LINEAR PREDICTION OF K-RECORD VALUES FROM GENERALIZED PARETO DISTRIBUTION

Laji Muraleedharan[1] and Manoj Chacko[2]

•

1. Department of Statistics, Government Arts College, Thiruvananthapuram, Kerala, India
2. Department of Statistics, University of Kerala, Thiruvananthapuram, Kerala, India
[1] lajikmr@gmail.com, [2] manojchacko02@gmail.com

Abstract

In this paper, the upper k-record values arising from a generalized Pareto distribution is considered. After considering the means, variances and covariances of the upper k-record values, the best linear unbiased estimators and best lnear invariant estimators of the location and scale parameters of the generalized Pareto distribution is discussed under the assumption that the shape parameter is assumed to be known. The marginal best linear unbiased predictor and best linear invariant predictor of future upper k-record value and the joint best linear unbiased predictor and best linear invariant predictor of a pair of future upper k-record values are also determined. Finally, a real dataset is considered to illustrate the proposed inference procedures developed in this paper.

Keywords: Generalized Pareto distribution; best linear unbiased estimation; best linear invariant estimation; joint best linear unbiased prediction; joint best linear invariant prediction.

1. Introduction

Record values are observed in many situations of daily life as well as many statistical phinomena. These associated statistics are greater importance in live problems such as data related to meteorology, hydrology, athletic events and mining etc. The statistical study of record values started first time in the literature when [6] studied the stochastic behaviour of independent and identically distributed (iid) random variables having an absolutely continuous cumulative distribution function (cdf). So many researchers have adressed the statistical applications of record values in the literature pertain many live problems. Interested surveys are given by [3] and [9].

Let $\{X_i, i \geq 1\}$ be a sequence of iid continuous random variables with cdf $F(x)$ and probability density function (pdf) $f(x)$. An observation X_j is called an upper record value if $X_j > X_i$, for every $i < j$. Estimation of parameters based on record values and prediction of future record values have been extensively studied by several authors. For more detailed discussion on these area, see, [14], [19] and [18].

One of the challenges in dealing with problems involving inference with record data is that the expected waiting time for every records after the first is infinite. Such an issue does not arise if we use the *k*-records proposed by [7]. We use the following formal definition of k-record values given by [3].

For a fixed positive integer k, the upper k-record times $\tau_{n(k)}$ and the upper k-record values $R_{n(k)}$ are defined as follows: .

Define $\tau_{1(k)} = k$ and $R_{1(k)} = X_{1:k}$, where $X_{i:m}$ denotes the ith order statistic in a sample of size m. Then for $n > 1$,

DOI: https://doi.org/10.24412/1932-2321-2025-489-757-771

$$\tau_{n(k)} = min\left\{i : i > \tau_{n-1(k)},\ X_i > X_{\tau_{n-1(k)}-k+1:\tau_{n-1(k)}}\right\}.$$

Then the sequence of upper k-record values $\left\{R_{n(k)}, n \geq 1\right\}$ is defined as

$$R_{n(k)} = X_{\tau_{n(k)}-k+1:\tau_{n(k)}}.$$

In an analogous way, one can define the lower k-record values. The ordinary record values are contained in the k-records as a special case when $k = 1$.
The pdf of nth upper k-record value $R_{n(k)}$ is given by (see, [3])

$$f_{n(k)}(x) = \frac{k^n}{\Gamma(n)} \{-\log[1 - F(x)]\}^{n-1} [1 - F(x)]^{k-1} f(x), \quad -\infty < x < \infty, \tag{1}$$

where $\Gamma(.)$ denotes the complete gamma function. The joint pdf of mth and nth upper k-record values, $R_{m(k)}$ and $R_{n(k)}$, for $m < n$, is given by

$$\begin{aligned} f_{m,n(k)}(x,y) &= \frac{k^n}{\Gamma(m)\Gamma(n-m)} \{-\log[1 - F(x)]\}^{m-1} \{-\log[1 - F(y)] + \log[1 - F(x)]\}^{n-m-1} \\ &\times \frac{[1 - F(y)]^{k-1}}{1 - F(x)} f(x)f(y),\ -\infty < x < y < \infty. \end{aligned} \tag{2}$$

Recently, k-record data has shown a significant role in statistical inference and future event prediction. [1] considered the problem of prediction of order statistics based on observed k-record data when the records are arising from an exponential distribution. [21] considered the same problem for single parameter generalized exponential distribution. [16] discussed the prediction problem for a Pareto distribution when the future sample size is both fixed and random.

In statistical inference, predicting future events based on the current knowledge is a fundamental problem. Many researchers have been expressed this sort of problem in different approaches. [17] studied the estimation and prediction problem of lognormal distribution based on record values. [13] discussed the problem of investigating the BLUP and best linear invariant predictor (BLIP) of future lower k-record values based on the observed lower k-record values arising from Unit-Gompertz distribution. [10] considered the same problem for an xgamma distribution based on the observed upper k-record values. In this paper, we consider the linear prediction of upper k-record value based on the observed upper k-record values arising from a generalized Pareto (GP) distribution with pdf given by

$$f(x) = \frac{1}{\sigma}\left[1 + \psi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\left(\frac{1}{\psi}+1\right)}, \tag{3}$$

where $\mu, \psi \in R$ and $\sigma > 0$. Here ψ is a shape parameter, σ is a scale parameter and μ is a location parameter. The GP distribution with pdf given in (3) will be denoted by GP(μ, σ, ψ). The cdf corresponding to the pdf (3) is given by

$$F(x) = 1 - \left[1 + \psi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\psi}}. \tag{4}$$

When $\psi = 0$, the GPD reduces the exponential distribution. When $\psi > 0$, the distribution has the support $\mu \leq x < \infty$ while for $\psi < 0$, $\mu \leq x \leq \mu - \frac{\sigma}{\psi}$. The distribution has heavy tailed when $\psi > 0$ and the distribution is short tailed if $\psi < 0$.

The GP distribution was introduced by Pickands in 1975. Some of its major applications include analysis of extreme events in the modeling of insurance claims and to describe the annual maximum flood at river gauging station. Many extensive works were done about the inference of the parameters of GP distribution based on upper record values. [20] derived the explicit expressions for the single, double, triple and quadruple moments of the upper record values from

GP distribution. [2] introduced minimum variance linear unbiased estimator and the best linear invariant estimator (BLIE) for the location and scale parameters of GP distribution when the shape parameter is known based on upper record values. [12] discussed the best linear unbiased predictor (BLUP) for future upper k-record values for a generalized Pareto distribution when the shape parameter is assumed to be known. More recently, [22] investigate the point estimation and confidence interval estimation for the heavy-tailed generalized Pareto distribution based on the upper record values.

Recently, [4] and [5] discussed the problem of simultanious prediction of two order statistics from a future sample and established the gain in efficiency of joint prediction over marginal prediction. When working with location- scale family of distributions, the joint prediction of future ordered observations provides more accurate, robust and insightful prediction than marginal prediction especially in the context of observed ordered observations are correlated or interdependent. This gives better understanding risks, optimizing resource allocation, improving prediction accuracy and making more informed strategic decisions. This provides a motivation for considering joint prediction of future k-record values.

The rest of the paper is organized as follows. In Section 2, we provide the linear estimators, which includes BLUEs and BLIEs for the location and scale parameters of GP distribution based on observed upper k-record values. In Section 3, we consider the BLUP and BLIP of future upper k-record values which incldes marginal predictor of future upper k-record values and the joint predictor of pair of future upper k-record values. The estimated mean square predictive error (MSPE) and the relative efficiencies for BLUP and BLIP are also discussed. In Section 4, a real dataset is considered to illustrate the inference procedures developed in this paper and finally some concluding remarks are made in Section 5.

2. Linear Estmation of Parameters

Let $\{Z_i, i \geq 1\}$ be a sequence of iid observations taken from GP$(0,1,\psi)$ with pdf and cdf defined in (3) and (4) respectively. Let $Z_{1(k)}, Z_{2(k)}, \ldots, Z_{n(k)}$ be the first n upper k-record values extracted from the sequence $\{Z_i, i \geq 1\}$. Then by using (1), the pdf of the nth upper k-record $Z_{n(k)}$ is obtained as

$$g_{n(k)}(x) = \frac{k^n}{\Gamma(n)\psi^{n-1}} \left[\log(1+\psi x)\right]^{n-1} (1+\psi x)^{-\left(\frac{k}{\psi}+1\right)}, \; x > 0, \psi > 0. \tag{5}$$

By using (2), the joint pdf of nth and mth upper k-records $Z_{n(k)}$ and $Z_{m(k)}$ for $m < n$ is given by

$$\begin{aligned} g_{m,n(k)}(x,y) &= \frac{k^n}{\Gamma(m)\Gamma(n-m)\psi^{n-2}} \left[\log(1+\psi x)\right]^{m-1} \\ &\times \left[\log(1+\psi y) - \log(1+\psi x)\right]^{n-m-1} \frac{(1+\psi y)^{-\left(\frac{k}{\psi}+1\right)}}{(1+\psi x)}, \; 0 \leq x < y < \infty. \end{aligned} \tag{6}$$

Since (5) is the pdf of nth upper k-records arising from GP$(0,1,\psi)$, if we denote $E\left(Z_{n(k)}\right)$, $Var\left(Z_{n(k)}\right)$ and $Cov\left(Z_{m(k)}, Z_{n(k)}\right)$ respectively by $\alpha_{n(k)}$, $\omega_{nn(k)}$ and $\omega_{mn(k)}$ then by [12], the explicit expressions for the means, variances and the product moments of the upper k-record values are given by the following.

$$\alpha_{n(k)} = \frac{1}{\psi}\left[\left(\frac{k}{k-\psi}\right)^n - 1\right], \text{ provided } k > \psi \tag{7}$$

$$\omega_{nn(k)} = \frac{1}{\psi^2}\left[\left(\frac{k}{k-2\psi}\right)^n - \left(\frac{k}{k-\psi}\right)^{2n}\right] \tag{8}$$

$$\omega_{mn(k)} = \frac{1}{\psi^2}\left[\frac{k^n}{(k-\psi)^{n-m}(k-2\psi)^m} - \left(\frac{k}{k-\psi}\right)^{m+n}\right], \tag{9}$$

where (8) and (9) are provided $k > 2\psi$. We have evaluated the values of means $\alpha_{n(k)}$ for $n = 1(1)9$; $\psi = 0.1, 0.2$ and $k = 2, 3$ and the values are presented in Table 1. We have computed

Table 1: *Means of upper k- record values arising from $GP(0, 1, \psi)$*

k	ψ	n								
		1	2	3	4	5	6	7	8	9
2	0.1	0.52632	1.08033	1.66351	2.27738	2.92355	3.60374	4.31973	5.07340	5.86673
	0.2	0.55556	1.17284	1.85871	2.62079	3.46754	4.40838	5.45375	6.61528	7.90587
3	0.1	0.34483	0.70155	1.07056	1.45231	1.84722	2.25574	2.67835	3.11553	3.56779
	0.2	0.35714	0.73980	1.14978	1.58905	2.05970	2.56396	3.10424	3.68312	4.30334

the values of $\omega_{nn(k)}$ and $\omega_{mn(k)}$ for $1 \leq m \leq n \leq 9$; $\psi = 0.1, 0.2$ and $k = 2, 3$ and the values are given in Table 2 and Table 3 respectively.

Table 2: *Variances and Covariances of upper k- record values from $GP(0, 1, \psi)$ for $\psi = 0.1$*

k	n	m								
		1	2	3	4	5	6	7	8	9
2	1	0.30779								
	2	0.32399	0.68302							
	3	0.34104	0.71897	1.13680						
	4	0.35899	0.75681	1.19663	1.68181					
	5	0.37788	0.79665	1.25961	1.77033	2.33262				
	6	0.39777	0.83857	1.32591	1.86351	2.45539	3.10586			
	7	0.41871	0.88271	1.39568	1.96158	2.58462	3.26932	4.02055		
	8	0.44074	0.92916	1.46914	2.06482	2.72065	3.44139	4.23216	5.09841	
	9	0.46393	0.97807	1.54646	2.17351	2.86384	3.62252	4.45491	5.36675	6.36421
3	1	0.12740								
	2	0.13179	0.27284							
	3	0.13634	0.28224	0.43823						
	4	0.14104	0.29198	0.45334	0.62567					
	5	0.14590	0.30205	0.46897	0.64724	0.83745				
	6	0.15093	0.31246	0.48514	0.66956	0.86632	1.07608			
	7	0.15613	0.32323	0.50187	0.69265	0.89619	1.11318	1.34429		
	8	0.16152	0.33438	0.51917	0.71653	0.92711	1.15156	1.39065	1.64511	
	9	0.16709	0.34591	0.53707	0.74123	0.95907	1.19127	1.43861	1.70183	1.98175

Let $R_{1(k)}, R_{2(k)}, \ldots, R_{n(k)}$ be the first n lower k-record values arising from GP(μ, σ, ψ) with pdf given in (3). Suppose $\mathbf{R}_{n(k)} = \left(R_{1(k)}, R_{2(k)}, \ldots, R_{n(k)}\right)'$ denotes the vector of n upper k-record values, Then by Aitken's generalized least-square approach, the BLUEs of μ and σ, denoted by μ^* and σ^* are respectively obtained as given below.

$$\begin{aligned} \mu^* &= \frac{\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha}\mathbf{1}'\boldsymbol{\Omega}^{-1} - \boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\mathbf{1}\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}}{(\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\boldsymbol{\Omega}^{-1}\mathbf{1}) - (\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\mathbf{1})^2}\mathbf{R}_{n(k)} \\ &= \sum_{i=1}^{n} a_i R_{i(k)} \end{aligned} \tag{10}$$

and

$$\begin{aligned} \sigma^* &= \frac{\mathbf{1}'\boldsymbol{\Omega}^{-1}\mathbf{1}\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1} - \mathbf{1}'\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha}\mathbf{1}'\boldsymbol{\Omega}^{-1}}{(\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\boldsymbol{\Omega}^{-1}\mathbf{1}) - (\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\mathbf{1})^2}\mathbf{R}_{n(k)} \\ &= \sum_{i=1}^{n} b_i R_{i(k)} \ , \end{aligned} \tag{11}$$

where $\boldsymbol{\alpha} = \left(\alpha_{1(k)}, \alpha_{2(k)}, ..., \alpha_{n(k)}\right)'$, $\mathbf{1}$ be a column vector of n ones and $\boldsymbol{\Omega} = \left(\omega_{ij(k)}\right)$ is the $n \times n$ dispersion matrix of $\mathbf{Z}_{n(k)} = \left(Z_{1(k)}, Z_{2(k)}, ..., Z_{n(k)}\right)'$. Furthermore, the variances and covariance

Table 3: *Variances and Covariances of upper k- record values from $GP(0,1,\psi)$ for $\psi = 0.2$*

k	n	m 1	2	3	4	5	6	7	8	9
	1	0.38580								
	2	0.42867	0.95855							
	3	0.47630	1.06506	1.78621						
	4	0.52922	1.18340	1.98468	2.95872					
2	5	0.58802	1.31489	2.20520	3.28747	4.59465				
	6	0.65336	1.46098	2.45022	3.65274	5.10516	6.84977			
	7	0.72595	1.62331	2.72247	4.05861	5.6724	7.61086	9.92823		
	8	0.80661	1.80368	3.02497	4.50956	6.30266	8.45651	11.03137	14.09673	
	9	0.89624	2.00409	3.36107	5.01062	7.00296	9.39612	12.25708	15.66303	19.7029
	1	0.14717								
	2	0.15769	0.33877							
	3	0.16895	0.36296	0.58483						
	4	0.18102	0.38889	0.62661	0.89745					
3	5	0.19395	0.41667	0.67136	0.96155	1.29111				
	6	0.20781	0.44643	0.71932	1.03023	1.38333	1.78314			
	7	0.22264	0.47831	0.77069	1.10382	1.48213	1.91051	2.39428		
	8	0.23854	0.51248	0.82575	1.18266	1.58800	2.04697	2.56530	3.14928	
	9	0.25558	0.54908	0.88472	1.26714	1.70143	2.19318	2.74854	3.37423	4.07764

of the above estimators are given by

$$Var(\mu^*) = \sigma^2 \left(\frac{\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha}}{(\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\boldsymbol{\Omega}^{-1}\mathbf{1}) - (\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\mathbf{1})^2} \right), \tag{12}$$

$$Var(\sigma^*) = \sigma^2 \left(\frac{\mathbf{1}'\boldsymbol{\Omega}^{-1}\mathbf{1}}{(\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\boldsymbol{\Omega}^{-1}\mathbf{1}) - (\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\mathbf{1})^2} \right) \tag{13}$$

and

$$Cov(\mu^*, \sigma^*) = \sigma^2 \left(\frac{-\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\mathbf{1}}{(\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\boldsymbol{\Omega}^{-1}\mathbf{1}) - (\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\mathbf{1})^2} \right). \tag{14}$$

By making use of Table 1, Table 2 and Table 3, we have evaluated the coefficients a_i and b_i, $i = 1, 2, \ldots, n$ of BLUEs of μ and σ for $n = 2(1)9$; $\psi = 0.1, 0.2$ and $k = 2, 3$ and the values are given in Table 4 to Table 7 respectively. If we denote

$$\nu_1^* = \frac{Var(\mu^*)}{\sigma^2} \tag{15}$$

$$\nu_2^* = \frac{Var(\sigma^*)}{\sigma^2} \tag{16}$$

and

$$\nu_3^* = \frac{Cov(\mu^*, \sigma^*)}{\sigma^2} \tag{17}$$

then the values of ν_1^*, ν_2^* and ν_3^* are evaluated for $n = 2(1)9$; $\psi = 0.1, 0.2, 0.3$ and $k = 2, 3$ are also computed from equations (12) to (14) and are given in Table 8.
Next, we focus on the BLIEs of the location and scale parameters of GP(μ, σ, ψ), denoted respectively by μ^{**} and σ^{**} when the shape parameter ψ is known. Based on the results of [11] and [3], the BLIEs of μ and σ are given by

$$\mu^{**} = \mu^* - \frac{\nu_3^*}{1 + \nu_2^*}\sigma^* \tag{18}$$

and

$$\sigma^{**} = \frac{\sigma^*}{1 + \nu_2^*}. \tag{19}$$

The variances of (18) and (19) are given by (see, [3])

$$Var\left(\mu^{**}\right) = \sigma^2 \left(\nu_1^* - \frac{\nu_3^{*2}\left(2+\nu_2^*\right)}{\left(1+\nu_2^*\right)^2} \right) \tag{20}$$

and

$$Var\left(\sigma^{**}\right) = \frac{\sigma^2 \nu_2^*}{\left(1+\nu_2^*\right)^2}, \tag{21}$$

where ν_1^*, ν_2^* and ν_3^* are defined in (15),(16) and (17) respectively.

Table 4: *Coefficients for the BLUE of μ for $\psi = 0.1$*

k	n	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
2	2	1.95000	-0.95000							
	3	1.47632	-0.02632	-0.45000						
	4	1.33900	-0.01845	-0.01661	-0.28395					
	5	1.24078	-0.01454	-0.01309	-0.01178	-0.20138				
	6	1.19420	-0.01222	-0.01097	-0.00990	-0.00888	-0.15221			
	7	1.16342	-0.01067	-0.00960	-0.00866	-0.00776	-0.00699	-0.11973		
	8	1.14168	-0.00959	-0.00862	-0.00778	-0.00697	-0.00628	-0.00568	-0.09676	
	9	1.12558	-0.00878	-0.00789	-0.00712	-0.00638	-0.00575	-0.00521	-0.00466	-0.07977
3	2	1.96667	-0.96667							
	3	1.48391	-0.01724	-0.46667						
	4	1.32324	-0.01188	-0.01109	-0.30026					
	5	1.24310	-0.00921	-0.00860	-0.00803	-0.21726				
	6	1.19514	-0.00754	-0.00716	-0.00664	-0.00619	-0.16760			
	7	1.16332	-0.00640	-0.00616	-0.00571	-0.00533	-0.00498	-0.13463		
	8	1.14069	-0.00573	-0.00547	-0.00507	-0.00468	-0.00444	-0.00405	-0.11124	
	9	1.12381	-0.00516	-0.00494	-0.00458	-0.00423	-0.00400	-0.00364	-0.00349	-0.09374

Table 5: *Coefficients for the BLUE of μ for $\psi = 0.2$*

k	n	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
2	2	1.90000	-0.90000							
	3	1.45556	-0.05556	-0.40000						
	4	1.32984	-0.04098	-0.03279	-0.23607					
	5	1.23875	-0.03388	-0.02710	-0.02168	-0.15610				
	6	1.19748	-0.02975	-0.0238	-0.01904	-0.01522	-0.10966			
	7	1.17106	-0.02711	-0.02168	-0.01735	-0.01387	-0.01111	-0.07993		
	8	1.15308	-0.02531	-0.02025	-0.0162	-0.01295	-0.01038	-0.00829	-0.05971	
	9	1.14033	-0.02404	-0.01923	-0.01538	-0.01229	-0.00985	-0.00787	-0.00629	-0.04536
3	2	1.93333	-0.93333							
	3	1.46905	-0.03571	-0.43333						
	4	1.31534	-0.02547	-0.02207	-0.26780					
	5	1.23115	-0.03321	-0.02447	-0.02168	-0.15218				
	6	1.19428	-0.01748	-0.01503	-0.01308	-0.01129	-0.13739			
	7	1.16477	-0.01551	-0.01333	-0.01159	-0.01002	-0.00877	-0.10554		
	8	1.14410	-0.01412	-0.01213	-0.01056	-0.00912	-0.00799	-0.00683	-0.08333	
	9	1.12897	-0.01311	-0.01126	-0.00980	-0.00846	-0.00742	-0.00634	-0.00551	-0.06704

Table 6: *Coefficients for the BLUE of σ for $\psi = 0.1$*

k	n	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9
2	2	-1.80500	1.80500							
	3	-0.90500	0.05000	0.85500						
	4	-0.60611	0.03505	0.03155	0.53950					
	5	-0.45749	0.02762	0.02486	0.02238	0.38262				
	6	-0.36897	0.02320	0.02087	0.01879	0.01690	0.28919			
	7	-0.31049	0.02026	0.01825	0.01643	0.01476	0.01331	0.22747		
	8	-0.26919	0.01821	0.01638	0.01475	0.01326	0.01194	0.01075	0.18386	
	9	-0.23861	0.01667	0.01502	0.01351	0.01216	0.01094	0.00985	0.00886	0.15158
3	2	-2.80333	2.80333							
	3	-1.40333	0.05000	1.35333						
	4	-0.93741	0.03447	0.03217	0.87077					
	5	-0.70500	0.02672	0.02494	0.02328	0.63006				
	6	-0.56597	0.02206	0.02063	0.01923	0.01799	0.48606			
	7	-0.47367	0.01900	0.01774	0.01653	0.01549	0.01441	0.39048		
	8	-0.40806	0.01680	0.01572	0.01467	0.01363	0.01281	0.01190	0.32251	
	9	-0.35911	0.01516	0.01419	0.01325	0.01232	0.01156	0.01072	0.01003	0.27184

Table 7: *Coefficients for the BLUE of σ for $\psi = 0.2$*

k	n	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9
2	2	-1.62000	1.62000							
	3	-0.82000	0.10000	0.72000						
	4	-0.55770	0.07377	0.05902	0.42492					
	5	-0.42976	0.06098	0.04878	0.03902	0.28098				
	6	-0.35546	0.05355	0.04283	0.03427	0.02741	0.19739			
	7	-0.30790	0.04879	0.03903	0.03122	0.02497	0.01999	0.14388		
	8	-0.27553	0.04555	0.03644	0.02915	0.02332	0.01866	0.01492	0.10747	
	9	-0.25257	0.04325	0.03461	0.02769	0.02214	0.01772	0.01417	0.01134	0.08165
3	2	-2.61333	2.61333							
	3	-1.31333	0.10000	1.21333						
	4	-0.88294	0.07131	0.06180	0.74984					
	5	-0.66993	0.05711	0.04949	0.04289	0.52044				
	6	-0.54389	0.04878	0.04216	0.03661	0.03168	0.38465			
	7	-0.46124	0.04325	0.03742	0.03243	0.02812	0.02437	0.29564		
	8	-0.40337	0.03938	0.03405	0.02956	0.02559	0.02221	0.01922	0.23335	
	9	-0.36100	0.03655	0.03162	0.02742	0.02376	0.02062	0.01783	0.01546	0.18771

Table 8: *Variances and Covariances of the BLUEs of μ and σ in terms of σ^2*

k	n	ψ=0.1			ψ=0.2			ψ=0.3		
		ν_1^*	ν_2^*	ν_3^*	ν_1^*	ν_2^*	ν_3^*	ν_1^*	ν_2^*	ν_3^*
2	2	0.58642	1.11698	-0.55864	0.70313	1.27813	-0.64063	0.98142	2.14117	-0.99586
	3	0.44022	0.58920	-0.28086	0.52951	0.71562	-0.32813	0.84022	0.95892	-0.58086
	4	0.39167	0.41392	-0.18861	0.47259	0.53120	-0.22567	0.79167	0.74139	-0.31886
	5	0.36753	0.32677	-0.14274	0.44483	0.44123	-0.17569	0.67525	0.62677	-0.19274
	6	0.35315	0.27485	-0.11542	0.42871	0.38899	-0.14666	0.54142	0.56468	-0.20612
	7	0.34365	0.24056	-0.09737	0.41838	0.35555	-0.12808	0.53062	0.53347	-0.18776
	8	0.33694	0.21634	-0.08462	0.41135	0.33279	-0.11544	0.52394	0.51419	-0.17642
	9	0.33197	0.19841	-0.07519	0.40637	0.31665	-0.10647	0.51957	0.50155	-0.16898
3	2	0.24660	1.07390	-0.35799	0.27613	1.16489	-0.38856	0.31466	1.27390	-0.45180
	3	0.18502	0.55604	-0.17942	0.20745	0.62643	-0.19625	0.28502	0.75604	-0.27892
	4	0.16453	0.38369	-0.11999	0.18471	0.44816	-0.13259	0.21645	0.50369	-0.19992
	5	0.15431	0.29772	-0.09035	0.17346	0.35993	-0.10107	0.19603	0.42977	-0.15035
	6	0.14819	0.24631	-0.07262	0.16678	0.30770	-0.08241	0.19054	0.38899	-0.09778
	7	0.14413	0.14413	-0.06084	0.16242	0.27347	-0.07019	0.18595	0.35555	-0.08539
	8	0.14124	0.18789	-0.05247	0.15937	0.24951	-0.06163	0.18282	0.33279	-0.07696
	9	0.13909	0.16978	-0.04623	0.15713	0.23196	-0.05536	0.18061	0.31665	-0.07098

3. Prediction of future k-record values

Prediction of future events on the basis of the past and present knowledge is a fundamental problem of Statistics, arising in many contexts and producing varied solutions. Prediction of future records becomes a problem of great interest. For example, while studying the minimum temperature, having observed the record values until the present time, we will be naturally interested in predicting the minimum temperature that is to be expected when the present record is broken for the first time in future. In this section, we consider prediction for future upper k-record value from generalized Pareto distribution based on past observed upper k-records.

3.1. Marginal Prediction of future k-record values

Suppose that the vector of upper k-record values $\mathbf{R_{n(k)}} = (R_{1(k)}, R_{2(k)}, \ldots, R_{n(k)})'$ have been observed from the generalized Pareto distribution. The problem of interest then is to predict the value of the next upper k-record $R_{n+1(k)}$ or, more generally, the value of the sth upper k-record value $R_{s(k)}, s > n$. For a location-scale family with location parameter μ and scale parameter σ, the BLUP of the sth upper k- record value, $R_{s(k)}$ is given by (see, [3])

$$R^*_{s(k)} = (\mu^* + \alpha_{s(k)}\sigma^*) + \boldsymbol{\omega}'_{\mathbf{s(k)}}\mathbf{\Omega}^{-1}\left(\mathbf{R}_{n(k)} - \mu^*\mathbf{1} - \sigma^*\boldsymbol{\alpha}\right), \tag{22}$$

where μ^* and σ^* are the BLUEs of μ and σ based on the first n upper k-record values and $\alpha_{s(k)}$ is the of mean of sth upper k-record value from the standard GP distribution, $\mathbf{\Omega}$ is the variance-covariance matrix of first n upper k-records and $\boldsymbol{\omega}'_{\mathbf{s(k)}} = \left(\omega_{1s(k)}, \ldots, \omega_{ns(k)}\right)$ is the vector of the covariances between the sth future record statistic and the first n upper k-record observations. The estimated MSEP of $R^*_{s(k)}$ is obtained as [15]

$$MSEP\left(R^*_{s(k)}\right) = \sigma^2\left(\Delta'\Omega^{-1}\Delta + \omega_{nn(k)} - 2\Delta'\omega_{s(k)}\right), \tag{23}$$

where

$$\Delta' = \left(1\ ,\ \alpha_{n(k)}\right)\left(\mathbf{A}'\Omega^{-1}\mathbf{A}\right)^{-1}\mathbf{A}'\Omega^{-1} + \boldsymbol{\omega}'_{\mathbf{s(k)}}\Omega^{-1}\left[\mathbf{I}_n - \mathbf{A}\left(\mathbf{A}'\Omega^{-1}\mathbf{A}\right)^{-1}\mathbf{A}'\Omega^{-1}\right] \tag{24}$$

and $\mathbf{I}_n$ is the identity matrix of order n. The BLIP of the sth upper k- record value, $R_{s(k)}$ is given by

$$R^{**}_{s(k)} = R^*_{s(k)} - \left(\frac{\nu^*_4}{1+\nu^*_2}\right)\sigma^*, \tag{25}$$

where $\nu^*_4 = \left(1 - \boldsymbol{\omega}'_{\mathbf{s(k)}}\mathbf{\Omega}^{-1}\mathbf{1}\right)\nu^*_1 + \left(\alpha_{s(k)} - \boldsymbol{\omega}'_{\mathbf{s(k)}}\mathbf{\Omega}^{-1}\boldsymbol{\alpha}\right)\nu^*_2$. The estimated MSEP of $R^{**}_{s(k)}$ is obtained as given below (see, [5]).

$$\begin{aligned} MSEP\left(R^{**}_{s(k)}\right) &= \sigma^2\left[\rho^2\left(\boldsymbol{\alpha}'\mathbf{1} - 1\right)^2 + 2\rho\left(\boldsymbol{\alpha}'\mathbf{1} - 1\right)\left(\boldsymbol{\alpha}'\boldsymbol{\alpha} - \alpha_{s(k)}\right)\right] \\ &+ \sigma^2\left[\boldsymbol{\alpha}'\mathbf{\Omega}\boldsymbol{\alpha} - 2\boldsymbol{\alpha}'\boldsymbol{\omega}_{\mathbf{s(k)}} + \omega_{ss(k)} + \left(\boldsymbol{\alpha}'\boldsymbol{\alpha} - \alpha_{s(k)}\right)^2\right], \end{aligned} \tag{26}$$

where $\rho = \frac{\mu}{\sigma}$ and $\omega_{ss(k)}$ is the variance of the sth upper k-record value.

3.2. Joint Prediction of future k-record values

In this subsection we obtain the joint prediction of the pair $\left(R_{s(k)}, R_{t(k)}\right)$ based on n observed upper k-record values arising from GP(μ, σ, ψ) for $n < s < t$. The following theorems provide the explicit expression for the joint BLUP and joint BLIP of the pair of future upper k-record values $\left(R_{s(k)}, R_{t(k)}\right)$, for $n < s < t$.

Theorem 1. Let $\mathbf{R_{n(k)}} = \left(R_{1(k)}, R_{2(k)}, \ldots, R_{n(k)}\right)'$ be the $1 \times n$ vector of first n observed upper k-record values arising from GP(μ, σ, ψ). Then $\left(R^*_{s(k)}, R^*_{t(k)}\right)$, for $n < s < t$, is the joint BLUP of the pair $\left(R_{s(k)}, R_{t(k)}\right)$ where $R^*_{s(k)} = \sum_{j=1}^{n} c_j R_{j(k)}$ and $R^*_{t(k)} = \sum_{j=1}^{n} d_j R_{j(k)}$. Then the coefficients c_j and d_j are given by

$$c_j = \frac{\sum_{i=1}^{n}\left(\alpha_{i(k)} - \alpha_{s(k)}\right)\left(S_{i(k)}T_{j(k)} - T_{i(k)}S_{j(k)}\right)}{(\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\boldsymbol{\Omega}^{-1}\mathbf{1}) - (\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\mathbf{1})^2}, j = 1, 2, \ldots, n \tag{27}$$

and

$$d_j = \frac{\sum_{i=1}^{n}\left(\alpha_{i(k)} - \alpha_{t(k)}\right)\left(T_{i(k)}S_{j(k)} - S_{i(k)}T_{j(k)}\right)}{(\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\boldsymbol{\Omega}^{-1}\mathbf{1}) - (\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\mathbf{1})^2}, j = 1, 2, \ldots, n, \tag{28}$$

where $T_{i(k)}$ and $S_{i(k)}$ are the ith entry of column vector $\Omega^{-1}\mathbf{1}$ and is $\Omega^{-1}\boldsymbol{\alpha}$, respectively.

Proof. We will derive the BLUPs jointly by minimizing the disperson matrix of $\left(R^*_{s(k)}, R^*_{t(k)}\right)$ with respect to the vectors $\mathbf{c}' = (c_1, c_2, \ldots, c_n)$ and $\mathbf{d}' = (d_1, d_2, \ldots, d_n)$. Let Σ denote the variance-covariance matrix of the BLUP $\left(R^*_{s(k)}, R^*_{t(k)}\right)$. Then Σ is of the form

$$\Sigma = Cov\left(R^*_{s(k)}, R^*_{t(k)}\right) = \sigma^2\begin{pmatrix}\mathbf{c}'\boldsymbol{\Omega}\mathbf{c} & \mathbf{c}'\boldsymbol{\Omega}\mathbf{d}\\ \mathbf{c}'\boldsymbol{\Omega}\mathbf{d} & \mathbf{d}'\boldsymbol{\Omega}\mathbf{d}\end{pmatrix} = \sigma^2\begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix} (say).$$

Then the joint BLUPs of $R_{s(k)}$ and $R_{t(k)}$ are obtained by minimizing $|\Sigma|$ subject to the unbiasedness constraints of BLUPs $\mathbf{c}'\mathbf{1} = 1$, $\mathbf{c}'\boldsymbol{\alpha} = \alpha_{s(k)}$, $\mathbf{d}'\mathbf{1} = 1$ and $\mathbf{d}'\boldsymbol{\alpha} = \alpha_{t(k)}$. Then the optimum values of $\mathbf{c}$ and $\mathbf{d}$ are obtained by minimizing the following generalized variance. We can leave the multiplication factor σ^2 in the minimization process.

$$\begin{aligned} Q(\mathbf{c}, \mathbf{d}, \boldsymbol{\lambda}) &= \left(\mathbf{c}'\boldsymbol{\Omega}\mathbf{c}\right)\left(\mathbf{d}'\boldsymbol{\Omega}\mathbf{d}\right) - \left(\mathbf{c}'\boldsymbol{\Omega}\mathbf{d}\right)^2 - 2\lambda_1\left(\mathbf{c}'\mathbf{1} - 1\right) - 2\lambda_2\left(\mathbf{c}'\boldsymbol{\alpha} - \alpha_{s(k)}\right)\\ &- 2\lambda_3\left(\mathbf{d}'\mathbf{1} - 1\right) - 2\lambda_4\left(\mathbf{d}'\boldsymbol{\alpha} - \alpha_{t(k)}\right), \end{aligned} \tag{29}$$

were $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ is the vector of Lagrangian multiplayers. Differentiating (29) with respect to $\mathbf{c}$ and $\mathbf{d}$ and equating them to $\mathbf{0}$, we obtain the following

$$\left(\mathbf{d}'\boldsymbol{\Omega}\mathbf{d}\right)\boldsymbol{\Omega}\mathbf{c} - \left(\mathbf{c}'\boldsymbol{\Omega}\mathbf{d}\right)\boldsymbol{\Omega}\mathbf{d} - \lambda_1\mathbf{1} - \lambda_2\boldsymbol{\alpha} = \mathbf{0} \tag{30}$$

and

$$\left(\mathbf{c}'\boldsymbol{\Omega}\mathbf{c}\right)\boldsymbol{\Omega}\mathbf{d} - \left(\mathbf{c}'\boldsymbol{\Omega}\mathbf{d}\right)\boldsymbol{\Omega}\mathbf{c} - \lambda_3\mathbf{1} - \lambda_4\boldsymbol{\alpha} = \mathbf{0}. \tag{31}$$

Pre-multiplying (30) by $\mathbf{c}'$ and (31) by $\mathbf{d}'$ then simplifying by using the constraints, we get

$$\lambda_1 + \lambda_2\alpha_{s(k)} = |\Sigma| \tag{32}$$

and

$$\lambda_3 + \lambda_4\alpha_{t(k)} = |\Sigma|. \tag{33}$$

Next, pre-multiplying (30) by $\mathbf{d}'$ and (31) by $\mathbf{c}'$ then simplifying by using the constraints, we get

$$\lambda_1 + \lambda_2\alpha_{t(k)} = 0 \tag{34}$$

and

$$\lambda_3 + \lambda_4\alpha_{s(k)} = 0. \tag{35}$$

Solving (32) and (34) for λ_1 and λ_2, we get

$$\lambda_1 = -\frac{|\Sigma|\alpha_{t(k)}}{\alpha_{s(k)} - \alpha_{t(k)}} \text{ and } \lambda_2 = \frac{|\Sigma|}{\alpha_{s(k)} - \alpha_{t(k)}}.$$

Similarly, solving (33) and (35) for λ_3 and λ_4, we get

$$\lambda_3 = -\frac{|\Sigma|\alpha_s}{\alpha_{t(k)} - \alpha_{s(k)}} \text{ and } \lambda_4 = \frac{|\Sigma|}{\alpha_{t(k)} - \alpha_{s(k)}}.$$

Pre-multiplying (30) by $\mathbf{\Omega}^{-1}$ and then substituting the values of λ_1 and λ_2, we get

$$\left(\mathbf{d}'\mathbf{\Omega}\mathbf{d}\right)\mathbf{c} - \left(\mathbf{c}'\mathbf{\Omega}\mathbf{d}\right)\mathbf{d} = \frac{|\Sigma|}{\alpha_{s(k)} - \alpha_{t(k)}}\mathbf{\Omega}^{-1}\left(\boldsymbol{\alpha} - \alpha_{t(k)}\mathbf{I}\right). \tag{36}$$

Similarly, by pre-multiplying (31) by $\mathbf{\Omega}^{-1}$ and then substituting the values of λ_3 and λ_4, we get

$$\left(\mathbf{c}'\mathbf{\Omega}\mathbf{c}\right)\mathbf{d} - \left(\mathbf{c}'\mathbf{\Omega}\mathbf{d}\right)\mathbf{c} = \frac{|\Sigma|}{\alpha_{t(k)} - \alpha_{s(k)}}\mathbf{\Omega}^{-1}\left(\boldsymbol{\alpha} - \alpha_{s(k)}\mathbf{I}\right). \tag{37}$$

Let us denote $\boldsymbol{\alpha}_{s(k)} = \boldsymbol{\alpha} - \alpha_{s(k)}\mathbf{I} = \left(\alpha_{1(k)} - \alpha_{s(k)}, \ldots, \alpha_{n(k)} - \alpha_{s(k)}\right)'$ and $\boldsymbol{\alpha}_{t(k)} = \boldsymbol{\alpha} - \alpha_{t(k)}\mathbf{I} = \left(\alpha_{1(k)} - \alpha_{t(k)}, \ldots, \alpha_{n(k)} - \alpha_{t(k)}\right)'$. Then (36) and (37) can be jointly expressed as

$$\begin{bmatrix}\mathbf{c} & \mathbf{d}\end{bmatrix}\begin{bmatrix}\mathbf{d}'\mathbf{\Omega}\mathbf{d} & -\mathbf{c}'\mathbf{\Omega}\mathbf{d} \\ -\mathbf{c}'\mathbf{\Omega}\mathbf{d} & \mathbf{c}'\mathbf{\Omega}\mathbf{c}\end{bmatrix} = \frac{|\Sigma|}{\alpha_{t(k)} - \alpha_{s(k)}}\mathbf{\Omega}^{-1}\begin{bmatrix}-\boldsymbol{\alpha}_{t(k)} & \boldsymbol{\alpha}_{s(k)}\end{bmatrix}.$$

Then we get

$$\begin{bmatrix}\mathbf{c} & \mathbf{d}\end{bmatrix} = \frac{\mathbf{\Omega}^{-1}}{\alpha_{t(k)} - \alpha_{s(k)}}\begin{bmatrix}-\boldsymbol{\alpha}_{t(k)} & \boldsymbol{\alpha}_{s(k)}\end{bmatrix}\Sigma.$$

Then the solution for **c** and **d** are

$$\mathbf{c} = \frac{1}{\alpha_{t(k)} - \alpha_{s(k)}}\left(-\Sigma_{11}\mathbf{\Omega}^{-1}\boldsymbol{\alpha}_{t(k)} + \Sigma_{12}\mathbf{\Omega}^{-1}\boldsymbol{\alpha}_{s(k))}\right) \tag{38}$$

and

$$\mathbf{d} = \frac{1}{\alpha_{t(k)} - \alpha_{s(k)}}\left(-\Sigma_{12}\mathbf{\Omega}^{-1}\boldsymbol{\alpha}_{t(k)} + \Sigma_{22}\mathbf{\Omega}^{-1}\boldsymbol{\alpha}_{s(k)}\right). \tag{39}$$

Now by using the unbiasedness condition, $\mathbf{c}'\mathbf{1} = 1$ and $\mathbf{c}'\boldsymbol{\alpha} = \alpha_{s(k)}$, we obtain

$$\frac{1}{\alpha_{t(k)} - \alpha_{s(k)}}\left[-\Sigma_{11}\left(\boldsymbol{\alpha}_{t(k)}{}'\mathbf{\Omega}^{-1}\mathbf{1}\right) + \Sigma_{12}\left(\boldsymbol{\alpha}_{s(k)}{}'\mathbf{\Omega}^{-1}\mathbf{1}\right)\right] = 1 \tag{40}$$

and

$$\frac{1}{\alpha_{t(k)} - \alpha_{s(k)}}\left[-\Sigma_{11}\left(\boldsymbol{\alpha}_{t(k)}{}'\mathbf{\Omega}^{-1}\boldsymbol{\alpha}\right) + \Sigma_{12}\left(\boldsymbol{\alpha}_{s(k)}{}'\mathbf{\Omega}^{-1}\boldsymbol{\alpha}\right)\right] = \alpha_{s(k)}. \tag{41}$$

Solving (40) and (41) for Σ_{11} and Σ_{12}, we get

$$\Sigma_{11} = -\left(\alpha_{t(k)} - \alpha_{s(k)}\right)\frac{\boldsymbol{\alpha}_{s(k)}{}'\mathbf{\Omega}^{-1}\boldsymbol{\alpha}_{s(k)}}{\left(\boldsymbol{\alpha}_{t(k)}{}'\mathbf{\Omega}^{-1}\mathbf{1}\boldsymbol{\alpha}_{s(k)}{}' - \boldsymbol{\alpha}_{s(k)}{}'\mathbf{\Omega}^{-1}\mathbf{1}\boldsymbol{\alpha}_{t(k)}{}'\right)\mathbf{\Omega}^{-1}\boldsymbol{\alpha}} \tag{42}$$

and

$$\Sigma_{12} = -\left(\alpha_{t(k)} - \alpha_{s(k)}\right)\frac{\boldsymbol{\alpha}_{(t(k))}{}'\mathbf{\Omega}^{-1}\boldsymbol{\alpha}_{s(k)}}{\left(\boldsymbol{\alpha}_{t(k)}{}'\mathbf{\Omega}^{-1}\mathbf{1}\boldsymbol{\alpha}_{s(k)}{}' - \boldsymbol{\alpha}_{s(k)}{}'\mathbf{\Omega}^{-1}\mathbf{1}\boldsymbol{\alpha}_{t(k)}{}'\right)\mathbf{\Omega}^{-1}\boldsymbol{\alpha}}. \tag{43}$$

Let us denote

$$\mathbf{\Omega} = \begin{pmatrix}\sigma_{11(k)} & \sigma_{12(k)} & \cdots & \sigma_{1n(k)} \\ \sigma_{21(k)} & \sigma_{22(k)} & \cdots & \sigma_{2n(k)} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1(k)} & \sigma_{n2(k)} & \cdots & \sigma_{nn(k)}\end{pmatrix}$$

and

$$\mathbf{\Omega}^{-1} = \begin{pmatrix} \sigma^{11(k)} & \sigma^{12(k)} & \cdots & \sigma^{1n(k)} \\ \sigma^{21(k)} & \sigma^{22(k)} & \cdots & \sigma^{2n(k)} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^{n1(k)} & \sigma^{n2(k)} & \cdots & \sigma^{nn(k)} \end{pmatrix}.$$

Then

$$\mathbf{\Omega}^{-1}\mathbf{1} = \begin{pmatrix} \sigma^{11(k)} + \sigma^{12(k)} + \cdots + \sigma^{1n(k)} \\ \sigma^{21(k)} + \sigma^{22(k)} + \cdots + \sigma^{2n(k)} \\ \vdots \\ \sigma^{n1(k)} + \sigma^{n2(k)} + \cdots + \sigma^{nn(k)} \end{pmatrix} = \begin{pmatrix} T_{1(k)} \\ T_{2(k)} \\ \vdots \\ T_{n(k)} \end{pmatrix},$$

where T_i is the sum of the entries in the ith row of matrix Ω^{-1}. Similarly

$$\mathbf{\Omega}^{-1}\boldsymbol{\alpha} = \begin{pmatrix} \alpha_{1(k)}\sigma^{11(k)} + \alpha_{2(k)}\sigma^{12(k)} + \cdots + \alpha_{n(k)}\sigma^{1n(k)} \\ \alpha_{1(k)}\sigma^{21(k)} + \alpha_{2(k)}\sigma^{22(k)} + \cdots + \alpha_{n(k)}\sigma^{2n(k)} \\ \vdots \\ \alpha_{1(k)}\sigma^{n1(k)} + \alpha_{2(k)}\sigma^{n2(k)} + \cdots + \alpha_{n(k)}\sigma^{nn(k)} \end{pmatrix} = \begin{pmatrix} S_{1(k)} \\ S_{2(k)} \\ \vdots \\ S_{n(k)} \end{pmatrix}.$$

Thus we obtain

$$\boldsymbol{\alpha_{t(k)}}'\mathbf{\Omega}^{-1}\mathbf{1}\boldsymbol{\alpha_{s(k)}}' = \left(\sum_{i=1}^{n} T_{i(k)}\left(\alpha_{i(k)} - \alpha_{t(k)}\right)\right)\left[\alpha_{1(k)} - \alpha_{s(k)} \quad \alpha_{2(k)} - \alpha_{s(k)} \quad \cdots \quad \alpha_{n(k)} - \alpha_{s(k)}\right] \tag{44}$$

and

$$\boldsymbol{\alpha_{s(k)}}'\mathbf{\Omega}^{-1}\mathbf{1}\boldsymbol{\alpha_{t(k)}}' = \left(\sum_{i=1}^{n} T_{i(k)}\left(\alpha_{i(k)} - \alpha_{s(k)}\right)\right)\left[\alpha_{1(k)} - \alpha_{t(k)} \quad \alpha_{2(k)} - \alpha_{t(k)} \quad \cdots \quad \alpha_{n(k)} - \alpha_{t(k)}\right]. \tag{45}$$

Therefore,

$$\begin{aligned} \boldsymbol{\alpha_{t(k)}}'\mathbf{\Omega}^{-1}\mathbf{1}\boldsymbol{\alpha_{s(k)}}' - \boldsymbol{\alpha_{s(k)}}'\mathbf{\Omega}^{-1}\mathbf{1}\boldsymbol{\alpha_{t(k)}}' &= \left(\alpha_{s(k)} - \alpha_{t(k)}\right) \\ &\times \left(\alpha_{1(k)}T - T^*, \quad \alpha_{2(k)}T - T^*, \quad \cdots, \quad \alpha_{n(k)}T - T^*\right) \end{aligned}$$

where $T = \sum_{i=1}^{n} T_{i(k)}$ and $T^* = \sum_{i=1}^{n} T_{i(k)}\alpha_{i(k)}$. We observe that $T = \mathbf{1}'\mathbf{\Omega}^{-1}\mathbf{1}$, $T^* = \boldsymbol{\alpha}'\mathbf{\Omega}^{-1}\mathbf{1}$, $\sum_{i=1}^{n} S_{i(k)} = \mathbf{1}'\mathbf{\Omega}^{-1}\boldsymbol{\alpha} = \boldsymbol{\alpha}'\mathbf{\Omega}^{-1}\mathbf{1}$ and $\sum_{i=1}^{n} \alpha_i S_{i(k)} = \boldsymbol{\alpha}'\mathbf{\Omega}^{-1}\boldsymbol{\alpha}$. Then

$$\begin{aligned} \left(\boldsymbol{\alpha_{t(k)}}'\mathbf{\Omega}^{-1}\mathbf{1}\boldsymbol{\alpha_{s(k)}}' - \boldsymbol{\alpha_{s(k)}}'\mathbf{\Omega}^{-1}\mathbf{1}\boldsymbol{\alpha_{t(k)}}'\right)\mathbf{\Omega}^{-1}\boldsymbol{\alpha} &= \left(\alpha_{s(k)} - \alpha_{t(k)}\right) \\ &\times \left[\alpha_{1(k)}T - T^* \quad \alpha_{2(k)}T - T^* \quad \cdots \quad \alpha_{n(k)}T - T^*\right] \\ &\times \begin{pmatrix} S_{1(k)} \\ S_{2(k)} \\ \vdots \\ S_{n(k)} \end{pmatrix} \\ &= \left(\alpha_{s(k)} - \alpha_{t(k)}\right)\left(T\sum_{i=1}^{n} \alpha_{i(k)}S_{i(k)} - T^*\sum_{i=1}^{n} S_{i(k)}\right) \\ &= \left(\alpha_{s(k)} - \alpha_{t(k)}\right)\left[(\boldsymbol{\alpha}'\mathbf{\Omega}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\mathbf{\Omega}^{-1}\mathbf{1}) - (\boldsymbol{\alpha}'\mathbf{\Omega}^{-1}\mathbf{1})^2\right]. \end{aligned}$$

Hence we can write

$$\Sigma_{11} = \frac{\boldsymbol{\alpha_{s(k)}}'\mathbf{\Omega}^{-1}\boldsymbol{\alpha_{s(k)}}}{(\boldsymbol{\alpha}'\mathbf{\Omega}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\mathbf{\Omega}^{-1}\mathbf{1}) - (\boldsymbol{\alpha}'\mathbf{\Omega}^{-1}\mathbf{1})^2} \tag{46}$$

and

$$\Sigma_{12} = \frac{\boldsymbol{\alpha}_{t(k)}{}'\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha}_{s(k)}}{(\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\boldsymbol{\Omega}^{-1}\mathbf{1}) - (\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\mathbf{1})^2}. \tag{47}$$

Now by using the unbiasedness condition, $\mathbf{d}'\mathbf{1} = 1$ and $\mathbf{d}'\boldsymbol{\alpha} = \alpha_t$ and proceeded exactly in the same way, we obtain

$$\Sigma_{12} = \frac{\boldsymbol{\alpha}_{t(k)}{}'\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha}_{s(k)}}{(\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\boldsymbol{\Omega}^{-1}\mathbf{1}) - (\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\mathbf{1})^2} \tag{48}$$

and

$$\Sigma_{22} = \frac{\boldsymbol{\alpha}_{t(k)}{}'\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha}_{t(k)}}{(\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\boldsymbol{\Omega}^{-1}\mathbf{1}) - (\boldsymbol{\alpha}'\boldsymbol{\Omega}^{-1}\mathbf{1})^2}. \tag{49}$$

Substituting the value of Σ_{11} and Σ_{12} in (38) and simplyfying, we obtain the explicit expression for the coefficienr vector $\mathbf{c}$. Similary, substituting the value of Σ_{12} and Σ_{22} in (39) and simplyfying, we obtain the explicit expression for the coefficienr vector $\mathbf{d}$. Hence the theorem. ■

Theorem 2. Let $\mathbf{R}_{n(k)} = \left(R_{1(k)}, R_{2(k)}, \ldots, R_{n(k)}\right)'$ be the $1 \times n$ vector of first n observed upper k-record values arising from GP(μ, σ, ψ) with $E\left(\mathbf{R}_{n(k)}\right) = \mu\mathbf{1} + \Sigma\boldsymbol{\alpha}$ and $D\left(\mathbf{U}_{n(k)}\right) = \sigma^2\boldsymbol{\Omega}$. Then $\left(R^{**}_{s(k)}, R^{**}_{t(k)}\right)$, for $n < s < t$, is the joint BLIP of the pair $\left(R_{s(k)}, R_{t(k)}\right)$ where $R^{**}_{s(k)} = \boldsymbol{\theta}'\mathbf{R}_{n(k)}$ and $R^{**}_{t(k)} = \boldsymbol{\phi}'\mathbf{R}_{n(k)}$, where the coefficients $\boldsymbol{\theta}' = (\theta_1, \theta_2, \ldots, \theta_n)$ and $\boldsymbol{\phi}' = (\phi_1, \phi_2, \ldots, \phi_n)$ are given by

$$\boldsymbol{\theta} = \Lambda^{-1}\Delta_s \tag{50}$$

and

$$\boldsymbol{\phi} = \Lambda^{-1}\Delta_t \tag{51}$$

where $\Lambda = \boldsymbol{\Omega} + (\boldsymbol{\alpha} + \rho\boldsymbol{I_n})(\boldsymbol{\alpha} + \rho\boldsymbol{I_n})'$, $\Delta_s = \boldsymbol{\omega_{s(k)}} + (\alpha_s + \rho)(\boldsymbol{\alpha} + \rho\boldsymbol{I_n})$ and $\Delta_t = \boldsymbol{\omega_{t(k)}} + (\alpha_t + \rho)(\boldsymbol{\alpha} + \rho\boldsymbol{I_n})$.

Proof. The MSPE matrix is given by

$$\begin{pmatrix} E_1 & E_3 \\ E_3 & E_2 \end{pmatrix}.$$

where

$$\begin{aligned} E_1 &= E\left[\left(\tilde{R}_{s(k)} - R_{s(k)}\right)^2\right] \\ &= \sigma^2\left[\rho^2\left(\boldsymbol{\theta}'\mathbf{1} - 1\right)^2 + 2\rho\left(\boldsymbol{\theta}'\mathbf{1} - 1\right)\left(\boldsymbol{\theta}'\boldsymbol{\alpha} - \alpha_s\right)\right] \\ &+ \sigma^2\left[\boldsymbol{\theta}'\boldsymbol{\Omega}\boldsymbol{\theta} - 2\boldsymbol{\theta}'\boldsymbol{\omega_{s(k)}} + \omega_{ss} + \left(\boldsymbol{\theta}'\boldsymbol{\alpha} - \alpha_s\right)^2\right]. \end{aligned} \tag{52}$$

$$\begin{aligned} E_2 &= E\left[\left(\tilde{R}_{t(k)} - R_{t(k)}\right)^2\right] \\ &= \sigma^2\left[\rho^2\left(\boldsymbol{\phi}'\mathbf{1} - 1\right)^2 + 2\rho\left(\boldsymbol{\phi}'\mathbf{1} - 1\right)\left(\boldsymbol{\phi}'\boldsymbol{\alpha} - \alpha_t\right)\right] \\ &+ \sigma^2\left[\boldsymbol{\phi}'\boldsymbol{\Omega}\boldsymbol{\phi} - 2\boldsymbol{\phi}'\boldsymbol{\omega_{t(k)}} + \omega_{tt} + \left(\boldsymbol{\phi}'\boldsymbol{\alpha} - \alpha_t\right)^2\right]. \end{aligned} \tag{53}$$

$$\begin{aligned} E_3 &= E\left[\left(\tilde{R}_{s(k)} - R_{s(k)}\right)\left(\tilde{R}_{t(k)} - R_{t(k)}\right)\right] \\ &= \sigma^2\left[\rho^2\left(\boldsymbol{\theta}'\mathbf{1} - 1\right)\left(\boldsymbol{\phi}'\mathbf{1} - 1\right) + \rho\left(\boldsymbol{\theta}'\mathbf{1} - 1\right)\left(\boldsymbol{\phi}'\boldsymbol{\alpha} - \alpha_t\right) + \rho\left(\boldsymbol{\alpha}'\boldsymbol{\alpha} - \alpha_s\right)\left(\boldsymbol{\phi}'\mathbf{1} - 1\right)\right] \\ &+ \sigma^2\left[\boldsymbol{\theta}'\boldsymbol{\Omega}\boldsymbol{\phi} - \boldsymbol{\theta}'\boldsymbol{\omega_{t(k)}} - \boldsymbol{\phi}'\boldsymbol{\omega_{s(k)}} + \omega_{st} + \left(\boldsymbol{\theta}'\boldsymbol{\alpha} - \alpha_s\right)\left(\boldsymbol{\phi}'\boldsymbol{\alpha} - \alpha_t\right)\right]. \end{aligned} \tag{54}$$

The determinant of the MSEP matrix is $E = E_1E_2 - E_3^2$ which needs to be minimized with respect to $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$. Differentiating E with respect to $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ and equating them to $\mathbf{0}$, we obtain the following

$$(\Lambda\boldsymbol{\theta} - \Delta_s)\, E_2 - (\Lambda\boldsymbol{\phi} - \Delta_t)\, E_3 = \mathbf{0} \tag{55}$$

and

$$(\Lambda\boldsymbol{\phi} - \Delta_t)\, E_1 - (\Lambda\boldsymbol{\theta} - \Delta_s)\, E_3 = \mathbf{0}, \tag{56}$$

where Λ, Δ_s and Δ_t are defined in the statement of the theorem. From (55) and (56), we obtain

$$(\Lambda\boldsymbol{\theta} - \Delta_s)\left(E_1E_2 - E_3^2\right) = \mathbf{0}, \tag{57}$$

Assuming $E_1E_2 - E_3^2 \neq 0$, we readly obtain $\boldsymbol{\theta} = \Lambda^{-1}\Delta_s$. Prceeding similary, from (55) and (56), we obtain $\boldsymbol{\phi} = \Lambda^{-1}\Delta_t$. Hence the proof. ■

3.3. Comparison between BLUP and BLIP

In this Subsection, We define two types of relative efficiencies for the performance of joint BLIP and joint BLUP based on the MSPE matrix as follows.

$$\begin{aligned} \text{D-Efficiency}\,(\epsilon_1) &= \frac{\text{Determinant of the MSEP matrix of } \left(\tilde{U}_{s(k)}, \tilde{U}_{t(k)}\right)}{\text{Determinant of the MSEP matrix of } \left(U^{**}_{s(k)}, (U^{**}_{t(k)}\right)} \\ &= \frac{E_1E_2 - E_3^2}{\Sigma_{11}\Sigma_{22} - \Sigma_{12}^2} \end{aligned} \tag{58}$$

and

$$\begin{aligned} \text{T-Efficiency}\,(\epsilon_2) &= \frac{\text{Trace of the MSEP matrix of } \left(\tilde{U}_{s(k)}, \tilde{U}_{t(k)}\right)}{\text{Trace of the MSEP matrix of } \left(U^{**}_{s(k)}, (U^{**}_{t(k)}\right)} \\ &= \frac{E_1 + E_2}{\Sigma_{11} + \Sigma_{22}}. \end{aligned} \tag{59}$$

The quantity ρ plays an important role in the relative performance of the BLIP against the BLUP. Note that the range of ρ is $(-\infty, +\infty)$. However, due to symmetry, we can focus on the interpretation of the relative efficiencies $(0, +\infty)$. Both ϵ_1 and ϵ_2 always possess an unique maximum ρ^* which are calculated numerically and analytically. If both efficiencies are always greater than 1 indicating that BLUP performs better than BLIP. If for some $\rho_* > 0$, BLUP performs better in an interval $(0, \rho_*)$ and then BLIP performs better in the interval (ρ_*, ∞) then for overall comparative assessment, we define an integrated efficiency measure (IEM) which is simply an average of all efficiencies calculated on a finite interval $(0, \rho_{max})$ for ρ. If both ϵ_1 and ϵ_2 are less than 1 which indicates that BLIP has overall better performance than BLUP.

4. Illustration with Real Data

In this section, we consider a numerical study to illustrate how the proposed predictors work in practical situation. We consider the following dataset by [8] represents the exceedances $\left(kg\ mm^{-2}\right)$ of the testing thresholds in tensile-strength tests for a random sample size 15 nylon carpet fibres. The dataset is

$$0.051, 0.140, 0.365, 0.561, 0.030, 0.268, 0.184, 0.100, 0.876, 0.092, 0.011, 0.200, 0.518, 0.338, 0.056$$

The Kolmogorov-Smirnov goodness- of- fit test shows that the above data fits GPD$(0, 0.283, 0.1)$ very well with the Kolmogorov-Smirnov test statistic is 0.107 and p-value 0.988.

The upper k- record values extracted from the dataset for $k = 2, 3$ are given by the following table.

Table 9: *Upper k- record values extracted from the dataset*

n	1	2	3	4	5
$R_{n(2)}$	0.051	0.14	0.365	0.561	-
$R_{n(3)}$	0.051	0.14	0.268	0.315	0.518

The marginal BLUP and BLIP of future upper k- record values based on the n observed upper k-record values along with their estimated MSPE are given by the following table.

Table 10: *Marginal BLUP and BLIP for the real dataset with their estimated MSPE*

k	n	s	Marginal Predictors		MSPE	
			$R^*_{s(k)}$	$R^{**}_{s(k)}$	BLUP	BLIP
2	4	5	0.7470	0.7060	0.0580	0.0260
3	5	6	0.6400	0.5970	0.7820	0.0520

From Table 10, we can observe that the marginal BLIP performs better than the marginal BLUP in the sense of estimated MSPE. The Joint BLUP and BLIP for the real dataset along with the Determinant and Trace efficiencies are given in Table 11.

Table 11: *Joint BLUP and BLIP for the real dataset with Determinant and Trace efficiencies*

k	n	(s,t)	Joint Predictors		ϵ_1	ϵ_2
			$\left(R^*_{s(k)}, R^*_{t(k)}\right)$	$\left(R^{**}_{s(k)}, R^{**}_{t(k)}\right)$		
2	4	(6,7)	$(0.9378, 1.1441)$	$(0.8173, 0.9556)$	0.0009	0.0246
		(6,8)	$(0.9378, 1.3611)$	$(0.8173, 1.1013)$	0.0011	0.1078
		(7,9)	$(1.4412, 1.6117)$	$(0.9556, 1.2546)$	0.0056	0.1173
3	5	(7,8)	$(0.7854, 0.9333)$	$(0.7097, 0.8106)$	0.0026	0.0812
		(7,9)	$(0.7854, 1.0853)$	$(0.7097, 0.9150)$	0.0053	0.1531
		(8,9)	$(0.9333, 1.0853)$	$(0.8106, 0.9151)$	0.0084	0.1851

Since both ϵ_1 and ϵ_2 are less than 1 indicate that BLIP has overall better performance than BLUP.

5. Conclusion

Serious difficulties for the statistical inference based on records arise due to the fact that the occur rences of record data are very rare in practical situations and the expected waiting time is infinite for every record after the first. These problems are avoided if we consider the model of k-record statistics. In this work we considered the upper k- record values arising from generalized Pareto distribution. We have obtained the BLUEs and BLISs of the location and scale parameters of a generalized Pareto distribution based on the observed upper k-record values. The marginal BLUP and BLIP of future upper k-record values and the joint BLUP and BLIP of pair of future upper k-record values are also derived. A real dataset is used to illustrate the inferential procedures developed in this paper and concluded that the BLIP performs well in terms of estimated MSPE and relative efficiencies.

References

[1] Ahmadi, J., MirMostafaee, S. M. T. K. and Balakrishnan, N. (2011). Bayesian prediction of order statistics based on k-record values from exponential distribution. *Statistics*, 45(4): 375-387.

[2] Ahsanullah, M.and Nevzorov, V. B. (2015). *Estimation of Parameters and Predictions of Records. In Records via Probability Theory* (pp. 143-181). Paris: Atlantis Press.

[3] Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1998). *Records*, John Wiley and Sons, New York.

[4] Balakrishnan, N. and Bhattacharya, R. (2022). D-optimal joint best linear unbiased prediction of order statistics: Joint BLUP. *Metrika*, 85(2): 253-267.

[5] Balakrishnan, N. and Bhattacharya, R. (2022). Simultaneous best linear invariant prediction of future order statistics for location-scale and scale families and associated optimality properties. *Statistics*, 56(6): 1409-1426.

[6] Chandler, K. N. (1952). The distribution and frequency of record values. *Journal of the Royal Statistical Society, Series B*, 14: 220-228.

[7] Dziubdziela, W. and Kopociński, B. (1976). Limiting properties of the k-th record values. *Applicationes Mathematicae*, 15: 187-190.

[8] Grimshaw, S. D. (1993). Computing maximum likelihood estimates for the generalized Pareto distribution. *Technometrics*, 35(2): 185-191.

[9] Gulati, S.and Padgett, W. J. (2003). *Parametric and Nonparametric Inference from Record-Breaking Data*. Springer, New York.

[10] Jamkhaneh, M. B., MirMostafaee, S. M. T. K. and Jadidi, M. (2024). On the xgamma k-record values and associated inference. *Annals of Data Science*, 1-29 (in press).

[11] Mann, N. R. (1969). Optimum estimators for linear functions of location and scale parameters. *The Annals of Mathematical Statistics*, 2149-2155.

[12] Muraleedharan, L. and Chacko, M. (2022). Estimation and prediction based on k-record values from a generalized Pareto distribution. *International Journal of Mathematics and Statistics*, 23(1): 59-75.

[13] Ormoz, E., Akhter, Z., Alam, M. and MirMostafae, S. M. T. K. (2023). Lower k-record values from unit-Gompertz distribution and associated inference. *arXiv preprint arXiv*:2309.13279.

[14] Raqab, M. (2002). Inference for generalized exponential distribution based on record statistics. *Journal of Statistical Planning and Inference*, 104(2): 339-350.

[15] Raqab, M. Z. (1996). Linear predictors of future order statistics. *Communications in Statistics-Theory and Methods*, 25(7): 1553-1568.

[16] Shafay, A. R., Balakrishnan, N. and Ahmadi, J. (2017). Bayesian prediction of order statistics with fixed and random sample sizes based on k-record values from Pareto distribution. *Communications in Statistics-Theory and Methods*, 46(2): 721-735.

[17] Singh, S., Mani Tripathi, Y. and Wu, S. J. (2017). Bayesian estimation and prediction based on lognormal record values. *Journal of Applied Statistics*, 44(5): 916-940.

[18] Soliman, A., Amin, E. A. and Abd-El Aziz, A. A. (2010). Estimation and prediction from inverse Rayleigh distribution based on lower record values. *Applied Mathematical Sciences*, 4(62): 3057-3066.

[19] Sultan, K. S., Al-Dayian, G. R. and Mohammad, H. H. (2008). Estimation and prediction from gamma distribution based on record values. *Computational Statistics and Data Analysis*, 52(3): 1430-1440.

[20] Sultan, K. S. and Moshref, M. E. (2000). Record values from generalized Pareto distribution and associated inference. *Metrika*, 51(2): 105-116.

[21] Vidović, Z. (2019). Bayesian prediction of order statistics based on k-record values from a generalized exponential distribution. *Stats*, 2(4): 447-456.

[22] Zhao, X., Wei, S., Cheng, W., Zhang, P., Zhang, Y. and Xu, Q. (2023). Upper record values from the generalized Pareto distribution and associated statistical inference. *Communications in Statistics-Simulation and Computation*, 52(2): 369-391.