i \leqslant \psi(x,S_i)|N_{12}=n\Big] = \\ &\int_{t_1}^{t_2}\cdots\int_{t_1}^{t_2}\left[F_X(\psi(x,v_1))\cdots F_X(\psi(x,v_n))\right] f(v_1)\ldots f(v_n)\,\mathrm{d}v_1\ldots\mathrm{d}v_n \end{aligned} \tag{C.7}$$

The multi-dimensional integral is now separable in n one-dimensional integrals, which can be written in the following compact form:

$$\mathbb{P}\Big[\bigcap_{i=1}^{n} X_i \leqslant \psi(x,S_i)|N_{12}=n\Big] = \left[\int_{t_1}^{t_2} F_X(\psi(x,s))f_V(s)\,\mathrm{d}s\right]^n = \left[q(x,t_1,t_2)\right]^n \tag{C.8}$$

Substituting for $f_V(s)$ given by Eq. (C.6), the function, $q(x,t_1,t_2)$, can be written as

$$q(x,t_1,t_2) = \int_{t_1}^{t_2} F_X(\psi(x,s))f_V(s)\,\mathrm{d}s = \frac{a\alpha}{\mu_{12}}\int_{t_1}^{t_2}\frac{\lambda(s)}{\beta_s}F_X(\psi(x,s))\mathrm{d}s \tag{C.9}$$

Substituting this result in Eq. (C.3) along with $\mathbb{P}\left[N_{12}=n\right]$ from Eq. (14) leads to

$$F_{max}(x,t_1,t_2) = \sum_{n=0}^{\infty} \left[q(x,t_1,t_2)\right]^n \frac{\Gamma(\alpha+n)}{n!\Gamma(\alpha)} (\beta_{12}^*)^{\alpha} (1-\beta_{12}^*)^n$$
$$= (\beta_{12}^*)^{\alpha} \sum_{n=0}^{\infty} \frac{\Gamma(\alpha+n)}{n!\Gamma(\alpha)} \left(q(x,t_1,t_2)(1-\beta_{12}^*)\right)^n \quad \text{(C.10)}$$

The infinite sum can be simplified using the property of the negative binomial distribution given by Eq. (6), which leads to a more concise expression,

$$F_{max}(x,t_1,t_2) = (\beta_{12}^*)^{\alpha} \left[1 - q(x,t_1,t_2)(1-\beta_{12}^*)\right]^{-\alpha} \quad \text{(C.11)}$$

Recall from Eq. (16) that $1-\beta_{12}^* = \mu_{12}\beta_{12}^*/\alpha$, and substituting for $q(x,t_1,t_2)$ from Eq. (C.9) leads to the following expression

$$F_{max}(x,t_1,t_2) = (\beta_{12}^*)^{\alpha} \left[1 - a\beta_{12}^* q^*(x,t_1,t_2)\right]^{-\alpha} \quad \text{(C.12)}$$

where

$$q^*(x,t_1,t_2) = \int_{t_1}^{t_2} \frac{\lambda(s)}{\beta_s} F_X(\psi(x,s))\mathrm{d}s \quad \text{(C.13)}$$

The above expression for $F_{max}(x,t_1,t_2)$ can be written in different forms, either in terms of the parameter β_{12}^* as

$$F_{max}(x,t_1,t_2) = \left[\frac{1}{\beta_{12}^*} - aq^*(x,t_1,t_2)\right]^{-\alpha} \quad \text{(C.14)}$$

or in terms of μ_{12}, which is the expected number of events in $(t_1,t_2]$, as

$$F_{max}(x,t_1,t_2) = \left[1 + \frac{\mu_{12}}{\alpha} - aq^*(x,t_1,t_2)\right]^{-\alpha} \quad \text{(C.15)}$$

Note that the above expression is obtained by substituting $(1/\beta_{12}^*) = 1 + (\mu_{12}/\alpha)$.

In summary, the maximum value distribution depends on: (1) the mean number of events in the interval, and (2) the distribution of the magnitude of load, as included in the function, $q^*(x,t_1,t_2)$.

Data availability

Data will be made available on request.

References

[1] Kharin VV, Zwiers FW. Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere–ocean GCM. J Clim 2000;13(21):3760–88.

[2] Kharin VV, Zwiers FW. Estimating extremes in transient climate change simulations. J Clim 2005;18(8):1156–73.

[3] Li C, Zwiers F, Zhang X, Li G, Sun Y, Wehner M. Changes in annual extremes of daily temperature and precipitation in CMIP6 models. J Clim 2021;34(9):3441–60.

[4] Li SH. Effect of nonstationary extreme wind speeds and ground snow loads on the structural reliability in a future canadian changing climate. Struct Saf 2023;101(102296):1–15.

[5] Hong HP, Tang Q, Yang SC, Cui XZ, Cannon AJ, Lounis Z, Irwin P. Calibration of the design wind load and snow load considering the historical climate statistics and climate change effects. Struct Saf 2021;93:102135.

[6] Alhamid AK, Akiyama M, Aoki K, Koshimura S, Frangopol DM. Stochastic renewal process model of time-variant tsunami hazard assessment under nonstationary effects of sea-level rise due to climate change. Struct Saf 2022;99(102263):1–17.

[7] Pandey MD, Lounis Z. Stochastic modelling of non-stationary environmental loads for reliability analysis under the changing climate. Struct Saf 2023;103:102348, 1–11.

[8] Badía FG, Mercier S, Sangüesa C. Extensions of the generalized Pólya process. Methodol Comput Appl Probab 2019;21:1057–85.

[9] Daley DJ, Vere-Jones D. An introduction to the theory of point processes, vol. 1, New York: Springer-Verlag; 2003.

[10] Taylor HM, Karlin S. An introduction to stochastic modeling. 3rd ed.. San Diego, CA: Academic Press; 1998.

[11] Cha JH. Characterization of the generalized Pólya process and its applications. Adv in Appl Probab 2014;46:1148–71.

[12] Le Gat Y. Extending the yule process to model recurrent pipe failures in water supply networks. Urban Water J 2014;11(8):617–30.

[13] Konno H. On the exact solution of a generalized Pólya process. Adv Math Phys 2010;2010:504267, 1–12.

[14] Johnson N, Kotz S, Kemp A. Univariate discrete distributions. New York, NY: John Wiley & Sons; 1992.

[15] Gradshteyn IS, Ryzhik IM. Table of integrals, series, and products. 5th ed.. San Diego, CA: Academic Press; 1994.

[16] CSA. Development, interpretation, and use of rainfall intensity-duration-frequency(idf) information: Guideline for canadian water resources practitioners. Tech. rep., Toronto, ON: Canadian Standards Association; 2019, 4013:19.

[17] Sobie SR, Zwiers FW, Curry CL. Climate model projections for Canada: A comparison of CMIP5 and CMIP6. Atmos-Ocean 2021;59(4–5):269–84.

[18] Le Gat Y. Recurrent event modeling based on the yule process: application to water network asset management. John Wiley & Sons; 2015.

[19] Rider PR. The negative binomial distribution and the incomplete beta function. Amer Math Monthly 1962;69(4):302–4.

[20] Ross SM. Introduction to probability models. 9th ed.. Orlando, FL: Academic Press; 2007.

[21] Serfozo R. Basics of applied stochastic processes. Springer-Verlag Berlin Heidelberg; 2009.