STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 202x, pp 0–19.
Published online in International Academic Press (www.IAPress.org)

Prediction Methods for Future Record Values from Two-Parameter Kies Distribution

Nesreen M. Al-Olaimat [1,*], Hatim Solayman Migdadi[2], Husam A. Bayoud [3], Mohammad Z. Raqab [4]

[1] *Department of Applied Science, Ajloun College, Al-Balqa Applied University, Ajoun 26816, Jordan*
E-mail:n.malolaimat@bau.edu.jo; nesrreenolaimat@gmail.com. ORCID: https://orcid.org/0000-0003-2862-3739
[2] *Department of Mathematics, Faculty of Science, The Hashemite University, Zarqa 13133, Jordan*
Email: hatims@hu.edu.jo. ORCID: https://orcid.org/0000-0002-1989-5153
[3] *Department of Mathematics, College of Sciences and Humanities, Fahad Bin Sultan University, Tabuk, Saudi Arabia*
E-mail: hbayoud@fbsu.edu.sa; husam.awni@yahoo.com. ORCID: https://orcid.org/0000-0002-0066-7206
[4] *Department of Mathematics, Department of Mathematics, The University of Jordan, Amman 11942, Jordan*
E-mail: mraqab@ju.edu.jo; mraqab4@gmail.com. ORCID: https://orcid.org/0000-0002-3047-1466

Abstract In this paper, we address the prediction problem of the future records based on observed data from two-parameter, shape and scale parameter, Kies distribution. Various point predictors, including maximum likelihood, conditional median, best unbiased and Bayesian predictors of the future records are derived. The corresponding prediction intervals are developed using pivotal quantity, Highest Conditional Density, Shortest Length and Bayesian prediction intervals. The Monte Carlo algorithm is employed to compute simulation consistent Bayesian prediction intervals for future unobserved records. The performance of the obtained point predictors and prediction intervals are compared via experimental numerical simulation. The criteria considered for comparison purposes are mean square prediction error and prediction bias for point predictors and coverage probability and the average length for prediction intervals. A real and simulated data sets are performed for illustrative purposes.

Keywords Kies Distribution, Records, Maximum Likelihood Predictor, Conditional Median Predictor, Best Unbiased Predictor, Bayesian Prediction

AMS 2010 subject classifications 62F10, 62F15, 60E05, 62G30, 62M20

DOI: 10.19139/soic-2310-5070-2123

1. Introduction

In connection with the study of braking strength of glass, Kies (1958)[15] proposed a functional form of the Weibull distribution(WD), which later became known in the literature as the Kies distribution. Kumar and Dharmaja (2014)[17] studied the properties of the Kies distribution along with its applications in the areas of medical and engineering sciences. Moreover, they shown that due to of flexibility of the bathtub-shaped hazard function, it provides a better alternative than other extended versions of the WD, namely the generalized Weibull (GW) distribution, modified Weibull (MW) distribution, beta Weibull (BW) distribution and beta generalized Weibull (BGW) distribution, for modeling the lifetime data sets. Furthermore, they considered the estimation of Kies parameters using maximum likelihood estimation method.
Kumar and Dharmaja (2013)[16] studied a special case of the Kies distribution, called the reduce Kies distribution (RKD), and showed that the RKD enjoys certain characteristic properties similar to those of the WD. In 2017, they introduced and studied a generalized version of the extended reduced Kies distribution, called the modified

*Correspondence to: Nesreen M. Al-Olaimat (Email: nmolaimat@anu.edu.jo). Department of Mathematics,Department of Mathematics, Faculty of Science, Ajloun National University, Ajloun 26810, Jordan.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

Kies distribution (MKD), see Kumar and Dharmaja (2017a)[18]. In addition, Kumar and Dharmaja (2017b)[19] proposed an exponentiated reduced Kies distribution with two parameters. Al-Olaimat et al. (2021)[5] derived some properties of record statistics from the two-parameter Kies distribution. They also studied the estimation of the two-parameter Kies distribution under record values using classical and Bayesian estimation methods, comparing the performance of these estimators for several samples using extensive simulations.
The cumulative distribution function (CDF), probability density function (PDF), hazard rate and cumulative hazard rate functions of the two-parameter Kies distribution,K(λ, β), are given by:

$$F(x;\lambda,\beta) = 1 - e^{-\lambda\left(\frac{x}{1-x}\right)^{\beta}}, \tag{1}$$

$$f(x;\lambda,\beta) = \frac{\beta\lambda x^{\beta-1}}{(1-x)^{\beta+1}} e^{-\lambda\left(\frac{x}{1-x}\right)^{\beta}}, \tag{2}$$

$$h(x;\lambda,\beta) = \frac{\beta\lambda x^{\beta-1}}{(1-x)^{\beta+1}}, \tag{3}$$

and

$$H(x;\lambda,\beta) = \lambda\left(\frac{x}{1-x}\right)^{\beta}, \tag{4}$$

respectively, where $0 < x < 1$, $\lambda > 0$ and $\beta > 0$. The Kies distribution has a bounded range, which makes it appropriate model for fitting real data sets with a bounded range. However, there are many situations where observations can only take values within a limited range, such as fractions, percentages or proportions. Papke and Wooldridge (1996)[22] pointed out that variables in many economic applications like the proportion of income spent on non-durable consumption, the fraction of total weekly hours spent on working, a fraction of land area allocated to agriculture and industry market shares are all bounded between zero and one. Furthermore, Genc (2013)[14] indicated that when the reliability is measured as a ratio or percentage, it is important to have models defined on the unit interval in order to have reasonable results.
Records play an important role in several fields of statistics which date back to Chandler (1952)[9], who first defined and provided the groundwork for the mathematical theory of records. Let $\{X_j, j \geq 1\}$ be a sequence of independent and identically distributed (iid) continuous random variables (r.v.'s) with CDF $F(x)$ and PDF $f(x)$. An observation X_j is defined to be an upper record if $X_j > X_i$ for every $j > i$, and an analogous definition can be given for lower records (with the inequality being reversed). By convention, the first record X_1 is called the trivial record because it is serves as both an upper and a lower record value simultaneously.
The set of the upper record values is given by the r.v.'s $X_{U(k)}$ for $k \geq 1$ where

$$U(1) = 1,\ U(k) = min\{j : j > U(k-1), X_j > X_{U(k-1)}\}.$$

Suppose we have a random sample (not ordered) of size n, say $\{X_1, X_2, ..., X_n\}$, the set

$$\left\{X_{U(1)} = X_1, X_{U(2)}, ..., X_{U(m)}\right\},$$

presents a set of upper record values with size $1 \leq m \leq n$ that is obtained from the random sample. The sequence $U(k),\ k \geq 1$ is called the sequence of upper record times. For simplicity, we denote the sequence of upper record values $\{X_{U(j)}\}_{j=1}^{m}$ by $\mathbf{Y} = \{Y_j\}_{j=1}^{m}$.
Record statistics arise in many practical fields, including hydrology, meteorology, sports and athletic events, wherein only records are usually considered. For example, record values are applied in estimating the strength of the material, predicting sports achievements, and the natural disasters. Al-Olaimat et al. (2021)[5] addressed the estimation problem for the two-parameter Kies distribution based on record data, specifically using Bayesian and non-Bayesian methods. For further details and applications on record statistics, readers may refer to Arnold et al. (1998)[6], Ahsanullah (2004)[1], Ahsanullah and Raqab (2006)[3] and Ahsanullah and Nevzorov (2015)[2].
The prediction problem of future events based on the past and present knowledge is of great interest in statistics. Considerable several of work has been done on prediction of record values. For example, Bayesian predictive distributions of future records from an exponential distribution provided by Dunsmore (1983)[12]. Nagaraja

(1988)[21] discussed the predictors of future records from three extreme value distributions. Awad and Raqab (2000)[7] considered the prediction interval problem of the future record from exponential distribution. Al-Hussaini and Ahmed (2003)[4] studied Bayesian prediction interval for the future generalized order statistics (including record values as a special case). The problem of Bayesian prediction of temperature records using the Pareto model was considered by Madi and Raqab (2004)[20]. Raqab et al. (2007)[25] considered the problem of predicting the future record values, either point or interval prediction, from the two-parameter Pareto distribution, based on the past record values observed. Dey et al. (2017)[11] discussed prediction intervals for future record values from the generalized Rayleigh distribution using both frequentist and Bayesian approaches. This study emphasizes the practical utility of record-based inference and highlights the comparative strengths of frequentist and Bayesian methods in predicting future records. Volovskiy and Kamps (2020)[27] investigated the point prediction of future upper record values for absolutely continuous distribution with a strictly increasing cumulative distribution functions. They derived a general predictor by maximizing the observed predictive likelihood function. Furthermore, they illustrated the results for exponential, extreme-value, and power-function distributions, comparing the performance of the obtained predictors against maximum likelihood predictors using the mean squared error and Pitman's measure of closeness criteria. Empacher et al. (2023)[13] studied the point prediction of future record values based on sequences of previous records using the maximum product of spacings method. Their study focused on the power function and Pareto distributions, examining both exact and approximate prediction intervals in terms of their expected lengths and coverage percentages. Their work emphasizes the growing importance of statistical predictions in sports analytics, an area that has traditionally relied on extreme value theory for forecasting athletic records. They applied their methods to various athletic data sets as well as to American football data and discussed the implications of their findings alongside the choice of underlying distributions.

This paper is motivated by the need to enhance predictive methodologies for future record values, specifically within the framework of the two-parameter Kies distribution. While previous research has made significant strides in the field of record statistics, there remains a notable lack of focus on this particular distribution, which possesses unique properties that make it well-suited for various applications. The main contributions of this study include deriving several predictive methods, including maximum likelihood, modified maximum likelihood, conditional median, best unbiased, and Bayesian predictors. in addition, this study presents a robust framework for obtaining prediction intervals, which enhances the reliability of our predictions. Using Monte Carlo simulation studies, comprehensive numerical comparisons of the proposed methods are conducted, validating their effectiveness with both simulated and real-world data.

The remainder of this paper is organized as follows: In Section(2) we derive the maximum likelihood, modified maximum likelihood, conditional median, best unbiased and Bayesian predictors for future records based on observed sample. In section(3) we propose various procedures for obtaining the prediction intervals. The Monte Carlo simulation study that conducts numerical comparisons are performed in section(5). Section(4) presents the numerical results from both simulated and real data sets for illustrative purpose. Finally, the conclusions of the paper are summarized in section(6).

2. Point Prediction

Let $y_s, s > m$ be the future record. This future record will be predicting via several point processers based on $\mathbf{Y} = (y_1, y_2, ..., y_m)$, for simplicity, let $R_i = \frac{y_i}{1-y_i}$, $i = 1, 2, ..., s$.

The prediction of future records y_s based on the a sequence of the first m observed records, $\mathbf{Y} = \{y_1, y_2, ..., y_m\}$, mainly depends on the conditional predictive density function of y_s given the observed record data. Using the Markovian property of record data, the conditional distribution of y_s given $data$, is just the conditional distribution of y_s given y_m, see Arnold et al. (1998),which has the pdf

$$f(y_s|y_m;\lambda,\beta) = \frac{[H(y_s) - H(y_m)]^{s-m-1}}{\Gamma(s-m)} \frac{f(y_s|\lambda,\beta)}{1 - F(y_m|\lambda,\beta)}, \tag{5}$$

Hence, using Eqs.(1), (2) and (4), Eq.(5) reduces to

$$f(y_s|y_m;\lambda,\beta)=\lambda^{s-m}\beta\frac{R_s^{\beta}}{y_s(1-y_s)}\frac{(R_s^{\beta}-R_m^{\beta})^{s-m-1}}{\Gamma(s-m)}e^{-\lambda[R_s^{\beta}-R_m^{\beta}]} \tag{6}$$

where $0<y_m<y_s<1$.

2.1. Maximum Likelihood Predictor

In this subsection, we will study point predict of $y_s, s>m$ using maximum likelihood predictor(MLP) method. Let $\mathbf{Y}=\{y_1,y_2,...,y_m\}$ be a sequence of observed records from a population with PDF $f(y_s;\theta)$ and $F(y_s;\theta)$ where $\theta=(\lambda,\beta)$, then the predictive likelihood function (PLF) of y_s, λ and β, which is given by Basak and Balakrishnan (2003)[8] as:

$$L(y_s;\theta,data)=\prod_{i=1}^{m}h(y_i,\theta)\frac{[H(y_s;\theta)-H(y_m;\theta)]^{s-m-1}}{\Gamma(s-m)}f(y_s;\theta). \tag{7}$$

Generally, if $\hat{y}_{MLP}=u(\mathbf{Y})$, $\hat{\lambda}=v(\mathbf{Y})$, and $\hat{\beta}=w(\mathbf{Y})$ are statistics for which

$$L(u(\mathbf{Y}),v(\mathbf{Y}),w(\mathbf{Y})|\mathbf{Y})=\underset{y_s,\lambda,\beta}{sup}\,L(y_s,\lambda,\beta|\mathbf{Y}), \tag{8}$$

then $u(\mathbf{Y})$ is said to be the MLP of y_s, $1<m<s$, and $v(\mathbf{Y})$ and $w(\mathbf{Y})$ are the predictive maximum likelihood estimators (PMLEs) of λ and β, respectively. Using the Eqs.(2), (3)and (4), Eq.(7) will be

$$L(y_s,\lambda,\beta)=\lambda^{s}\beta^{m+1}\prod_{i=1}^{m}\frac{R_i^{\beta}}{y_i(1-y_i)}\frac{(R_s^{\beta}-R_m^{\beta})^{s-m-1}}{\Gamma(s-m)}\frac{R_s^{\beta}}{y_s(1-y_s)}e^{-\lambda R_s^{\beta}} \tag{9}$$

Regardless of the constant terms, the predictive log-likelihood function is given by

$$\begin{aligned}\log(L(y_s,\lambda,\beta))\propto\; & s\log(\lambda)+(m+1)\log(\beta)+\beta\sum_{i=1}^{m}\log(R_i)\\ & +(s-m-1)\log(R_s^{\beta}-R_m^{\beta})\\ & +(\beta-1)\log(\frac{R_s}{1+R_s})-(\beta+1)\log(\frac{1}{1+R_s})-\lambda R_s^{\beta}\end{aligned} \tag{10}$$

By using Eq.(10), the predictive likelihood equations (PLEs) for y_s, λ and β are derived and presented, respectively, as follows:

$$\frac{\partial\log(L(y_s,\lambda,\beta))}{\partial\lambda}=\frac{s}{\lambda}-R_s^{\beta}=0 \tag{11}$$

$$\begin{aligned}\frac{\partial\log(L(y_s,\lambda,\beta))}{\partial\beta}=\; & \frac{m+1}{\beta}+\sum_{i=1}^{m}\log R_i+(s-m-1)\frac{R_s^{\beta}\log R_s-R_m^{\beta}\log R_m}{R_s^{\beta}-R_m^{\beta}}\\ & +(1-\lambda R_s^{\beta})\log R_s=0\end{aligned} \tag{12}$$

$$\begin{aligned}\frac{\partial\log(L(y_s,\lambda,\beta))}{\partial y_s}=\; & (s-m-1)\beta\frac{(\frac{y_s}{1-y_s})^{\beta}}{y_s(1-y_s)[(\frac{y_s}{1-y_s})^{\beta}-R_m^{\beta}]}+\frac{\beta+2y_s-1}{y_s(1-y_s)}\\ & -\lambda\beta\frac{{y_s}^{\beta-1}}{(1-y_s)^{\beta+1}}=0\end{aligned} \tag{13}$$

The PMLE of λ is obtained from Eq.(11) and it is given by

$$\hat{\lambda}=\frac{s}{R_s^{\beta}} \tag{14}$$

The PMLE of β, $\hat{\beta}$, and MLP of y_s, $\hat{y}_{MLP}$, can be obtained by substituting Eq.(14) in Eq.(12) and Eq.(13) then using numerical methods can be solve these simultaneous equations with respect to β and y_s.

2.2. *Modified Maximum Likelihood Predictor*

In practical application, the experimenters wish to use a simple and quick predictor, for this the modified maximum likelihood predictor (MMLP) of y_s is suggested. Therefore, the PLF of y_s, λ and β can be decomposed into product of two functions, the first one L_1 is the likelihood function of $\mathbf{Y}$, λ and β, which is viewed as a function of λ and β. The MLEs of λ and β were computed by Al-Olaimat et al. (2021)[5] and presented as follows:

$$\hat{\lambda} = \frac{m}{R_m^{\hat{\beta}}}, \tag{15}$$

and

$$\hat{\beta} = \frac{m}{\sum_{i=1}^{m-1} \log(\frac{R_m}{R_i})}, \tag{16}$$

respectively, For more details interested readers may refer to [5] Section (3). The second function L_2 is the conditional PDF of y_s given $\mathbf{Y}$, λ and β. Based on m observed record values, we compute the MLEs of λ and β, $\hat{\lambda}$ and $\hat{\beta}$, and substitute their values onto L_2. Using the modified L_2 we can easily find a MMLP of y_s. Therefor, the MMLP of y_s is obtained by solving the following equation:

$$\frac{1}{y_s(1-y_s)}\left[(s-m-1)\hat{\beta}\frac{R_s^{\hat{\beta}}}{R_s^{\hat{\beta}} - R_m^{\hat{\beta}}} + 2y_s - \hat{\lambda}\hat{\beta}R_s^{\hat{\beta}} + \hat{\beta} - 1\right] = 0, \tag{17}$$

where $y_s > y_m$. Since Eq.(17) can not be solved analytically, a numerical method is needed to compute the MMLP of y_s, $y_{M\hat{M}LP}$. For a special case when $s = m+1$, we can see $y_{M\hat{M}LP} = y_m$.

2.3. *Conditional Median Predictor*

Another possible predictor called conditional median predictor (CMP), is proposed in this subsection following the lines of Raqab (1992)[23]. A predictor $\hat{Y}_{CMP}$ is said to be the CMP of y_s if it is the median of the conditional distribution of y_s given $\mathbf{Y}$, i.e.,

$$P((y_s|y_m;\lambda,\beta) \leq \hat{Y}_{CMP}) = P((y_s|y_m;\lambda,\beta) \geq \hat{Y}_{CMP}) = \frac{1}{2}. \tag{18}$$

Consequently, assume $\hat{Y}_{CMP} = k(y_m;\lambda,\beta)$ which is a function of y_m, then from Eq.(6), we have

$$\int_{y_m}^{k(y_m,\lambda,\beta)} \lambda^{s-m}\beta\frac{R_s^{\beta}}{y_s(1-y_s)}\frac{(R_s^{\beta}-R_m^{\beta})^{s-m-1}}{\Gamma(s-m)}e^{-\lambda(R_s^{\beta}-R_m^{\beta})}dy_s = \frac{1}{2}$$

Setting $R_s^{\beta} - R_m^{\beta} = t$, we obtain

$$\int_0^{[\frac{(K(R_m,\lambda,\beta))^{\beta}}{1-(K(R_m,\lambda,\beta))^{\beta}} - R_m^{\beta}]} \frac{\lambda^{s-m}}{\Gamma(s-m)}t^{s-m-1}e^{-\lambda t}dt = \frac{1}{2}.$$

Thus

$$\hat{Y}_{CMP} = \frac{(Med(W)+R_m^{\beta})^{\frac{1}{\beta}}}{1+(Med(W)+R_m^{\beta})^{\frac{1}{\beta}}},$$

where $W \sim Gamma(s-m, \frac{1}{\lambda})$.
Assume that we are interested in predicting the first future prediction, i.e $s = m+1$, then $W \sim Exp(\lambda)$, with Median(W)=$\frac{1}{\lambda}\log 2$, therefore

$$\hat{Y}_{CMP} = \frac{(\log 2^{\frac{1}{\lambda}} + R_m^{\beta})^{\frac{1}{\beta}}}{1+(\log 2^{\frac{1}{\lambda}} + R_m^{\beta})^{\frac{1}{\beta}}},$$

2.4. *Best Unbiased Predictor*

The predictor $\hat{Y}_{BUP}$ of y_s, $s > m$ is called the best unbiased predictor (BUP)if the prediction error $\hat{Y}_{BUP} - y_s$ has a mean zero and it's prediction variance, var($\hat{Y}_{BUP} - y_s$), is less than or equal to that of any other unbiased predictor of y_s. For known λ and β, the BUP of y_s is given by

$$\hat{Y}_{BUP} = E(Y_s|Y_m;\lambda,\beta). \tag{19}$$

Therefore, using Eq.(6) and the binomial expansion, we can obtain $\hat{Y}_{BUP}$ as the following:

$$\begin{aligned}\hat{Y}_{BUP} &= \int_{y_m}^{1} y_s f(y_s|y_m;\lambda,\beta)dy_s \\ &= \int_{y_m}^{1} \lambda^{s-m}\beta R_s^{\beta}\frac{1}{1-y_s}\frac{(R_s^{\beta}-R_m^{\beta})^{s-m-1}}{\Gamma(s-m)}e^{-\lambda(R_s^{\beta}-R_m^{\beta})}dy_s \\ &= \frac{\lambda^{s-m}\beta}{\Gamma(s-m)}e^{-\lambda R_m^{\beta}}\sum_{i=0}^{s-m-1}\binom{s-m-1}{i}(-1)^{s-m-i-1}R_m^{\beta(s-m-i-1)} \\ &\times \int_{y_m}^{1}\frac{R_s^{\beta(i+1)}}{1-y_s}e^{-\lambda R_s^{\beta}}dy_s,\end{aligned} \tag{20}$$

when the parameters λ and β are unknown, the BUP of y_s can be approximated by replacing both the parameters λ and β by their corresponding MLEs.

2.5. *Bayesian Prediction*

In this section, we use the Bayesian approach to predict unknown future records based on the observed current sequence from $K(\lambda,\beta)$ distribution. The important task in Bayesian inference is the selection of an appropriate prior for the unknown parameter. Therefore, we want to assume the parameters λ and β are independent and follow gamma distributions; namely $Gamma(a_1, b_1)$ and $Gamma(a_2, b_2)$, respectively, where the hyper parameters a_1, b_1,a_2 and b_2 are preselected and non negative real numbers that are chosen to reflect prior knowledge about λ and β. The choice of the gamma distribution is done for illustrative purposes only and any other suitable prior can be used instead of this. Moreover, dependent priors can also be assumed. Therefore, the joint prior distribution of λ and β is obtained as follows:

$$g(\lambda,\beta) \propto \lambda^{a_1-1}e^{-b_1\lambda}\beta^{a_2-1}e^{-b_2\beta}. \tag{21}$$

In order to conduct a Bayesian analysis, we want to use the squared error(SE)loss function, the most commonly used loss function, which is given as follows $L(\hat{\theta},\theta) = (\hat{\theta}-\theta)^2$. The SE loss function is a symmetric loss function, it leads to the identical penalization for overestimation and underestimation, so that it is not an appropriate in some practical situations. Therefore, several loss functions have been introduced to handle such a problem, for example, Varian (1975) [26] proposed an extremely helpful asymmetric loss function it is called linear exponential(LINEX) loss function, which is given as follows $L(\hat{\theta},\theta) = b[e^{\nu(\hat{\theta}-\theta)} - \nu(\hat{\theta}-\theta) - 1]$ where $\nu \neq 0$ is the shape parameter and b is the scale parameter, in our study we assumed $b = 1$. The LINEX loss function reduce to SE loss function when ν close to zero and therefore almost symmetric. Furthermore, when $\nu > 0$, overestimation more serious than underestimation and when $\nu < 0$, underestimation more serious than overestimation.
Let $\mathbf{Y} = \{y_1, y_2, \ldots, y_m\}$ be the first m upper record values arising from a sequence of iid K(λ, β) distribution with CDF, PDF and hazard rate being defined in Eqs. (1), (2) and (3), respectively., The likelihood function of $\mathbf{Y}$ is given by (see Arnold et al. (1998).

$$\begin{aligned}L(\mathbf{Y};\lambda,\beta) &= f(y_m;\lambda,\beta)\prod_{i=1}^{m-1}h(y_i;\lambda,\beta) \\ &= \beta^m\lambda^m e^{-\lambda(\frac{y_m}{1-y_m})^{\beta}}\prod_{i=1}^{m}\frac{y_i^{\beta-1}}{(1-y_i)^{\beta+1}}.\end{aligned} \tag{22}$$

where $0 < y_1 \leq y_2 \leq ... \leq y_m < 1$, $\lambda > 0$ and $\beta > 0$.
In light of the observed upper record, $\mathbf{Y} = \{y_1, y_2, ..., y_m\}$, and by combining Eqs. (21) and (22), we obtain the joint posterior density of λ and β as following:

$$\pi(\lambda, \beta|\mathbf{Y}) \propto \frac{1}{C}\lambda^{m+a_1-1}\beta^{m+a_2-1}e^{-\beta(b_2-\sum_{i=1}^{m}\log R_i)}e^{-\lambda(b_1+R_m^{\beta})}, \tag{23}$$

where C is the normalizing constant and it is obtained as

$$C = \int_0^\infty \int_0^\infty \lambda^{m+a_1-1}\beta^{m+a_2-1}e^{-\beta(b_2-\sum_{i=1}^{m}\log R_i)}e^{-\lambda(b_1+R_m^{\beta})}d\lambda d\beta. \tag{24}$$

The joint posterior distribution in Eq.(23) can be rewritten as follows:

$$\pi(\lambda, \beta|data) \propto \pi_1(\beta|data)\pi_2(\lambda|\beta, data), \tag{25}$$

where

$$\pi_1(\beta|data) \propto \frac{\beta^{m+a_2-1}e^{-\beta b_2}\prod_{i=1}^{m}R_i^{\beta}}{\left(b_1 + R_m^{\beta}\right)^{m+a_1}}, \tag{26}$$

and $\pi_2(\lambda|\beta, data)$ is a gamma density with shape and scale parameters are $m + a_1$ and $\left[b_1 + R_m^{\beta}\right]^{-1}$, respectively. Here we are mainly interested in obtaining the posterior predictive density of y_s, $f_s^P(y_s|\mathbf{Y})$, given observed data $\mathbf{Y}$. The posterior predictive density of y_s is given by

$$\begin{aligned} f_s^P(y_s|\mathbf{Y}) &= E_{posterior}(f(Y_s|\mathbf{Y}, \lambda, \beta)) \\ &= \int_0^\infty \int_0^\infty f(y_s|\mathbf{Y}, \lambda, \beta)\pi(\lambda, \beta|\mathbf{Y})d\lambda d\beta \end{aligned} \tag{27}$$

where $f(y_s|\mathbf{Y}, \lambda, \beta)$ and $\pi(\lambda, \beta|\mathbf{Y})$ are given in Eqs.(6) and (23), respectively. Substituting these equations in Eq.(27), then the posterior predictive density function $f_s^P(y_s|\mathbf{Y})$ becomes

$$\begin{aligned} f_s^P(y_s|\mathbf{Y}) = \frac{1}{C}\int_0^\infty \int_0^\infty & \frac{\lambda^{s-m}\beta}{\Gamma(s-m)}\frac{R_s^{\beta}}{y_s(1-y_s)}(R_s^{\beta} - R_m^{\beta})^{s-m-1}e^{-\lambda(R_s^{\beta}-R_m^{\beta})} \\ & \times \lambda^{m+a_1-1}e^{-\lambda(b_1+R_m^{\beta})}\beta^{m+a_2-1}e^{-\beta(b_2-\sum_{i=1}^{m}\log(R_i))}d\lambda d\beta, \end{aligned} \tag{28}$$

Since

$$\int_0^\infty \lambda^{s+a_1-1}e^{-\lambda(b_1+R_s^{\beta})}d\lambda = \frac{\Gamma(s+a_1)}{(b_1+R_s^{\beta})^{s+a_1}}$$

and

$$\begin{aligned} C &= \int_0^\infty \int_0^\infty \lambda^{m+a_1-1}\beta^{m+a_2-1}e^{-\lambda(b_1+R_m^{\beta})}e^{-\beta(b_2-\sum_{i=1}^{m}\log(R_i))}d\lambda d\beta \\ &= \frac{\Gamma(m+a_1)\Gamma(m+a_2)}{(b_2-\sum_{i=1}^{m}\log R_i)^{m+a_2}} \times E_{\pi_{\beta^\star}}[J(\beta)], \end{aligned}$$

then Eq.(28) reduces to the form

$$\begin{aligned} f_s^P(y_s|\mathbf{Y}) = & \frac{(b_2-\sum_{i=1}^{m}\log R_i)^{m+a_2}}{\Gamma(m+a_1)\Gamma(m+a_2)}\frac{\Gamma(s+a_2)\Gamma(m+a_2+1)}{\Gamma(s-m)(b_2-\sum_{i=1}^{m}\log R_i)^{m+a_2+1}} \\ & \times \frac{1}{y_s(1-y_s)}\frac{E_{\pi_1^*}[I(y_s, \beta)]}{E_{\pi_2^\star}[J(\beta)]}, \end{aligned} \tag{29}$$

where $I(y_s,\beta)=\frac{(R_s^{\beta}-R_m^{\beta})^{s-m-1}R_s^{\beta}}{(R_s^{\beta}+b_1)^{s+a_1}}$, and π_1^* is the gamma density function with the shape and scale parameters are $m+a_2+1$, $\frac{1}{b_2-\sum_{i=1}^{m}\log R_i}$, respectively. $J(\beta)=\frac{1}{(b_1+R_m^{\beta})^{m+a_1}}$, and $\pi_2{}^{\star}$ is the gamma density function with the shape and scale parameters are $m+a_2$, $\frac{1}{b_2-\sum_{i=1}^{m}\log R_i}$, respectively.
Notice that Eq(29)can not be computed explicitly. Therefore, as an approximate of the expected value we take the mean of β, then if we replace β by it's corresponding means say β_1 and β_2 in $I(y_s,\beta)$ and $J(\beta)$ from π_1^* and π_2^*, respectively, then an approximate of $f_s^p(y_s|\mathbf{Y})$, which is denoted by $\widehat{f_s^*(y_s|\mathbf{Y})}$, is obtained as follows

$$\begin{aligned}\widehat{f_s^*(y_s|\mathbf{Y})} &= \frac{1}{g(y_m)}\frac{\Gamma(s+a_2)(m+a_2)}{\Gamma(m+a_1)\Gamma(s-m)(b_2-\sum_{i=1}^{m}\log R_i)}\\ &\quad\times\frac{1}{y_s(1-y_s)}\frac{I(y_s,\beta_1)}{J(\beta_2)},\end{aligned}\tag{30}$$

where $g(y_m)=\int_{y_m}^{1}\widehat{f_s^*(y_s|\mathbf{Y})}dy_s$.
If $\hat{Y}$ is a predictor of y_s, $0<y_m<y_s<1$, then the Bayes predictive estimators of y_s under SE loss, $\hat{Y}_{SEP}$ and LINEX loss functions, $\hat{Y}_{LEP}$ are given by:

$$\begin{aligned}\hat{Y}_{SEP} &= E_{\widehat{f_s^*}}(Y_s|\mathbf{Y})\\ &= \int_{y_m}^{1} y_s\widehat{f_s^*(y_s|\mathbf{Y})}dy_s\\ &= \frac{1}{g(y_m)}\frac{\Gamma(s+a_2)(m+a_2)}{J(\beta_2)\Gamma(m+a_1)\Gamma(s-m)(b_2-\sum_{i=1}^{m}\log R_i)}\\ &\quad\times\int_{y_m}^{1}\frac{I(y_s,\beta_1)}{(1-y_s)}dy_s,\end{aligned}\tag{31}$$

and

$$\begin{aligned}\hat{Y}_{LEP} &= \frac{-1}{\nu}\log E_{\widehat{f_s^*}}\left(e^{-\nu Y_s}|\mathbf{Y}\right)\\ &= \frac{-1}{\nu}\log\int_{y_m}^{1} e^{-\nu y_s}\widehat{f_s^*(y_s|\mathbf{Y})}dy_s\\ &= \frac{-1}{\nu}\log[\frac{1}{g(y_m)}\frac{\Gamma(s+a_2)(m+a_2)}{J(\beta_2)\Gamma(m+a_1)\Gamma(s-m)(b_2-\sum_{i=1}^{m}\log R_i)}\\ &\quad\times\int_{y_m}^{1} e^{-\nu y_s}\frac{I(y_s,\beta_1)}{y_s(1-y_s)}dy_s],\end{aligned}\tag{32}$$

respectively. Furthermore, since we are often interested in predict the first unobserved record value, substitute $s=m+1$ in Eqs.(31) and (32) and using binomial expansion on $(R_s^{\beta_1}-R_m^{\beta_1})^{s-m-1}$,we will get the following:

$$\begin{aligned}\hat{Y}_{m+1}^{SEP2} &= \frac{1}{g(y_m)}\frac{(b_1+R_m^{\beta_2})^{m+a_1}\Gamma(s+a_2)(m+a_2)}{\Gamma(m+a_1)(b_2-\sum_{i=1}^{m}\log R_i)}\\ &\quad\times\int_{y_m}^{1}\frac{1}{1-y_s}\frac{R_s^{\beta_1}}{(b_1+R_s^{\beta_1})^{s+a_1}}dy_s,\end{aligned}$$

and

$$\begin{aligned}\hat{Y}_{m+1}^{LEP2} &= \frac{-1}{\nu}\log[\frac{1}{g(y_m)}\frac{(b_1+R_m^{\beta_2})^{m+a_1}\Gamma(s+a_2)(m+a_2)}{\Gamma(m+a_1)(b_2-\sum_{i=1}^{m}\log R_i)}\\ &\quad\times\int_{y_m}^{1} e^{-\nu y_s}\frac{1}{y_s(1-y_s)}\frac{R_s^{\beta_1}}{(b_1+R_s^{\beta_1})^{s+a_1}}dy_s].\end{aligned}$$

3. Prediction Intervals

The second way to use the previous data to predict a future observation from the same distribution is to construct an interval, which will likely to contain a future observation given what has already been observed such these intervals are called prediction intervals(PIs). In this section we consider several methods to obtain the PI_s based on the observed record sample $\mathbf{Y} = (y_1, y_2, ..., y_m)$.

3.1. Pivotal Method

Let us take the random variable Z as

$$Z = R_s^\beta - R_m^\beta,$$

it is easily to observe that $Z|y_m \sim Gamma(s-m, \frac{1}{\lambda})$, using jacobian transformation of $Z|y_m$ in Eq.(6). Hence when the parameters λ and β are known and y_m is given then the pivotal quantity $2\lambda Z|y_m \sim \chi^2(2(s-m))$. From this, the exact $(1-\tau)100\%$ PI of y_s is $(L1(y_m), U1(y_m))$ where

$$L_1(y_m) = \frac{(\frac{\chi^2_{\frac{\tau}{2}}(2(s-m))}{2\lambda} + R_m^\beta)^{\frac{1}{\beta}}}{1 + ((\frac{\chi^2_{\frac{\tau}{2}}(2(s-m))}{2\lambda} + R_m^\beta)^{\frac{1}{\beta}})}, \qquad U_1(y_m) = \frac{(\frac{\chi^2_{1-\frac{\tau}{2}}(2(s-m))}{2\lambda} + R_m^\beta)^{\frac{1}{\beta}}}{1 + ((\frac{\chi^2_{1-\frac{\tau}{2}}(2(s-m))}{2\lambda} + R_m^\beta)^{\frac{1}{\beta}})}, \tag{33}$$

when λ and β are unknown, the parameters in Eqs.(33), have to be estimated by their MLEs. So an approximate $(1-\tau)100\%$ PI is obtained as follows:

$$\hat{L}_1(y_m) = \frac{(1 + \frac{\chi^2_{\frac{\tau}{2}}(2(s-m))}{2m})^{\frac{1}{\beta}} R_m}{1 + (1 + \frac{\chi^2_{\frac{\tau}{2}}(2(s-m))}{2m})^{\frac{1}{\beta}} R_m}, \qquad \hat{U}_1(y_m) = \frac{(1 + \frac{\chi^2_{1-\frac{\tau}{2}}(2(s-m))}{2m})^{\frac{1}{\beta}} R_m}{1 + (1 + \frac{\chi^2_{1-\frac{\tau}{2}}(2(s-m))}{2m})^{\frac{1}{\beta}} R_m} \tag{34}$$

Since we are usually interest in predict the first prediction, when $s = m+1$, then using the pivotal quantity $\lambda Z|y_m \sim Exp(1)$, the $(1-\tau)100\%$ exact PI and approximate PI of y_{m+1}, respectively, are given by

$$L_2(y_m) = \frac{(R_m^\beta - \frac{1}{\lambda}\log(1-\frac{\tau}{2}))^{\frac{1}{\beta}}}{1 + [R_m^\beta - \frac{1}{\lambda}\log(1-\frac{\tau}{2})]^{\frac{1}{\beta}}}, \qquad U_2(y_m) = \frac{(R_m^\beta - \frac{1}{\lambda}\log(\frac{\tau}{2}))^{\frac{1}{\beta}}}{1 + [R_m^\beta - \frac{1}{\lambda}\log(\frac{\tau}{2})]^{\frac{1}{\beta}}} \tag{35}$$

and

$$\hat{L}_2(y_m) = \frac{R_m(1 - \frac{1}{m}\log(1-\frac{\tau}{2}))^{\frac{1}{\beta}}}{1 + R_m[1 - \frac{1}{m}\log(1-\frac{\tau}{2})]^{\frac{1}{\beta}}}, \qquad \hat{U}_2(y_m) = \frac{1 - \frac{1}{m}\log(\frac{\tau}{2}))^{\frac{1}{\beta}}}{1 + R_m[1 - \frac{1}{m}\log(\frac{\tau}{2})]^{\frac{1}{\beta}}} \tag{36}$$

3.2. Highest Conditional Density Method

An interval in which the value of the conditional PDF of y_s given observed data, $f(y_s|\mathbf{Y})$, at every point inside it is greater than that for every point outside it is called the highest conditional density (HCD) interval, Raqab (2001) [24]. If we replace λ and β in Eq.(6) with their corresponding MLEs we will obtain the approximate PDF of y_s given y_m as follows:

$$\begin{aligned}\widehat{f(y_s|y_m, \hat{\lambda}, \hat{\beta})} &= \hat{\beta}(\frac{m}{R_m^{\hat{\beta}}})^{s-m} R_s^{\hat{\beta}} \frac{1}{y_s(1-y_s)} \\ &\times \frac{[R_s^{\hat{\beta}} - R_m^{\hat{\beta}}]^{s-m-1}}{\Gamma(s-m)} e^{-\frac{m}{R_m^{\hat{\beta}}}[R_s^{\hat{\beta}} - R_m^{\hat{\beta}}]},\end{aligned} \tag{37}$$

since the approximation conditional density given in Eq.(37) is unimodal function of $U = \frac{m(R_s^{\hat{\beta}} - R_m^{\hat{\beta}})}{R_m^{\hat{\beta}}}$, and the distribution of U given y_m is $G(s-m, 1)$ with PDF

$$g(u) = \frac{u^{s-m-1}e^{-u}}{\Gamma(s-m)}, u > 0.$$

therefore, an interval $[w_1, w_2]$ is called a HCD-PI of content $1-\tau$ $(0<\tau<1)$, if $[w_1, w_2]$ is given by

$$\{w : w \in [0,\infty), g(w) \geq k\} \subset [0,\infty), \text{where} \int_{w_1}^{w_2} g(w)dw = 1-\tau,$$

for some $k > 0$. If $s > m+1$, then g(w) is a unimodal PDF and it attains its maximum value at $\xi = s-m-1 \in (0,\infty)$. In this context, the HCD method requires finding two cut of points $w_1 = w_{\frac{\tau}{2}} \leq \xi \leq w_2 = w_{1-\frac{\tau}{2}}$, as $(\frac{\tau}{2})100th$ and $(1-\frac{\tau}{2})100th$ percentiles of $G(s-m, 1)$ distribution, respectively, satisfying

$$1-\tau = \int_{w_1}^{w_2} g(u)du, \tag{38}$$

and

$$g(w_1) = g(w_2). \tag{39}$$

Eqs.(38) and (39)can be simplified as follows:

$$1-\tau = \Gamma(s-m, w_1) - \Gamma(s-m, w_2), \tag{40}$$

and

$$(\frac{w_1}{w_2})^{s-m-1} = e^{-(w_2-w_1)}, \tag{41}$$

where $\Gamma(a,b) = \frac{1}{\Gamma(a)}\int_b^{\infty} t^{a-1}e^{-t}dt$, which is the upper incomplete gamma function. In consequence of that, the $(1-\tau)100\%$ PI of y_s based on HCD method is computed to be $(L_3(y_m), U_3(y_m))$ where

$$L_3(y_m) = \frac{R_m(1+\frac{w_1}{m})^{\frac{1}{\hat{\beta}}}}{1+R_m(1+\frac{w_1}{m})^{\frac{1}{\hat{\beta}}}}, \qquad U_3(y_m) = \frac{R_m(1+\frac{w_2}{m})^{\frac{1}{\hat{\beta}}}}{1+R_m(1+\frac{w_2}{m})^{\frac{1}{\hat{\beta}}}} \tag{42}$$

Now, let us consider a special case where $s = m+1$, it may be noted that, Eq.(41) yields $w_1 = w_2$ and then no prediction interval can be constructed. In this case to avoid this problem, we can note the density $g(w)$ is a decreasing function with $g(0) = 1$ and $g(\infty) = 0$. Then, the HCD method involves finding $[0, w_2]$ where

$$\int_0^{w_2} g(w)dw = 1-\tau, \text{which is equivalent to } w_2 = -\log(\tau).$$

This leads that the $(1-\tau)100\%$ HCD PI of y_{m+1} where it's bounds are $L_4(y_m)$ and $U_4(y_m)$ as follows:

$$(y_m, \frac{\frac{y_m}{1-y_m}(1-\frac{\log(\tau)}{m})^{\frac{1}{\hat{\beta}}}}{1+\frac{y_m}{1-y_m}(1-\frac{\log(\tau)}{m})^{\frac{1}{\hat{\beta}}}}).$$

3.3. The Shortest Length Prediction Intervals

Another related PI is the shortest length(SL)PI. Using the fact that the distribution of $V = \frac{2m(R_s^{\hat{\beta}}-R_m^{\hat{\beta}})}{R_m^{\hat{\beta}}} \sim \chi^2_{2(s-m)}$, we first choose the constants c_1 and c_2 where $c_1 < c_2$ satisfying the following:

$$P(c_1 < \chi^2_{2(s-m)} < c_2) = 1-\tau \tag{43}$$

which is equivalent to

$$P(c_1 < \frac{2m(R_s^{\hat{\beta}}-R_m^{\hat{\beta}})}{R_m^{\hat{\beta}}} < c_2) = 1-\tau, \tag{44}$$

or equivalently

$$P(\frac{R_m(1+\frac{c_1}{2m})^{\frac{1}{\hat{\beta}}}}{1+R_m(1+\frac{c_1}{2m})^{\frac{1}{\hat{\beta}}}} < y < \frac{R_m(1+\frac{c_2}{2m})^{\frac{1}{\hat{\beta}}}}{1+R_m(1+\frac{c_2}{2m})^{\frac{1}{\hat{\beta}}}}) = 1-\tau. \tag{45}$$

This implies that a $(1-\tau)100\%$ PI for y_s is derived to be $(L_5(y_m), U_5(y_m))$, where

$$L_5(y_m) = \frac{R_m(1+\frac{c_1}{2m})^{\frac{1}{\beta}}}{1+R_m(1+\frac{c_1}{2m})^{\frac{1}{\beta}}}, \text{ and } U_5(y_m) = \frac{R_m(1+\frac{c_2}{2m})^{\frac{1}{\beta}}}{1+R_m(1+\frac{c_2}{2m})^{\frac{1}{\beta}}}. \tag{46}$$

The best choices for c_1 and c_2 are ones minimizing the width of PI, $U_5(y_m) - L_5(y_m)$. The shortest length $(1-\tau)100\%$ PI can be obtained by minimizing the Lagrangian multipliers function by imposing (43) as follows:

$$L(c_1, c_2, \omega) = U_5(y_m) - L_5(y_m) - \omega[\int_{c_1}^{c_2} g_{\chi^2_{2(s-m)}}(v)dv - (1-\tau)], \tag{47}$$

which is equivalent to

$$L(c_1, c_2, \omega) = \frac{R_m(1+\frac{c_2}{2m})^{\frac{1}{\beta}}}{1+R_m(1+\frac{c_2}{2m})^{\frac{1}{\beta}}} - \frac{R_m(1+\frac{c_1}{2m})^{\frac{1}{\beta}}}{1+R_m(1+\frac{c_1}{2m})^{\frac{1}{\beta}}} - \omega[\int_{c_1}^{c_2} g_{\chi^2_{2(s-m)}}(v)dv - (1-\tau)], \tag{48}$$

where $g_{\chi^2_{2(s-m)}}$ is the PDF of $\chi^2_{2(s-m)}$ distribution and ω is a Lagrangian multiplier. The constants c_1 and c_2 can be derived from equating the partial derivative of $L(c_1, c_2, \omega)$, with respect to c_1, c_2 and ω, to zero as follows:

$$\frac{\partial L}{\partial c_1} = \frac{\frac{-R_m}{2m\beta}(1+\frac{c_1}{2m})^{\frac{1}{\beta}-1}}{(1+R_m(1+\frac{c_1}{2m})^{\frac{1}{\beta}})^2} + \omega g_{\chi^2_{2(s-m)}}(c_1) = 0 \tag{49}$$

$$\frac{\partial L}{\partial c_2} = \frac{\frac{R_m}{2m\beta}(1+\frac{c_2}{2m})^{\frac{1}{\beta}-1}}{(1+R_m(1+\frac{c_2}{2m})^{\frac{1}{\beta}})^2} - \omega g_{\chi^2_{2(s-m)}}(c_2) = 0 \tag{50}$$

$$\frac{\partial L}{\partial \omega} = -[\int_{c_1}^{c_2} g_{\chi^2_{2(s-m)}}(v)dv - (1-\tau)] = 0 \tag{51}$$

After some algebraic computations on Eqs.(49)and (50), we reach to

$$(\frac{c_2}{c_1})^{s-m-1} e^{\frac{-1}{2}(c_2-c_1)} = (\frac{2m+c_2}{2m+c_1})^{\frac{1}{\beta}-1} \times \left[\frac{1+R_m(1+\frac{c_1}{2m})^{\frac{1}{\beta}}}{1+R_m(1+\frac{c_2}{2m})^{\frac{1}{\beta}}}\right]^2, \tag{52}$$

also from Eq.(51) we get

$$\int_{c_1}^{c_2} g_{\chi^2_{2(s-m)}}(v)dv = (1-\tau). \tag{53}$$

Now, c_1 and c_2 of the shortest PI can be computed simultaneously by solving Eqs.(52)and (53) numerically.
We now consider the case where $s = m+1$. In this case, $g_{\chi^2_{2(s-m)}}(v)$ is decreasing function with $g_{\chi^2_{2(s-m)}}(0) = \frac{1}{2}$ and $g_{\chi^2_{2(s-m)}}(\infty) = 0$. Consequently, the lower endpoint of the PI can be chosen simply as $L_5(y_m) = y_m$, this leads that the $(1-\tau)100\%$ PI is $(y_m, U_5(y_m))$ as a modified of SLPI for y_{m+1}.

3.4. Bayesian Prediction Interval

Our aim of this section is to obtain the Bayes predictive intervals for the s^{th} future record, $s > m$, based on observed records data from $K(\lambda, \beta)$ distribution under SE and LINEX loss functions. Let us define the survival prediction function of y_s, $s > m$, based on the observed record sample, $\mathbf{Y} = (y_1, y_2, ..., y_m)$, as

$$\begin{aligned} S^P(y_s|\mathbf{Y}, \lambda, \beta) &= E_{posterior}[S(y_s|\mathbf{Y}, \lambda, \beta)] \\ &= \int_0^\infty \int_0^\infty S(y_s|\mathbf{Y}, \lambda, \beta) \times \pi(\lambda, \beta|\mathbf{Y}) d\lambda d\beta. \end{aligned} \tag{54}$$

Where $S(y_s|\mathbf{Y},\lambda,\beta)$ is the survival function of y_s, and $S(y_s|\mathbf{Y},\lambda,\beta)$= $S(y_s|y_m,\lambda,\beta)$ due to Markovian property of record statistics, thus

$$\begin{aligned} S(y_s|y_m,\lambda,\beta) &= Pr(Y > y_s|y_m) \\ &= \int_{y_s}^{1} f(t|y_m,\lambda,\beta)dt \\ &= \int_{y_s}^{1} \frac{[H(t)-H(y_m)]^{s-m-1}}{\Gamma(s-m)} \frac{f(t|\lambda,\beta)}{1-F(y_m|\lambda,\beta)}dt \end{aligned}$$

By making the transformation $v = H(t) - H(y_m)$ and using the relation between the incomplete gamma function and the poisson distribution, the survival function $S(y_s|y_m,\lambda,\beta)$ reduces to the form

$$S(y_s|y_m,\lambda,\beta) = \sum_{i=0}^{s-m-1} \frac{\frac{1-F(y)}{1-F(y_m)}[\log(\frac{1-F(y_m)}{1-F(y)})]^i}{i!} \tag{55}$$

using the Eqs.(55) and (1), the survival function $S(y_s|y_m,\lambda,\beta)$ is obtained as follows:

$$S(y_s|y_m,\lambda,\beta) = \sum_{i=0}^{s-m-1} e^{-\lambda(R_s^{\beta}-R_m^{\beta})}\lambda^i \frac{(R_s^{\beta}-R_m^{\beta})^i}{i!} \tag{56}$$

Now, by substitute the Eq.(56)in the Eq.(54), then the survival prediction function of y_s, is given by:

$$S^P(y_s|y_m,\lambda,\beta) = \int_0^{\infty}\int_0^{\infty}[\sum_{i=0}^{s-m-1} e^{-\lambda(R_s^{\beta}-R_m^{\beta})}\lambda^i \frac{(R_s^{\beta}-R_m^{\beta})^i}{i!}] \times \pi(\lambda,\beta|data)d\lambda d\beta \tag{57}$$

It is obvious that Eq.(57) can not be expressed in a closed form and hence it can not be evaluated analytically. For this, we propose to approximate Eq.(57) by using an importance sampling technique as suggested by Chen (1999)[10]. We need the following lemma for further development.

Lemma 3.1
The conditional distribution of β given the observed records, $\pi_1(\beta|data)$, is log concave.

Proof
The log likelihood of conditional distribution of β given the observed records, Eq. (26), is given by:

$$\log \pi_1(\beta|data) \propto -(m+a_1)\log(b_1+R_m^{\beta}) + (m+a_2-1)\log(\beta) - \beta(b_2 - \sum_{i=1}^{m}\log R_i) \tag{58}$$

By differentiating $\log \pi_1(\beta|data)$ twice with respect to β, we get:

$$\frac{\partial^2 \log \pi_1(\beta|data)}{\partial \beta^2} = -(m+a_1)\left(\frac{(b1+R_m^{\beta})R_m^{\beta}(\log R_m)^2 - (R_m \log R_m)^2}{(b_1+R_m^{\beta})^2}\right) - \frac{m+a_2-1}{\beta^2} \tag{59}$$

Since $\frac{\partial^2 \log \pi_1(\beta|data)}{\partial \beta^2} < 0$, this follows that $\pi_1(\beta|data)$ is log-concave density. □

Since $\pi_1(\beta|data)$ has a log-concave density, using the idea of Devroye (1984) it is possible to generate a sample from $\pi_1(\beta|data)$. Moreover, since $\pi_2(\lambda|\beta,data)$ follows gamma, it is quite simple to generate from $\pi_2(\lambda|\beta,data)$. Now we would like to provide the importance sampling procedure to compute the survival prediction estimators and also to construct the Bayesian PIs of y_s. Using Lemma (3.1), a simulation consistent estimator for the predictive survival function of y_s can be obtained using the following algorithm:

algorithm 3.2
step 1: Generate β from the log concave density function $\pi_1(\beta|data)$, Eq. (26), by using the method proposed by Devroy (1984) as follows:

(a) Compute $c = \pi_1(m|data)$ where m is the mode of $\pi_1(.|data)$, also compute $d = \log c$.

(b) Generate U uniform on $[0, 2]$ and E exponential random variate where E independent of U.
(c) If $U \leq 1$, then $\beta = U$ and $T = -E$, else $\beta = 1 + E^*$ and $T = -E - E^*$, where E^* is a new exponential random variate.
(d) Set $\beta = m + \frac{\beta}{c}$ and if $T \leq \log \pi_1(\beta|data) - d$ then β is a sample from $\pi_1(.|data)$ else go to step (b).

step 2: For each β obtained in step (1), generate λ from the marginal posterior density function of λ given β, in step (1), and data, $\pi_2(\lambda|\beta, data)$.
step 3: Repeat step (1) and step (2) M times, and obtain MCMC-samples, $(\lambda_j, \beta_j), j = 1, 2, ..., M$.

Then the survival prediction estimators of y_s under SE loss function $\hat{S}_{SES}$, and under LINEX loss function $\hat{S}_{LES}$, are obtained, respectively, as

$$\hat{S}_{SES}(y_s) = \frac{1}{M} \sum_{j=1}^{M} [\sum_{i=0}^{s-m-1} e^{-\lambda_j (R_s^{\beta_j} - R_m^{\beta_j})} \lambda_j^i \frac{(R_s^{\beta_j} - R_m^{\beta_j})^i}{i!}] \tag{60}$$

$$\hat{S}_{LES}(y_s) = \frac{-1}{\nu} \log[\frac{1}{M} \sum_{j=1}^{M} e^{-\nu \sum_{i=0}^{s-m-1} e^{-\lambda_j (R_s^{\beta_j} - R_m^{\beta_j})} \lambda_j^i \frac{(R_s^{\beta_j} - R_m^{\beta_j})^i}{i!}}] \tag{61}$$

Therefore, the $(1-\tau)\%$ Bayesian predictive interval for $y_s, s > m$ under SE loss is given by $(L_5(y_m), U_5(y_m))$ where $L_5(y_m)$ and $U_5(y_m))$ can be obtained by solving the following non-linear equations simultaneously

$$\begin{aligned} Pr(Y > L(y_m)|y_m) = 1 - \frac{\tau}{2} \Leftrightarrow \hat{S}_{SES}(L(y_m)) = 1 - \frac{\tau}{2} \\ Pr(Y > U(y_m)|y_m) = \frac{\tau}{2} \Leftrightarrow \hat{S}_{SES}(U(y_m)) = \frac{\tau}{2} \end{aligned} \tag{62}$$

and, the $(1-\tau)\%$ Bayesian predictive interval for $y_s, s > m$ under LINEX loss function is given by $(L_6(y_m), U_6(y_m))$ where $L_6(y_m)$ and $U_6(y_m))$ can be obtained by solving the following non-linear equations simultaneously

$$\begin{aligned} Pr(Y > L|data) = 1 - \frac{\tau}{2} \Leftrightarrow \hat{S}_{LES} L(y_m) = 1 - \frac{\tau}{2} \\ Pr(Y > U|data) = \frac{\tau}{2} \Leftrightarrow \hat{S}_{LES} U(y_m) = \frac{\tau}{2} \end{aligned} \tag{63}$$

Thus we need to apply an appropriate numerical technique to solve these non-linear equations (62)and (63). As a special case when $s = m + 1$, $(1 - \gamma)\%$ PI for y_{m+1}, can be obtained by setting $s = m + 1$ in the Eqs.(62)and (63).

4. Data Analysis

In this section, we study the proposed prediction classical and Bayesian methods for records from real and simulated data sets from two-parameter Kies distribution.

4.1. Real Data: Total Annual Rainfall

In this example, we analyze the total annual rainfall (in inches) during 25 years from 1984-2008 recorded at Los Angeles Civic Center. This data is given below, see
http : // www.laalmanac.com/weather/we08aa.php:

12.82	17.86	7.66	2.48	8.08	7.35	11.99	21.00	7.36
8.11	24.35	12.44	12.40	31.01	9.09	11.57	17.94	4.42
16.42	9.25	37.96	13.19	3.21	13.53	9.08		

Firstly, all observations have been divided over 100, where we can also divide by any number greater than 38, in order to transform them to be in $(0, 1)$, the support of K(λ, β) distribution. Then, the well known Kolmogorov-Smirnov (K-S) goodness of fit test is used to test whether the Kies distribution adequately fits this data set or not. The MLEs of λ and β have been computed based on the complete sample using Newton Raphson method and found to be 11.1410 and 1.4171, respectively. The corresponding K-S test statistic and the associated P-value are equal to 0.1674 and 0.4851, respectively. Accordingly, one cannot reject the hypothesis that the data set is coming from K(λ, β) distribution.
It can be easily seen that the upper records obtained from this data set are: 0.1282, 0.1786, 0.2100, 0.2435, 0.3101, 0.3796. Based on these records, we compute the value of the predictors of the 7^{th}, 8^{th} and 9^{th} future records using point and interval prediction methods including, MLP, MMLP, BUP, CMP and Bayes predictor as well as the pivotal quantity, HCD, SL and Bayesian PIs. To study how sensitive are the Bayes estimates and the Bayes predictors for the choice of the hyper-parameters, we consider two priors as follows: $Prior\ 1: a_1 = 24, b_1 = 2, a_2 = 7, b_2 = 5$, and $Prior\ 2: a_1 = 12, b_1 = 1, a_2 = 12, b_2 = 9$.
Tables (1) and (2) summarize the results of point and interval predictors of the 7^{th}, 8^{th} and 9^{th} future records, respectively, based on both the classical and the Bayesian approaches.

Table 1. predicted values for the 7^{th}, 8^{th} and 9^{th} future records based on the real data set(I)

m	Y_s	MLP	MMLP	BUP	CMP	Bayes predictor							
						$Prior\ 1$				$prior\ 2$			
						BE_{SE}	BE_{LE}			BE_{SE}	BE_{LE}		
							$\nu = -0.01$	$\nu = 0.5$	$\nu = 2$		$\nu = -0.01$	$\nu = 0.5$	$\nu = 2$
m= 6	Y_7	-	0.3796	0.4043	0.3971	0.4046	0.4046	0.4047	0.4048	0.4446	0.4446	0.4443	0.4433
	Y_8	0.4014	0.4090	0.4263	0.4215	0.4306	0.4306	0.4315	0.4344	0.4966	0.4966	0.4963	0.4952
	Y_9	0.4207	0.4338	0.4463	0.4426	0.4497	0.4497	0.4505	0.4530	0.5346	0.5346	0.5342	0.5328

Table 2. 95% PIs for the 7^{th}, 8^{th} and 9^{th} future records based on the real data set(I)

m	Y_s	Pivot	HCD	SPL	BPIs							
					$Prior\ 1$				$prior\ 2$			
					BE_{SE}	BE_{LE}			BE_{SE}	BE_{LE}		
						$\nu = -0.01$	$\nu = 0.5$	$\nu = 2$		$\nu = -0.01$	$\nu = 0.5$	$\nu = 2$
m= 6	y_7	(0.3803, 0.4620)	(0.3796, 0.4490)	(0.3796, 0.3917)	(0.1701, 0.3997)	(0.1701, 0.3998)	(0.1677, 0.3994)	(0.1596, 0.3987)	(0.2114, 0.5122)	(0.2114, 0.5122)	(0.2093, 0.5112)	(0.2028, 0.5084)
	y_8	(0.3862, 0.4934)	(0.3808, 0.4806)	(0.3816, 0.4812)	(0.3796, 0.4205)	(0.3647, 0.4205)	(0.3637, 0.4202)	(0.3654, 0.4192)	(0.3796, 0.5655)	(0.3791, 0.5656)	(0.3786, 0.5639)	(0.3794, 0.5595)
	y_9	(0.3961, 0.5170)	(0.4447, 0.51394)	(0.3908, 0.5078)	(0.2593, 0.4370)	(0.2593, 0.4370)	(0.2578,0.4367)	(0.2522, 0.4357)	(0.175, 0.6025)	(0.1752, 0.6026)	(0.1659, 0.6004)	(0.1334, 0.5947)

4.2. Real Data II: Size of Rocks

In this example, we study the proposed prediction methods, based on the following data of Dunsmore (1983)[12] which is discussed by Awad and Raqab (2000)[7]. This data shows the sizes of rocks to be crushed at any operation. If the size of the rock being crushed is greater than any that has been crushed before then a crushing machine has to be rest. These data are presented as follows:

9.3	0.6	24.4	18.1	6.6	9.0
14.3	6.6	13.0	2.4	5.6	33.8

Firstly, all observations have been divided over 100 in order to transform them to be in $(0, 1)$, the support of K(λ, β) distribution. Then, to check the validity of the use of K(λ, β) distribution to fit this data set, The K-S test is applied. The K-S distance and its respective p-value are computed to be K-S= 0.1184 and p-value= 0.9078,

respectively. Hence, it is quite reasonable to indicate that K(λ, β) distribution is adequately fitting this data.
The MLEs of λ and β have been computed based on the complete sample numerically using Newton Raphson method to be 7.8460 and 1.1105, respectively. The record values extracted from the original data set are: 0.093, 0.244, 0.338.
Based on the proposed prediction methods presented in the previous sections, the point predictors and PIs of the 4^{th}, 5^{th} and 6^{th} future records are computed and presented in Tables (3) and (4), respectively. In the context of Bayesian predictors, to examine the sensitivity of the hyperparameters (a_1, b_1, a_2, b_2) we used two different choices of the hyperparameters: $Prior\ 1: a_1 = 16, b_1 = 8, a_2 = 2, b_2 = 7$ and $Prior\ 2: a_1 = 4, b_1 = 0.5, a_2 = 5.5, b_2 = 5$, under SE and LINEX loss functions.

Table 3. predicted values for the 4^{th}, 5^{th} and 6^{th} future records based on the real data set(II)

m	Y_s	MLP	MMLP	BUP	CMP	Bayes predictor							
						$Prior\ 1$				$prior\ 2$			
						BE_{SE}	BE_{LE}			BE_{SE}	BE_{LE}		
							$\nu = -0.01$	$\nu = 0.5$	$\nu = 2$		$\nu = -0.01$	$\nu = 0.5$	$\nu = 2$
	Y_4	-	0.3380	0.3800	0.3707	0.3848	0.3848	0.3848	0.3850	0.3911	0.3911	0.3918	0.3935
3	Y_5	0.3704	0.3936	0.4152	0.4094	0.4233	0.4233	0.4236	0.4246	0.4329	0.4329	0.4341	0.4374
	Y_6	0.3968	0.4336	0.4453	0.4418	0.4585	0.4584	0.4593	0.4616	0.4698	0.4698	0.4715	0.4757

Table 4. 95% PIs for the 4^{th}, 5^{th} and 6^{th} future records based on the real data set(II)

m	Y_s	Pivot	HCD	SPL	BPIs							
					$Prior\ 1$				$prior\ 2$			
					BE_{SE}	BE_{LE}			BE_{SE}	BE_{LE}		
						$\nu = -0.01$	$\nu = 0.5$	$\nu = 2$		$\nu = -0.01$	$\nu = 0.5$	$\nu = 2$
	y_4	(0.3393, 0.4699)	(0.3380, 0.4512)	(0.3380, 0.3598)	(0.1059, 0.4098)	(0.1060, 0.4098)	(0.1034, 0.4095)	(0.0951, 0.4085)	(0.1156, 0.4355)	(0.1156, 0.4355)	(0.1128, 0.4344)	(0.1034, 0.4314)
3	y_5	(0.3501, 0.5125)	(0.3402, 0.4955)	(0.3428, 0.4974)	(0.3380, 0.4575)	(0.3301, 0.4575)	(0.3378, 0.4571)	(0.3379, 0.4559)	(0.3380, 0.4903)	(0.3332, 0.4903)	(0.3231, 0.4885)	(0.3317, 0.4839)
	y_6	(0.3674, 0.5427)	(0.3480, 0.5209)	(0.3612, 0.5338)	(0.1479, 0.4890)	(0.1480, 0.4891)	(0.1446, 0.4886)	(0.1339, 0.4873)	(0.1558, 0.5338)	(0.1559, 0.5338)	(0.1484, 0.5307)	(0.3480, 0.5230)

4.3. Simulated Data

Here we illustrate the usefulness of the proposed prediction methods for a simulated random sample of size 20 generated from Kies distribution with $\lambda = 2$ and $\beta = 1$ as follows: The upper record values extracted from the

0.1811	0.1293	0.5248	0.3133	0.4067	0.5709	0.2763	0.2429	0.0219	0.0574
0.5127	0.5185	0.1941	0.4240	0.0435	0.1036	0.0065	0.6788	0.3354	0.0133

above data set are: 0.1811, 0.5248, 0.5709, 0.6788.
Based on the above record values and based on the proposed prediction methods, we computed the point predictors and PIs of the 5^{th}, 6^{th} and 7^{th} future records and these results are presented in Tables (5) and (6). In context of Bayesian procedure ,two priors are considered assuming

$$Prior\ 1: a_1 = 20, b_1 = 10, a_2 = 9, b_2 = 8,$$
$$Prior\ 2: a_1 = 10, b_1 = 5, a_2 = 14, b_2 = 11,$$

under SE and LINEX loss function and using Different choices of LINEX parameter ν; namely -0.01, 0.5 and 2.
From Tables (5) and (6) we observed that the point predictors of y_5, y_6 and y_7 are lying within the so obtained prediction intervals, also, we can see the PIs become wider as s increases. It is evident from tables, the length of the classical PIs are shorter than the Bayesian PIs, and the best of them is SL PIs in terms of the length interval.

5. Simulation Results of the Prediction Methods

In this section we present a simulation study to assess the performance of the proposed prediction methods which were discussed in chapter 4. Performances are measured in terms of mean square prediction errors (MSPEs) and

Table 5. predicted values for the 5^{th}, 6^{th} and 7^{th} future records based on the simulated data set

m	Y_s	MLP	MMLP	BUP	CMP	Bayes predictor							
						$Prior$ 1				$prior$ 2			
						BE_{SE}	BE_{LE}			BE_{SE}	BE_{LE}		
							$\nu=-0.01$	$\nu=0.5$	$\nu=2$		$\nu=-0.01$	$\nu=0.5$	$\nu=2$
4	Y_5	-	0.6788	0.7188	0.7101	0.7145	0.7145	0.7151	0.7167	0.7128	0.7127	0.7133	0.7148
	Y_6	0.7133	0.7297	0.7506	0.7461	0.7437	0.7437	0.7450	0.7487	0.7366	0.7366	0.7374	0.7394
	Y_7	0.7408	0.7623	0.7763	0.7753	0.7649	0.7648	0.7664	0.7712	0.7605	0.7605	0.7617	0.7651

Table 6. 95% PIs for the 5^{th}, 6^{th} and 7^{th} future records based on the simulated data

m	Y_s	Pivot	HCD	SPL	BPIs							
					$Prior$ 1				$prior$ 2			
					BE_{SE}	BE_{LE}			BE_{SE}	BE_{LE}		
						$\nu=-0.01$	$\nu=0.5$	$\nu=2$		$\nu=-0.01$	$\nu=0.5$	$\nu=2$
4	y_5	(0.6801, 0.7983)	(0.6788, 0.7719)	(0.6788, 0.6980)	(0.6381,0.7275)	(0.6404, 0.7276)	(0.6111, 0.7269)	(0.5813, 0.7249)	(0.6539, 0.7469)	(0.6567, 0.7470)	(0.6236, 0.7456)	(0.5872, 0.7420)
	y_6	(0.6908, 0.8305)	(0.6807, 0.8036)	(0.6842, 0.8059)	(0.6788, 0.7598)	(0.6788, 0.7598)	(0.6788, 0.7585)	(0.6788, 0.7550)	(0.6788, 0.7741)	(0.6778, 0.7742)	(0.6788, 0.7727)	(0.6788, 0.7689)
	y_7	(0.7078, 0.8513)	(0.7251, 0.7738)	(0.6501, 0.8132)	(0.6463, 0.7855)	(0.6485, 0.7855)	(0.6188, 0.7837)	(0.5895, 0.7790)	(0.6488, 0.7947)	(0.6530, 0.7947)	(0.6218, 0.7930)	(0.5947, 0.7886)

the average biases of the predictors. We also compare the PIs, which are presented in the previous sections, in terms of coverage probabilities (CPs) and the average lengths (ALs). For conducting the Bayesian analysis, under the SE and LINEX loss functions, we assume four different priors as follows: $Prior$ 0: $a_1 = 0$, $b_1 = 0$, $a_2 = 0$, $b_2 = 0$.
For $\lambda = 1$, $\beta = 2$: $Prior$ 1: $a_1 = 20$, $b_1 = 20$, $a_2 = 16$, $b_2 = 8$.
For $\lambda = 2$, $\beta = 1$: $Prior$ 2: $a_1 = 1$, $b_1 = 0.5$, $a_2 = 20$, $b_2 = 20$.
For $\lambda = 2$, $\beta = 2$: $Prior$ 3: $a_1 = 1$, $b_1 = 0.5$, $a_2 = 10$, $b_2 = 5$.
These priors are proposed so as λ has the same mean but different variances, similarly for β. The main purpose of this is to reflect the sensitivity of our inferences to the choice of the hyper-parameters. The shape parameter of LINEX loss function ν is assumed to equal -0.01, 0.5 and 2, separately. In each case, we compute the value of the point predictor (classical and Bayesian). We also compute 95% PIs based on the pivotal quantity, HCD, SL and Bayesian methods.
Record samples from the Kies distribution were randomly generated using

$$Y_m \overset{D}{=} \frac{(\frac{1}{\lambda}\sum_{i=1}^{m} X_i^*)^{\frac{1}{\beta}}}{1 + (\frac{1}{\lambda}\sum_{i=1}^{m} X_i^*)^{\frac{1}{\beta}}}, \tag{64}$$

where $\{X_i^*\}_{i=1}^m$ is a sequence of $i.i.d.$ $Exp(1)$ random variables. The simulation process is repeated 1000 times. Using these random samples, MSPEs and prediction biases of the predictors are reported. Moreover, the CPs and ALs of the PIs are computed. The obtained results involving MSPEs, prediction biases, CPs, and ALs are presented in Tables(7) to (12).

From Tables (7), (8) and (9), and by considering the prediction average biases as an optimality criterion, there is a clear evidence that the BUPs are the most preferred classical point predictors. When comparing among the classical methods, one can see that the prediction average biases of the CMP are less than those of the MLP and MMLP for all the considered cases. Further, the prediction average biases of the MMLP are lower than those of MLP across all the considered cases.
When comparing Bayesian and frequentist methods, we observe that Bayes predictors perform better under different error loss functions and priors in terms of bias compared with MLPs and MMLPs. By considering MSPEs as an optimality criterion, it is also observed that Bayes predictors outperform MLPs and MMLPs. Additionally, Bayes predictors under the informative priors ($Prior$ 1, $Prior$ 2 and $Prior$ 3) are more efficient than the corresponding Bayes predictors under $Prior$ 0. Finally, when comparing among classical methods, the BUPs provide lower MSPEs than the other predictors.
From Tables (10), (11) and (12), and by considering the average length (AL) as an optimality criterion, the SL method is shown to be more efficient than the other methods for obtaining PIs. Among the HCD and the pivotal quantity methods, the HCD PIs are superior to the pivotal PIs in most considered cases in terms of ALs. When

Table 7. MSPEs and Average Bias from simulations of $\lambda = 1$ and $\beta = 2$

m	Y_s	Criterion	MLP	MMLP	BUP	CMP	$Prior\ 0$				$prior\ 1$			
							BE_{SE}	BE_{LE}			BE_{SE}	BE_{LE}		
								$\nu=-0.01$	$\nu=0.5$	$\nu=2$		$\nu=-0.01$	$\nu=0.5$	$\nu=2$
m=5	Y_7	MSPE	0.001492	0.001276	0.000459	0.000465	0.001538	0.001541	0.001432	0.00127	0.000724	0.000725	0.000699	0.000649
		Bias	-0.025749	-0.020434	-0.000026	-0.002417	-0.014546	-0.014617	-0.011204	-0.002874	-0.006494	-0.006517	-0.005386	-0.002143
	Y_8	MSPE	0.002023	0.001688	0.000458	0.000459	0.002091	0.002096	0.001909	0.001649	0.000846	0.000847	0.000811	0.000743
		Bias	-0.029308	-0.021296	-0.000250	-0.001470	-0.018323	-0.018420	-0.013771	-0.002445	-0.007216	-0.007246	-0.005708	-0.001262
m=6	Y_8	MSPE	0.000849	0.000745	0.000256	0.000258	0.000459	0.000460	0.000436	0.000426	0.000333	0.000334	0.000328	0.000320
		Bias	-0.019653	-0.015785	0.000655	-0.001638	-0.004980	-0.005027	-0.002732	0.003049	-0.001814	-0.001830	-0.001048	0.001246
	Y_9	MSPE	0.001666	0.001460	0.000420	0.000426	0.000616	0.000617	0.000566	0.000529	0.000383	0.000384	0.000372	0.000353
		Bias	-0.025196	-0.019210	-0.002515	-0.003844	-0.007818	-0.007883	-0.004729	0.003205	-0.003305	-0.003325	-0.002780	0.000744
m=7	Y_9	MSPE	0.000674	0.000589	0.000298	0.000313	0.000388	0.000389	0.000378	0.000375	0.000286	0.000286	0.000283	0.000277
		Bias	-0.017886	-0.015077	-0.002828	-0.004944	-0.007485	-0.007520	-0.005830	-0.003474	-0.005342	-0.005355	-0.004736	-0.002978
	Y_{10}	MSPE	0.000716	0.000604	0.000292	0.000296	0.000539	0.000539	0.000522	0.000528	0.000367	0.000368	0.000365	0.000363
		Bias	-0.017047	-0.012547	-0.000504	-0.001886	-0.005522	-0.005561	-0.003252	0.002697	-0.001536	-0.001553	-0.000724	0.001703
m=8	Y_{10}	MSPE	0.000529	0.000478	0.000209	0.000215	0.000332	0.000332	0.000322	0.000311	0.000264	0.000264	0.000261	0.000252
		Bias	-0.014086	-0.011842	-0.000702	-0.002672	-0.003652	-0.003677	-0.002404	0.000932	-0.002214	-0.002224	-0.001756	-0.000378
	Y_{11}	MSPE	0.000866	0.000762	0.000292	0.000299	0.000551	0.000552	0.000529	0.000501	0.000419	0.000419	0.000412	0.000395
		Bias	-0.017215	-0.013541	-0.002641	-0.003970	-0.005774	-0.005810	-0.004032	0.000622	-0.003266	-0.003271	-0.002619	0.002931

Table 8. MSPEs and Average Bias from simulations of $\lambda = 2$ and $\beta = 1$

m	Y_s	Criterion	MLP	MMLP	BUP	CMP	$Prior\ 0$				$prior\ 2$			
							BE_{SE}	BE_{LE}			BE_{SE}	BE_{LE}		
								$\nu=-0.01$	$\nu=0.5$	$\nu=2$		$\nu=-0.01$	$\nu=0.5$	$\nu=2$
m=5	Y_7	MSPE	0.006668	0.005880	0.001866	0.001870	0.004541	0.004554	0.004346	0.003902	0.002720	0.002720	0.002717	0.002713
		Bias	-0.048639	-0.038692	-0.000893	-0.004069	-0.026404	-0.026460	-0.023726	-0.016903	-0.013969	-0.013973	-0.013787	-0.013326
	Y_8	MSPE	0.007759	0.006576	0.001544	0.001536	0.005861	0.005868	0.005525	0.004803	0.002859	0.002859	0.002854	0.002844
		Bias	-0.052920	-0.039338	-0.001004	-0.001672	-0.030690	-0.030766	-0.027083	-0.017871	-0.013394	-0.013401	-0.013067	-0.012201
m=6	Y_8	MSPE	0.004181	0.003571	0.001033	0.001082	0.001787	0.001788	0.001747	0.001676	0.0012774	0.001277	0.001278	0.001271
		Bias	-0.043753	-0.036647	-0.005959	-0.009032	-0.017547	-0.017584	-0.015776	-0.011111	-0.008661	-0.008673	-0.008453	-0.007864
	Y_9	MSPE	0.004781	0.003882	0.000988	0.000991	0.002315	0.002317	0.002224	0.002049	0.001393	0.001394	0.001391	0.001385
		Bias	-0.044832	-0.034681	-0.004063	-0.005064	-0.019082	-0.019132	-0.016684	-0.010374	-0.006791	-0.006807	-0.006462	-0.005522
m=7	Y_9	MSPE	0.002482	0.002163	0.000934	0.000964	0.000728	0.000729	0.000705	0.000658	0.000553	0.000553	0.000552	0.000541
		Bias	-0.032610	-0.027503	-0.003014	-0.005911	-0.009754	-0.009778	-0.008530	-0.005255	-0.005776	-0.005778	-0.005645	-0.005315
	Y_{10}	MSPE	0.002956	0.002462	0.001021	0.001034	0.000945	0.000946	0.000903	0.000818	0.000582	0.000582	0.000581	0.000580
		Bias	-0.034864	-0.027463	-0.003136	-0.004354	-0.013391	-0.013434	-0.011725	-0.007236	-0.007706	-0.007710	-0.007472	-0.006851
m=8	Y_{10}	MSPE	0.001291	0.001143	0.000531	0.000553	0.000965	0.000965	0.000939	0.000871	0.000771	0.000771	0.000768	0.000760
		Bias	-0.021682	-0.018040	-0.002648	-0.005155	-0.010264	-0.010283	-0.009383	-0.006975	-0.007069	-0.007072	-0.006903	-0.006431
	Y_{11}	MSPE	0.001623	0.001382	0.000515	0.000528	0.001115	0.001116	0.001075	0.000982	0.000826	0.000826	0.000822	0.000811
		Bias	-0.024281	-0.018911	-0.004914	-0.006263	-0.011421	-0.011455	-0.010221	-0.006917	-0.007066	-0.007070	-0.006832	-0.006185

Table 9. MSPEs and Average Bias from simulations of $\lambda = 2$ and $\beta = 2$

m	Y_s	Criterion	MLP	MMLP	BUP	CMP	$Prior\ 0$				$prior\ 3$			
							BE_{SE}	BE_{LE}			BE_{SE}	BE_{LE}		
								$\nu=-0.01$	$\nu=0.5$	$\nu=2$		$\nu=-0.01$	$\nu=0.5$	$\nu=2$
m=5	Y_7	MSPE	0.001955	0.001670	0.000757	0.000779	0.000909	0.000910	0.000855	0.000846	0.000641	0.000642	0.000635	0.000634
		Bias	-0.031369	-0.025596	-0.001186	-0.004084	-0.007873	-0.007952	-0.004162	0.005065	0.001336	0.001311	0.002576	0.005931
	Y_8	MSPE	0.003108	0.002570	0.000814	0.000834	0.001273	0.001276	0.001182	0.001188	0.000819	0.000819	0.000815	0.000837
		Bias	-0.039313	-0.030320	-0.004476	-0.006068	-0.019799	-0.019909	-0.014669	-0.009371	-0.008746	-0.008778	-0.007192	-0.005370
m=6	Y_8	MSPE	0.001339	0.001179	0.000547	0.000565	0.000720	0.000721	0.000696	0.000695	0.000549	0.000549	0.000545	0.000544
		Bias	-0.023939	-0.019788	-0.001523	-0.004253	-0.007077	-0.007131	-0.004493	0.002078	-0.002288	-0.002305	-0.003835	-0.001974
	Y_9	MSPE	0.001727	0.001483	0.000524	0.000529	0.001027	0.001028	0.000988	0.001006	0.000718	0.000719	0.000718	0.000732
		Bias	-0.026193	-0.019552	-0.000711	-0.002417	-0.010211	-0.010287	-0.006628	0.002478	-0.002664	-0.002688	-0.001468	0.001625
m=7	Y_9	MSPE	0.000794	0.000690	0.000271	0.000295	0.000666	0.000667	0.000639	0.000601	0.000516	0.000516	0.000501	0.000497
		Bias	-0.019820	-0.016521	-0.001796	-0.004363	-0.006901	-0.006940	-0.005001	0.004139	-0.004167	-0.004182	-0.003465	-0.002487
	Y_{10}	MSPE	0.001097	0.000926	0.000315	0.000324	0.000952	0.000953	0.000901	0.000868	0.000672	0.000672	0.000658	0.000631
		Bias	-0.021694	-0.016319	-0.001743	-0.003457	-0.007434	-0.007489	-0.004794	0.002123	-0.004070	-0.004090	-0.003104	-0.001907
m=8	Y_{10}	MSPE	0.000787	0.000712	0.000261	0.000271	0.000402	0.000402	0.000387	0.000371	0.000309	0.000309	0.000305	0.000295
		Bias	-0.017853	-0.015357	-0.000679	-0.003128	-0.004693	-0.004723	-0.003256	0.002792	-0.002634	-0.002645	-0.002062	-0.001984
	Y_{11}	MSPE	0.001197	0.001065	0.000353	0.000357	0.000842	0.000843	0.000808	0.000762	0.000623	0.000623	0.000611	0.000587
		Bias	-0.020459	-0.016292	-0.000984	-0.002671	-0.007673	-0.007715	-0.005653	0.004162	-0.005007	-0.005023	-0.004231	-0.002256

adopting the CPs as the optimality criterion, the simulated CPs of PIs based on the pivotal quantity method are higher than those associated with the other methods in most considered cases. Moreover, one can see that the pivotal quantity PIs perform very well when compared to the Bayesian PIs in most the considered cases. However, Bayesian PIs outperform HCD and SL PIs in terms of CPs.

Table 10. CPs and ALs from simulations of $\lambda = 1$ and $\beta = 2$

m	Y_s	Criterion	Pivot	HCD	SL	BPIs			
						BE_{SE}	BE_{LE}		
							$\nu = -0.01$	$\nu = 0.5$	$\nu = 2$
m=5	y_7	CP	0.94	0.69	0.68	0.97	0.97	0.97	0.90
		AL	0.081834	0.068411	0.054774	0.091521	0.107440	0.095436	0.093739
	y_8	CP	0.95	0.61	0.63	0.99	0.99	0.99	0.97
		AL	0.08628	0.069513	0.068958	0.104510	0.121960	0.101750	0.096548
m=6	y_8	CP	0.96	0.78	0.76	0.971429	0.971429	0.971429	0.957143
		AL	0.067859	0.052533	0.046157	0.078527	0.099947	0.092039	0.090749
	y_9	CP	0.98	0.64	0.64	0.99	0.99	0.99	0.97
		AL	0.072816	0.058481	0.058322	0.087110	0.108130	0.085808	0.082046
m=7	y_9	CP	0.93	0.67	0.67	0.94	0.94	0.86	0.86
		AL	0.060137	0.046778	0.037941	0.071483	0.116410	0.066859	0.077446
	y_{10}	CP	0.97	0.67	0.67	0.98	0.99	0.80	0.78
		AL	0.065109	0.047001	0.045184	0.076375	0.124930	0.060163	0.067669
m=8	y_{10}	CP	0.99	0.86	0.84	0.97	0.97	0.94	0.88
		AL	0.049566	0.048812	0.038217	0.069515	0.130260	0.081471	0.081768
	y_{11}	CP	0.99	0.76	0.76	0.97	0.98	0.92	0.92
		AL	0.054562	0.050493	0.045384	0.076367	0.142340	0.070210	0.086240

Table 11. CPs and ALs from simulations of $\lambda = 2$ and $\beta = 1$

m	Y_s	Criterion	Pivot	HCD	SL	BPIs			
						BE_{SE}	BE_{LE}		
							$\nu = -0.01$	$\nu = 0.5$	$\nu = 2$
m=5	y_7	CP	0.95	0.73	0.69	0.93	0.93	0.91	0.90
		AL	0.138920	0.109680	0.093530	0.157380	0.157450	0.144780	0.133160
	y_8	CP	0.95	0.66	0.65	0.98	0.98	0.98	0.93
		AL	0.141140	0.122930	0.116990	0.175080	0.176500	0.156470	0.144770
m=6	y_8	CP	0.95	0.72	0.68	0.90	0.90	0.89	0.88
		AL	0.108720	0.095392	0.076062	0.121080	0.122890	0.117540	0.108920
	y_9	CP	0.96	0.69	0.72	0.91	0.91	0.88	0.87
		AL	0.112680	0.109600	0.107510	0.137210	0.138430	0.133780	0.125030
m=7	y_9	CP	0.95	0.79	0.76	0.89	0.89	0.87	0.87
		AL	0.092416	0.090503	0.066912	0.101680	0.104870	0.098422	0.090173
	y_{10}	CP	0.95	0.749	0.754	0.87	0.87	0.87	0.82
		AL	0.096927	0.096199	0.091193	0.114580	0.116100	0.111200	0.102800
m=8	y_{10}	CP	0.95	0.82	0.79	0.86	0.86	0.86	0.83
		AL	0.078396	0.074663	0.058354	0.084383	0.091762	0.081671	0.074334
	y_{11}	CP	0.95	0.75	0.77	0.91	0.91	0.91	0.90
		AL	0.083206	0.081510	0.080934	0.099775	0.099618	0.098116	0.090220

6. Conclusion

In this study, we have investigated the prediction of future records for the two-parameter Kies distribution. Both classical and Bayesian approaches were employed to develop point and interval predictors of the future records. The performance of these predictors was compared through Monte Carlo simulation studies. It was observed that, among all point predictors, the BUP showed the best performance in terms of bias, while the BUP and the CMP

Table 12. CPs and ALs from simulations of $\lambda = 2$ and $\beta = 2$

m	Y_s	Criterion	Pivot	HCD	SL	BPIs			
						BE_{SE}	BE_{LE}		
							$\nu = -0.01$	$\nu = 0.5$	$\nu = 2$
m=5	y_7	CP	0.96	0.81	0.77	0.91	0.91	0.90	0.88
		AL	0.092129	0.111960	0.062277	0.101150	0.127210	0.097757	0.088902
	y_8	CP	0.97	0.69	0.72	0.83	0.79	0.83	0.80
		AL	0.097837	0.086782	0.085165	0.117650	0.140730	0.114940	0.104500
m=6	y_8	CP	0.93	0.84	0.78	0.87	0.86	0.89	0.86
		AL	0.071806	0.081650	0.056904	0.086710	0.092262	0.084933	0.076393
	y_9	CP	0.94	0.72	0.71	0.87	0.87	0.86	0.83
		AL	0.078153	0.075926	0.073548	0.096249	0.098934	0.093238	0.085162
m=7	y_9	CP	0.96	0.73	0.72	0.93	0.91	0.91	0.90
		AL	0.065872	0.061657	0.045903	0.082202	0.129440	0.084075	0.076667
	y_{10}	CP	0.98	0.6	0.53	0.91	0.90	0.91	0.89
		AL	0.072133	0.066071	0.061670	0.094351	0.146290	0.092539	0.081635
m=8	y_{10}	CP	0.97	0.82	0.81	0.96	0.97	0.96	0.95
		AL	0.056281	0.053228	0.043322	0.084983	0.267850	0.091926	0.087719
	y_{11}	CP	0.97	0.78	0.75	0.91	0.92	0.90	0.90
		AL	0.062486	059308	059120	0.088974	0.250540	0.087110	0.078726

were quit close to each other in terms of MSPEs. The MLP and the MMLP also performed similarly.
Additionally, it was noted that the Bayesian predictors outperformed the MLP and MMLP in terms of both bias and MSPEs, especially under SE and LINEX loss functions. In the context of prediction intervals, the SL method was found to be the most suitable for obtaining PIs of the unobserved future records when adopting ALs as the optimality criterion. When adopting the CPs as the optimality criterion, it was observed that the pivotal quantity method proved to be an efficient technique for constructing PIs in most of the considered cases.

Acknowledgement

The authors would like to express their sincere gratitude to the anonymous referees for their insightful recommendations and remarks.

REFERENCES

1. M. Ahsanullah, *Record values theory and applications*, University Press of America, 2004.
2. M., Ahsanullah, and V.B. Nevzorov, *Records via probability theory*, Springer, Berlin, 2015.
3. M. Ahsanullah, and M. Z. Raqab, *Bounds and characterizations of record statistics*, Nova Publishers, 2006.
4. E.K. Al-Hussaini, and A.A. Ahmed, *On Bayesian interval prediction of future records*, Test, vol. 12, pp. 79–99, 2003.
5. N. M. Al-Olaimat, H. A. Bayoud, and M. Z. Raqab, *Record data from Kies distribution and related statistical inferences*, Statistics in Transition New Series, vol. 22, no. 4, pp. 153–170, 2021.
6. B.C. Arnold, N. Balakrishnan, and H.N. Nagaraja, *Records*, Wiley, New york, 1998.
7. A. M. Awad, and M. Z. Raqab, *Prediction intervals for the future record values from exponential distribution: comparative study*, Journal of Statistical Computation and Simulation, vol. 65 , no. 1-4 , pp. 325-340, 2000.
8. P. Basak, and N. Balakrishnan, *Maximum likelihood prediction of future record statistic, Mathmatical and statistical methods in reliablity*, World Scientific Publishing, vol. 7, pp. 159-175, 2003.
9. KN. Chandler, *The distribution and frequency of record values*, Journal of the Royal Statistical Society: Series B (Methodological), vol. 14, no. 2, pp. 220-228, 1952.
10. M. H. Chen, and Q. M. Shao, *Monte Carlo estimation of Bayesian credible and HPD intervals*, Journal of Computational and Graphical Statistics, vol. 8, no. 1, pp. 69-92, 1999.
11. S. Dey, T. Dey, and D. Luckett, *Statistical inference for the generalized Rayleigh distribution based on upper record values*, Model Assisted Statistics and Applications, vol 1. , no. 12, pp. 15-29, 2017.

12. I.R. Dunsmore, *The Future Occurrence of Records*, Ann. Inst. Staist. Math., vol. 35, pp. 267-277, 1983.
13. C. Empacher, U. Kamps, and G. Volovskiy, *Statistical prediction of future sports records based on record values*, Stats, vol. 6, no. 1, pp. 131-147, 2023.
14. A. I. Genc, *Estimation of $P(X > Y)$ with Topp-Leone distribution*, Journal of Statistical Computation and Simulation, vol. 83, no. 2, pp. 326-339, 2013.
15. J. A. Kies, *The strength of glass performance*, Naval Research Lab Report, 1958.
16. C.S. Kumar, and S.H.S. Dharmaja, *On Reduced Kies Distribution*, Collection of recent statistical methods and applications, vol. , no. , pp. 111-123, 2013.
17. C. S. Kumar, and S. H. S. Dharmaja, , *On some properties of Kies distribution*, Metron, vol. 72, pp. 97-122, 2014.
18. C.S. Kumar, and S.H.S. Dharmaja, *On Modified Kies Distribution and its Applications*, Journal of Statistical Research, vol. 51, no. 1, pp. 41-60, 2017.
19. C. S. Kumar, and S. H. S. Dharmaja, *The exponentiated reduced Kies distribution: Properties and applications*, Communications in Statistics – Theory and Methods, vol. 46, no. 17, pp. 8778-8790, 2017.
20. M.T. Madi, and M.Z. Raqab, *Bayesian prediction of temperature records using the Pareto model*, Environmetrics, vol. 15, no. 7, pp. 701-710, 2004.
21. H. N. Nagaraja, *Record values and related statistics-a review*, Communications in Statistics-Theory and Methods, vol. 17, no. 7, pp. 2223-2238, 1988.
22. L. E. Papke, and J. M. Wooldridge, *Econometric methods for fractional response variables with an application to 401(K) plan participation rates*, Journal of Applied Econometrics, vol. 11, no. 6, pp. 619-632, 1996.
23. M.Z. Raqab, *Predictors of Future Order Statistics From Type II Censored Samples*, The Ohio State Universeity, 1992.
24. M. Z. Raqab, *Optimal prediction-intervals for the exponential distribution, based on generalized order statistics*, IEEE Transactions on Reliability, vol. 50, no. 1, pp. 112-115, 2001.
25. M. Z. Raqab, J. Ahmadi, and M. Doostparastm, *Statistical inference based on record data from Pareto model*, Statistics, vol. 41, no. 2, pp. 105-118, 2007.
26. H. R. Varian, *A Bayesian approach to real estate assessment*, Studies in Bayesian econometrics and statistics in honor of Leonard J. Savage, 1975.
27. G. Volovskiy, and U. Kamps, *Maximum observed likelihood prediction of future record values*, Test, vol. 29, no. 4, pp. 1072-1097, 2020.