STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 202x, pp 0–35.
Published online in International Academic Press (www.IAPress.org)

A Comprehensive Review and Comparative Analysis of Prediction Methods for Future Order Statistics

R. A. Aldallal [1,*], H. M. Barakat [2], O. M. Khaled [3], H. H. Sakr [4,5], K. S. Khalil [6], A. M. Salm [6], M. M. Elarbawy [7]

[1] *Department of Management, College of Business Administration Hawtat Bani Tamim, Prince Sattam bin Abdulaziz University, Saudi Arabia*
[2] *Department of Mathematics, Faculty of Science, Zagazig University, Egypt*
[3] *Department of Mathematics, Faculty of Science, Port Said University, Egypt*
[4] *Department of Management Information Systems, College of Business Administration Hawtat Bani Tamim, Prince Sattam bin Abdulaziz University, Saudi Arabia*
[5] *Mathematics Department, Faculty of Education, Ain Shams University, Cairo, 11341 Egypt*
[6] *Department of Basic Science, Faculty of Computer and Informatics, Suez Canal University, Egypt*
[7] *Department of computer Science, Faculty of Science, Suez Canal University, Egypt*

Abstract Future ordered observation prediction is fundamental in reliability, survival analysis, and quality control studies, particularly when only early failures or partially observed data are available. This review and comparative analysis synthesizes recent prediction intervals and point predictors for future order statistics under fixed and mixture lifetime models, conditional expectation (near-MMSE) methods, and cumulative hazard least-squares frameworks. These approaches are further examined within the broader context of record values and extreme order statistics. The reviewed methods are compared with respect to their underlying assumptions, computational complexity, prediction targets, and practical limitations. A reproducible simulation study implemented in `R` benchmarks six selected procedures under fixed-size Type-II censoring, using Root Mean Squared Error (RMSE), bias, coverage probability, and interval width. The paper provides a critical comparative synthesis and a scope-limited benchmark, not a new prediction procedure, theorem, or model class; conclusions are conditional on the methods, models, censoring schemes, and targets examined. Finally, recent developments from 2015 to 2026 are surveyed, emphasizing advances motivated by complex censoring schemes, heterogeneous populations, and improved uncertainty quantification.

Keywords Review and Comparative Analysis, Future Order Statistics, Predictive Inference, Type-II Censoring, Reliability Modeling

AMS 2010 subject classifications 62G30, 62N05

DOI: 10.19139/soic-2310-5070-5034

1. Introduction

Predicting future observations from a sample is a classical yet still active problem in statistical inference. In lifetime experiments, the data naturally appear as ordered failure times: the weakest item fails first, then the second weakest, and so on. If only the first r failures are observed, the practical question is to predict a later failure time $X_{s:n}$, where $r < s \leq n$. Such a prediction is useful when a life test is stopped early due to cost, time constraints, censoring, or incomplete follow-up.

*Correspondence to: R. A. Aldallal (Email: r.eldallal@psau.edu.sa). Department of Management, College of Business Administration Hawtat Bani Tamim, Prince Sattam bin Abdulaziz University, Saudi Arabia

ISSN 2310-5070 (online) ISSN 2311-004X (print)
Copyright © 202x International Academic Press

The papers reviewed here share one core idea: future observations should be predicted from the sample's ordered structure rather than from the original sampling order. The main methodological families are: pivotal quantities based on transformed order statistics; near-MMSE conditional-expectation predictors; mixture models for heterogeneous lifetime data; and cumulative-hazard least-squares methods that combine estimation and prediction into a single optimization problem. These families are related, but they are not interchangeable. Pivotal methods are elegant when the distributional transform is available; near-MMSE predictors are attractive when the target is point prediction under squared error loss; mixture methods are flexible when a single distribution is too simple; hazard-based least squares is attractive for censored and generalized ordered structures; and sequential refitting methods are useful when the objective is to predict many future observations rather than one.

The broader literature on record values is important for this review because record values are ordered extremes observed through time, and many prediction ideas for future order statistics have close analogs in the record setting. The subject begins with the classical work of Chandler [39], who studied the distribution and frequency of record values in independent sequences. Limit and asymptotic questions were later developed by Resnick [65], and Galambos [47], while Nevzorov [60], Nagaraja [58], Ahsanullah [7], Ahsanullah [8], Arnold et al. [19], Nevzorov and Balakrishnan [59], and Nevzorov [61] organized record theory into a mature mathematical framework. These sources are not all prediction papers, but they explain why the transformed exponential spacing structure keeps reappearing in prediction intervals for future ordered data.

The connection with order statistics is equally important. Standard references such as Johnson and Kotz [52] and Balakrishnan and Cohen [22] provide the distributional and estimation background used when moving from complete samples to censored ordered samples. Comprehensive reviews on the prediction of order statistics and progressive censoring are provided by Kaminsky and Nelson [53] and Balakrishnan and Cramer [24]. The study of concomitants by David [42] also matters conceptually because prediction problems often involve not only the future ordered value itself but also variables attached to it, such as stress, covariates, or environmental conditions at the time of failure. Related developments for generalized or modified record schemes, including k-record values studied by Dziubdziela and Kopocinski [45] and distributional characterizations studied by Balakrishnan and Stepanov [23], show that ordered extremes can be studied under several sampling mechanisms, not only the usual complete sample order statistic model.

Prediction for records developed into its own line of work. Dunsmore [44] studied the future occurrence of records, which is one of the earliest explicit links between record theory and prediction. General frequentist prediction intervals and predictive distributions were further formalized by Lawless and Fredette [56]. Later, Raqab and Balakrishnan [4] developed prediction intervals for future records, while Ahmadi and Balakrishnan [5] used record coverage to construct distribution-free prediction intervals for order statistics. For generalized order statistics under random sample sizes, Basiri and Ahmadi [36] constructed exact prediction intervals. Raqab [64] extended distribution-free prediction ideas to future current record statistics. These papers motivate an important distinction used throughout the present review: some methods target a future point value, while others target an interval with a required coverage probability.

Current records are particularly useful when both upper and lower extremes are monitored. The concept was initiated by Houchens [50] and later developed through record-range and current-record studies, including Basak [35], Aldallal [10], Barakat et al. [26], and Barakat et al. [28]. Exact prediction intervals for future current records and record ranges from any continuous distribution were given by Barakat et al. [29]; this is especially relevant here because it uses the same broad principle as pivotal order statistic prediction: transform the data to a scale where the distributional calculation becomes manageable, then invert the result back to the original lifetime scale. More recent work by Aldallal [12], Aldallal [13], and Aldallal [14] shows that estimation, moment recurrence, Bayesian prediction, and applied industrial data analysis remain active parts of the current record literature.

Applications also help explain why predicting ordered extremes is not merely a theoretical exercise. Record models appear in sports and performance analysis, as discussed by Glick [48], financial monitoring and technical trading contexts, as discussed by Brock et al. [38], reliability and preservation of aging properties, as discussed by Ahmadi and Balakrishnan [3], and stochastic process extremes such as random walks and Levy flights, as discussed by Wergen et al. [68]. The recent review of current records by Aldallal and Barakat [9] compiles many of these developments and emphasizes predictive inference as a major component of the record-value methodology. In this

paper, we use the broader literature as supporting context, but we keep the main focus on predicting future order statistics from censored or partially observed samples.

The present review is written from the perspective of a user who must choose a prediction method, not only from the perspective of a reader who wants to classify formulas. This distinction is important. In applications, a method is rarely selected solely because it is algebraically elegant. It is selected because it answers a particular practical question: a point estimate of the next failure, a conservative upper prediction bound, a full prediction interval with target coverage, or a sequence of future ordered failures. Therefore, the comparison below emphasizes the prediction target, the required distributional inputs, the available data structure, and the numerical steps needed to reproduce the method.

Another purpose of the review is to connect theory with reproducibility. Several of the cited papers contain exact or semi-exact distributional arguments, but the applied statistician still has to fit a model, compute inverse distribution functions, solve nonlinear equations, and decide how to report error measures. For this reason, the paper includes both a synthesis of the literature and a reproducible simulation/application component. The simulation is not intended to replace the original proofs. Instead, it acts as a common testing ground where pivotal, near-MMSE, mixture, hazard-based, and sequential ideas can be compared under the same notation and performance criteria.

The paper

i- organizes representative methods by prediction target, information structure, and distributional assumptions;
ii- compares six selected published procedures in a reproducible simulation under the scenarios specified here; and
iii- illustrates their differing predictions using aircraft windscreen and voltage stress data.

The resulting guidance is conditional on the prediction objective, censoring scheme, prediction horizon, and model assumptions. Some studies address linear unbiased Kumaraswamy record prediction, as in Aldallal [11], Kies record point and interval prediction, as in Al-Olaimat et al. [17], and copula-based nonparametric predictive inference for diagnostic tests, as in Muhammad et al. [57].

1.1. Literature search strategy

The present study is intended as a narrative review supplemented by a comparative simulation study and two real data applications. To identify representative contributions to the prediction of future order statistics, a broad literature search was conducted using major scientific databases, including Web of Science, Scopus, Google Scholar, MathSciNet, and publisher databases for statistics, reliability, and applied probability journals.

The search focused on publications on future order statistics, predictive inference, censored samples, record values, generalized order statistics, random sample-size models, mixture lifetime distributions, and reliability prediction. Representative search terms included:

> ‘future order statistics”, ‘prediction intervals”, ‘predictive inference”, ‘Type-II censoring”, ‘progressive censoring”, ‘record values”, ‘generalized order statistics”, ‘random sample size”, ‘mixture distributions”, ‘stress strength reliability”, ‘conditional expectation prediction”, ‘Bayesian prediction”, and “reliability prediction”.

The review primarily emphasizes peer-reviewed journal articles and foundational references that contributed substantially to the theoretical or methodological development of predictive procedures for future ordered observations. Particular attention was given to studies proposing exact prediction intervals, pivotal methods, conditional expectation predictors, cumulative hazard approaches, Bayesian methodologies, and prediction procedures under censoring and heterogeneous lifetime models.

The objective of the review is not to provide an exhaustive bibliographic catalog of every published contribution. Instead, the emphasis is placed on representative and influential developments that illustrate the major methodological directions within the field. When several closely related papers addressed essentially the same prediction framework, preference was given to the original contribution and to subsequent articles that introduced significant methodological extensions or practical applications.

The collected literature was subsequently organized into thematic categories, including pivotal prediction methods, conditional expectation prediction, mixture distribution approaches, cumulative hazard prediction procedures, and recent developments in predictive inference. This classification provides the foundation for the comparative discussion presented throughout the remainder of the paper.

1.2. Positioning relative to selected reviews and reference works

The literature includes reviews and reference works with complementary scopes. Kaminsky and Nelson [53] address prediction of order statistics; Aggarwala [2] reviews progressive censoring, and the monograph by Balakrishnan and Cramer [24] provides a broader account of its theory, methods, and applications. Cramer [41] surveys ordered and censored lifetime data in reliability, whereas Aldallal and Barakat [9] focuses on current record values and their applications. The present paper draws on these related areas but keeps the prediction of future order statistics as its central inferential target. It brings selected published procedures together to discuss their assumptions and targets, compare their performance under the simulation scenarios considered, and illustrate their differences using aircraft windscreen and voltage stress data. Table 1 compares scope rather than quality or completeness; it does not claim priority or that earlier sources lack relevant contributions.

Table 1. Positioning of the present paper relative to selected reviews and reference works.

Source	**Primary focus**	**Relation to the present paper**
Kaminsky and Nelson [53]	Prediction of order statistics	Directly addresses order statistic prediction; the present paper provides a comparative treatment of selected method families and a reproducible benchmark.
Aggarwala [2]	Progressive censoring review	Surveys progressive censoring research; the present paper centers on the prediction of future ordered observations across selected information structures.
Balakrishnan and Cramer [24]	Progressive censoring: theory, methods, and applications	Provides a broad reference on censoring; the present paper focuses on the predictive target and comparison of selected procedures.
Cramer [41]	Ordered and censored lifetime data in reliability	Surveys data structures, lifetime models, and selected inferential results; the present paper narrows its focus to prediction methods and their empirical comparison.
Aldallal and Barakat [9]	Current record values and their applications	Reviews current record methodology; the present paper treats records as context while keeping future order statistics as its principal target.
Present paper	Future order statistic prediction across selected method families	Combines a critical synthesis, a reproducible simulation comparison, and aircraft windscreen and voltage stress data applications.

2. Preliminaries and Notation

Let $X_1, \ldots, X_n$ be independent lifetimes from a continuous distribution function F, survival function $\overline{F}(x) = 1 - F(x)$, and density f. The ordered sample is written as

$$X_{1:n} < X_{2:n} < \cdots < X_{n:n}. \tag{1}$$

The observed data are the first r order statistics,

$$\tilde{x}_r = (x_{1:n}, \ldots, x_{r:n}), \qquad r < s \leq n,$$

and the target is the future order statistic $X_{s:n}$. When the total sample size is random, the notation becomes $X_{s:N}$, and all prediction statements are interpreted conditionally on $N \geq s$.

A central transformation in the reviewed papers is the cumulative hazard transform

$$H(x) = -\log \overline{F}(x). \tag{2}$$

If F is continuous, then $H(X_i)$ follows a standard exponential distribution. For order statistics, define

$$Y_{i:n} = H(X_{i:n}), \qquad Y_{0:n} = 0.$$

The normalized spacings are

$$Z_i = (n-i+1)\left(Y_{i:n} - Y_{i-1:n}\right), \qquad i = 1, \ldots, n. \tag{3}$$

These spacings are independent standard exponential random variables. This result is the backbone of the pivotal prediction intervals in Barakat et al. [27] and related papers. It means that prediction can be built on transformed spacings whose distribution does not depend on the original F, provided F or a reliable estimate of it is available.

The same transformation idea also explains why order-statistic prediction and record-value prediction are closely related. In record theory, the observed information is sparse because only new extremes are retained; in Type-II censoring, the observed information is sparse because the experiment stops after the first r ordered failures. In both cases, the statistical task is to use a structured partial history to infer an unobserved future ordered quantity. This is why record-based prediction intervals studied by Raqab and Balakrishnan [4], Barakat et al. [29], Dunsmore [44], Raqab [64] are useful background for the methods reviewed here, even when the present target is $X_{s:n}$ rather than a future record value.

2.1. Pivotal quantities for fixed sample size

For a fixed sample size n, define the future transformed spacing

$$W_{r,s}^n = Y_{s:n} - Y_{r:n} = H(X_{s:n}) - H(X_{r:n}), \tag{4}$$

and the accumulated observed spacing

$$T_{r,n} = \sum_{j=1}^{r} Z_j. \tag{5}$$

Two common pivotal quantities are then

$$U_{r,s}^n = \frac{W_{r,s}^n}{Y_{r:n}}, \qquad V_{r,s}^n = \frac{W_{r,s}^n}{T_{r,n}}. \tag{6}$$

Thus, $V_{r,s}^n$ is not just a symbol; it is the future hazard increment divided by the accumulated normalized information in the first r failures. The U-quantity depends mainly on the last observed transformed failure, whereas the V-quantity uses all observed normalized spacings through $T_{r,n}$.

If u_α is an upper quantile of $U_{r,s}^n$, then

$$0 \le U_{r,s}^n \le u_\alpha \quad \Longleftrightarrow \quad x_{r:n} \le X_{s:n} \le F^{-1}\left(1 - \overline{F}(x_{r:n})^{1+u_\alpha}\right). \tag{7}$$

Similarly, if v_α is an upper quantile of $V_{r,s}^n$, then

$$0 \le V_{r,s}^n \le v_\alpha \quad \Longleftrightarrow \quad x_{r:n} \le X_{s:n} \le F^{-1}\left(1 - e^{-v_\alpha t_{r,n}}\overline{F}(x_{r:n})\right), \tag{8}$$

where $t_{r,n}$ is the observed value of $T_{r,n}$. Equations (7) and (8) show why the pivotal quantities are useful: after a quantile is found, the prediction interval is obtained by inverting the distribution function.

2.2. Near-MMSE conditional expectation predictors

For point prediction, squared error loss leads naturally to conditional expectation. The manuscript by Barakat et al. [34] uses this idea under Type-II censoring and proposes two near-MMSE predictors for $X_{s:n}$ based on the last observed order statistic $X_{r:n} = x_{r:n}$. The first predictor is derived on the exponential spacing scale. Let

$$p_r = F(x_{r:n}), \qquad h_{r,s,n} = \sum_{j=n-s+1}^{n-r} \frac{1}{j}.$$

Then the exponential-based conditional expectation predictor is

$$\widehat{X}_{s:n}^{\text{CE-exp}} = F^{-1}\left(1 - \{1 - p_r\}\exp(-h_{r,s,n})\right). \tag{9}$$

This is the method implemented in Barakat et al. [34], where the Weibull distribution is used through $F(x) = 1 - \exp\{-(x/\lambda)^k\}$.

The second predictor uses the conditional mean of uniform order statistics. Given $F(X_{r:n}) = p_r$, the future transformed order statistic has conditional mean

$$p_r + (1 - p_r)\frac{s - r}{n - r + 1}.$$

Therefore,

$$\widehat{X}_{s:n}^{\text{CE-unif}} = F^{-1}\left(p_r + (1 - p_r)\frac{s - r}{n - r + 1}\right). \tag{10}$$

Equations (9) and (10) are especially useful as point prediction benchmarks because they do not require solving the pivotal distribution of $U_{r,s}$ or $V_{r,s}$; once F is known or fitted, the predictors are obtained by a direct quantile inversion.

2.3. Random sample size

Several papers replace fixed n by a positive integer-valued random variable N, independent of the lifetimes. This is important in biological, agricultural, medical, and quality-control settings where the number of available observations may vary due to loss, censoring, or random enrollment. In that setting,

$$W_{r,s}^N = H(X_{s:N}) - H(X_{r:N}), \qquad V_{r,s}^N = \frac{W_{r,s}^N}{T_{r,N}}, \tag{11}$$

where the distribution is truncated to $N \geq s$. A useful way to read the random size formulas is as a weighted mixture over possible fixed sample sizes:

$$F_{Q^N}(q) = \frac{1}{\mathbb{P}(N \geq s)} \sum_{n=s}^{\infty} F_{Q^n}(q)\mathbb{P}(N = n), \tag{12}$$

where Q may be $U_{r,s}$ or $V_{r,s}$. Specific papers derive explicit forms of (12) when N follows negative binomial, binomial, or Poisson laws; see Barakat et al. [25, 27, 30]. The practical implication is simple: a random sample size usually produces wider intervals than a fixed nominal sample size, because the prediction must account for both lifetime variability and count variability.

2.4. Mixture lifetime models

Single lifetime distributions may be too rigid when the population is heterogeneous. A two-component mixture distribution has a density and a distribution function

$$f(x;\Theta) = \pi f_1(x;\theta_1) + (1 - \pi)f_2(x;\theta_2), \qquad 0 < \pi < 1, \tag{13}$$

$$F(x;\Theta) = \pi F_1(x;\theta_1) + (1-\pi)F_2(x;\theta_2). \tag{14}$$

The mixture exponential paper studies future lifetime prediction when F_1 and F_2 are exponential components, as described by Barakat et al. [32]. The gamma mixture and beta mixture paper extends this idea to more flexible component models and applies it to COVID-19 recovery rate data, as described in Khaled et al. [54]. The main computational difficulty is that mixture quantile functions are often not available in closed form. Therefore, prediction intervals usually require numerical inversion or Monte Carlo simulation from the fitted mixture distribution.

2.5. Cumulative hazard least squares

The cumulative hazard least squares method uses $H(x;\Theta)$ directly for estimation and prediction. Let

$$\mu^*_{i:n} = \mathbb{E}\{H(X_{i:n};\Theta)\}, \qquad (w^*_i)^{-1} = \mathrm{Var}\{H(X_{i:n};\Theta)\}.$$

For the first r observed order statistics, a weighted criterion can be written as

$$\mathrm{WL}^*_{H,r}(\Theta \mid \tilde{x}_r) = \sum_{i=1}^{r} w^*_i \left[H(x_i;\Theta) - \mu^*_{i:n}\right]^2 + (n-r)w^*_r \left[H(x_r;\Theta) - \mu^*_{r:n}\right]^2. \tag{15}$$

To predict $x_{s:n}$, the predictive criterion adds the unobserved point:

$$\begin{aligned}\mathrm{PWL}^*_{H,r,s}(\Theta, x_s \mid \tilde{x}_r) = &\sum_{i=1}^{r} w^*_i \left[H(x_i;\Theta) - \mu^*_{i:n}\right]^2 \\ &+ (s-r-1)w^*_r \left[H(x_r;\Theta) - \mu^*_{r:n}\right]^2 \\ &+ (n-s+1)w^*_s \left[H(x_s;\Theta) - \mu^*_{s:n}\right]^2.\end{aligned} \tag{16}$$

The predictor is obtained by minimizing (16) over Θ and x_s. This framework is especially useful for generalized ordered statistics and censored samples, and it is compared with maximum likelihood and ordinary weighted least squares using RMSE and Pitman's measure of closeness by Aly et al. [18].

2.6. A unified taxonomy for order statistic prediction methods

Despite the diversity of prediction methods in the literature, they can be organized along three useful dimensions. This descriptive taxonomy helps structure the comparison and the discussion of method selection; it is an organizing device, not a standalone methodological contribution.

These dimensions answer three distinct questions: what is predicted, which observations are available, and which distributional assumptions justify inference. They are useful organizing axes, not an exhaustive or uniquely optimal classification. Data requirements fall within information structure; computational cost and sensitivity to model misspecification are additional, cross-cutting evaluation criteria discussed in Sections 4, 6.2, and 8.

Dimension 1: Prediction Objective Methods differ fundamentally in what they aim to produce:

- *Point predictors:* Single numerical estimates (e.g., conditional expectation, median, mode).
- *Interval predictors:* Sets with nominal coverage probability (e.g., pivotal intervals, Bayesian credible intervals).
- *Predictive distributions:* Full density or distribution function for the future observation (e.g., Bayesian predictive density, nonparametric predictive inference).

Dimension 2: Information Structure The sampling and censoring scheme determines what information is available:

- *Fixed sample size Type-II censoring:* Exactly r failures observed from a fixed sample of size n.
- *Random sample size:* Sample size N is random, independent of lifetimes.
- *Record schemes:* Only record breaking observations are retained.
- *Progressive censoring:* Units may be removed at intermediate failure times.

Dimension 3: Distributional Assumption The strength of distributional assumptions varies considerably:

- *Fully parametric:* A specific parametric family is assumed (e.g., Weibull, exponential, gamma).
- *Semi-parametric :* The hazard transform or other structural assumptions are made, but not a full parametric form.
- *Nonparametric/distribution free:* No specific distributional form is assumed; validity holds across broad classes of distributions.

Mapping the Literature Table 2 maps the major methods reviewed in this paper onto this taxonomy. The axes describe different inferential requirements rather than establish an overall ranking. Near-MMSE predictors require a specified distribution for back transformation. For the reviewed U- and V-pivotal constructions, the transformed statistic has an F-free distribution, but prediction on the original scale requires F; this does not establish robustness to misspecification. The implemented cumulative hazard least squares procedure is parametric, and the bootstrap assumptions depend on the resampling scheme. Bayesian methods offer full predictive distributions but require model and prior assumptions and computational effort.

Table 2. Taxonomy of Prediction Methods for Future Order Statistics

Method	**Prediction Objective**	**Information Structure**	**Distributional Assumption**
Pivotal (U, V)	Interval	Fixed Type-II censoring	F-free pivot; specified F for inversion
Near-MMSE Conditional Expectation	Point	Fixed Type-II censoring	Parametric
Random Sample Size Pivotal	Interval	Random sample size	Specified F and sample size law
Mixture Based Prediction	Point/Interval	Fixed Type-II censoring	Parametric mixture
Cumulative Hazard Least Squares	Point in this benchmark	Fixed Type-II censoring	Parametric in this implementation
Bayesian Prediction	Full distribution	Various (including progressive)	Parametric
Nonparametric Predictive Inference	Interval	Various	Distribution free
Bootstrap Prediction	Interval	Various	Depends on resampling scheme

The taxonomy is used to organize the comparison, not as a standalone methodological contribution. It makes explicit how the reviewed procedures differ in their prediction targets, information structures, and distributional assumptions. The simulation study and the aircraft windscreen and voltage stress applications compare selected methods under their respective settings; their findings are conditional on those settings and do not establish a universally best method.

3. Comparative Review of Prediction Methods

The literature on the prediction of future order statistics has evolved through several methodological directions. Some studies established the theoretical foundations of pivotal prediction procedures, while others focused on conditional expectation-based predictors, generalized order statistics, record values, and censored samples. To provide a structured overview, the reviewed studies are grouped according to their primary methodological contributions. Table 3 summarizes the main characteristics of the most relevant works.

Table 3. Focused synthesis of the reviewed literature.

Reference	Main setting	Main contribution	Critical comment
Barakat et al. [25]	Exponential lifetimes; random generalized order statistics	Exact prediction intervals using pivotal quantities under random N	Strong theory, but tied to exponential/ generalized order structure.
Barakat et al. [27]	Any continuous distribution; fixed and random sample size	Distribution free transformation through normalized spacings and intervals based on $U_{r,s}$, $V_{r,s}$	Important foundation; practical use depends on fitting or estimating F.
Barakat et al. [30]	Exponential generalized order statistics with random sample size	Extends random size prediction for generalized ordered data	Useful for censored and record type schemes, but formulas are computationally heavy.
Barakat et al. [32]	Mixture of two exponential distributions	Point and interval prediction under heterogeneous lifetime populations	Better for heterogeneity than a single exponential, but still limited in shape flexibility.
Khaled et al. [54]	Gamma mixture and beta mixture distributions	Simulation-based prediction intervals; COVID-19 recovery rate application	Flexible and applied, but mixture quantiles require numerical work and careful estimation.
Barakat et al. [34]	Type-II censoring; any continuous distribution	Near-MMSE point predictors from conditional expectation on exponential and uniform scales	Strong point prediction benchmark; interval construction and far tail behavior still need careful assessment.
Barakat et al. [31]	Future order statistics; point prediction	New point predictors evaluated by simulation and real data	Useful comparator when the target is point prediction rather than interval coverage.
Aly et al. [18]	Generalized ordered statistics; cumulative hazard	Modified least squares estimation and prediction through $H(x)$	Strong bridge between estimation and prediction, especially under censoring.
Harpy et al. [49]	Iterative prediction with model reselection in R	Refits distribution after each predicted value; compares pivotal and least squares methods	Good candidate for reproduction because it already frames an R-based comparative algorithm.
El-Adll et al. [46]	Three-parameter Weibull generalized order statistics	Prediction intervals for Weibull future lifetimes	Useful parametric benchmark for reliability examples.
Barakat et al. [33]	Modified Muth distribution	Parsimonious lifetime model with estimation study and entropy measures	Valuable as a candidate lifetime model, but it is not itself a full future order prediction method.

3.1. Pivotal methods

Pivotal methods are mathematically attractive because they transform the prediction problem into the distribution of a statistic whose law is known or can be computed. The work of Barakat et al. [27] is central because it begins with an arbitrary continuous distribution and uses the hazard transform to obtain exponential normalized spacings. This gives a clear path from theory to prediction intervals: define U or V, derive or numerically evaluate its distribution, obtain a quantile, then invert F to produce the interval.

The advantage is transparency. The lower bound is usually the last observed failure $x_{r:n}$, and the upper bound is a transformed quantile. The weakness is that the method needs a reliable F or fitted $\widehat{F}$. If F is chosen poorly, the formal coverage may not hold in practice. This is why goodness-of-fit checking and model selection are not optional; they are part of the prediction method.

3.2. Near-MMSE conditional expectation

The near-MMSE conditional expectation approach of Barakat et al. [34] is different in emphasis from pivotal interval construction. Instead of first deriving the distribution of a pivotal statistic, it starts with the squared-error principle: the best predictor of a future random variable, among functions of the observed information, is its conditional expectation. The paper turns this principle into two practical predictors: one based on exponential spacings and another based on uniform order-statistic means.

Its strength is simplicity. Both formulas reduce prediction to evaluating $F(x_{r:n})$, applying a deterministic correction factor depending on r, s, n, and then using F^{-1}. This makes the method easy to implement in R and a natural competitor in the simulation section. Its limitation is that, as with the pivotal methods, performance depends on the chosen or fitted distribution F, and very-far-tail prediction can still be unstable.

3.3. Random sample size

Random sample size is not a minor technical extension. In fixed sample prediction, uncertainty comes from the future lifetime conditional on n. In random size prediction, uncertainty also comes from the distribution of N. The reviewed papers treat this by averaging over possible values of N, often under negative binomial, binomial, or Poisson assumptions, as in Barakat et al. [25, 27, 30]. In applied terms, this is useful when the number of test units, patients, or observed failures cannot be fixed in advance.

The primary disadvantage is increased computational complexity. Closed-form expressions can be long, and quantiles often require numerical solution. For a review paper, it is better to present the conceptual mixture formula (12), identify the count distributions used in the literature, and then compare their effect through simulation rather than reproducing every expanded series.

3.4. Mixture models

Mixture models address a different problem: heterogeneity in the lifetime distribution. A single exponential, Weibull, or gamma model can be misleading if the population contains subgroups with different failure mechanisms. The mixture exponential work provides an early focused treatment for future lifetime prediction in a two-component mixture, as described by Barakat et al. [32]. The gamma mixture and beta mixture paper generalizes this to richer shapes and bounded data, including an application to COVID-19 recovery rates, as described by Khaled et al. [54].

The main strength is flexibility. The main weakness is estimation instability: mixture parameters and weights can be difficult to estimate from small censored samples, and numerical inversion of $F(x; \widehat{\Theta})$ may be required. A good review should therefore treat mixture prediction as a simulation-based alternative rather than a simple closed-form replacement for pivotal methods.

3.5. Cumulative hazard least squares

The cumulative hazard approach of Aly et al. [18] is important because it gives a unified criterion for estimation and prediction. Instead of first estimating parameters and then separately predicting $X_{s:n}$, the predictive weighted

least squares function (16) can be minimized jointly over the parameters and the future value. This is attractive for generalized ordered statistics because the hazard scale simplifies the theoretical moments used in the criterion.

Compared with maximum likelihood, this method can be computationally convenient and robust in some settings. However, it still requires an explicit cumulative hazard $H(x;\Theta)$. It is therefore best used with distributions such as Weibull, modified Kies exponential, Gompertz Makeham, exponential families, and other lifetime models with tractable hazard functions.

3.6. Sequential refit and predict algorithms

The newer iterative method described by Harpy et al. [49] has a practical flavor: fit a distribution, predict the next value, append it to the data, refit, and repeat. This can produce many predicted observations, sometimes beyond the original observed count. It is attractive for applications where the goal is a predicted sequence rather than a single future order statistic.

The method should be used carefully. Appending predicted values can make later predictions depend on earlier model errors. Therefore, a review should not present this as automatically superior. It should be evaluated by simulation against simpler methods using MSE, coefficient of variation, interval width, and coverage. This makes it a useful reference point for the reproducible R comparison, even when the final simulation keeps the number of compared methods intentionally compact.

4. Discussion of Methodological Differences

The prediction methods reviewed in the previous sections differ in their objectives, assumptions, and computational requirements. Some procedures are designed for point prediction, whereas others focus on prediction intervals. In addition, the methods differ according to the assumed sampling framework, the treatment of distributional assumptions, and the nature of the available information. These distinctions are important when evaluating the applicability and performance of competing prediction procedures.

4.1. Point prediction versus interval prediction

The first decision is whether the analysis needs a point prediction or an interval prediction. Near-MMSE conditional expectation predictors are naturally suited to point prediction because they are motivated by squared error loss. Their output is a single predicted value, usually close to the conditional mean of the future order statistic under the fitted model. This makes them easy to compare through RMSE, MAE, bias, and coefficient of variation. They are especially attractive when a practitioner needs one expected future failure time for planning or replacement decisions.

Pivotal methods answer a different question. They are designed to construct intervals with a specified probability statement. The user obtains a lower and upper bound for $X_{s:n}$, often with the lower bound equal to the last observed failure and the upper bound obtained from a pivotal quantile. These intervals can be conservative or wide, but their purpose is uncertainty quantification rather than only the closeness of the center. Therefore, it is misleading to compare a pivotal interval method with a point predictor using RMSE alone. A fair comparison should report point accuracy for point methods and coverage probability, together with average width for interval methods.

Under a known continuous F and fixed Type-II censoring, the pivotal and near-MMSE procedures share a transformed order statistic structure, as described by Barakat et al. [27, 34]. Pivotal intervals calibrate probabilities via spacing-ratio quantiles, whereas near-MMSE predictors back-transform conditional means on the hazard or uniform scale. For $g = H$ or F, $g^{-1}\{\mathbb{E}[g(X_{s:n}) \mid \tilde{x}_r]\}$ need not equal $\mathbb{E}[X_{s:n} \mid \tilde{x}_r]$; only the latter is guaranteed to minimize conditional squared error loss when second moments exist. Thus, coverage calibration does not imply optimal point accuracy, nor does low RMSE guarantee interval coverage. A common transformation alone does not establish asymptotic equivalence, which would require a specified growth regime for n, r, s and regularity assumptions on F.

4.2. Distributional choice and quantile inversion

All reviewed parametric procedures depend, explicitly or implicitly, on a distribution function F, a survival function $\overline{F}$, a cumulative hazard H, or an inverse quantile function F^{-1}. In the theoretical derivations, this dependence may appear simple because the hazard transform converts the sample to exponential spacings. In an applied analysis, however, the analyst must still choose and estimate the lifetime model. For Weibull, exponential, gamma, beta, modified Kies exponential, and related distributions, this choice changes both the fitted probabilities $F(x_{r:n})$ and the back-transformed predictions.

The practical implication is that model checking should be treated as part of prediction, not as a preliminary decoration. A method that is exact under the assumed F can perform poorly when the fitted distribution misses tail behavior. This issue is particularly important when predicting large s, because the prediction relies on high quantiles, where small distributional errors are magnified. In real data applications, it is therefore useful to report the fitted parameters, the form of F, and whether the quantile was computed analytically or numerically.

4.3. Censoring level and prediction horizon

The difficulty of prediction increases with the gap $s - r$. Predicting $X_{r+1:n}$ from $X_{1:n}, \ldots, X_{r:n}$ is a local extrapolation; predicting $X_{n:n}$ from the same information is a tail extrapolation. Pivotal, near-MMSE, and hazard-based methods all reflect this fact through wider intervals, larger expected increments on the hazard scale, or greater numerical instability in the inverse transform. This is why a simulation study should include both short- and long-horizon prediction targets rather than reporting only one favorable case.

The amount of censoring also matters. When r is large, the observed ordered history contains substantial information about the tail; when r is small, the method depends more heavily on the assumed model. This distinction is important for Type-II censoring. A method that performs well when half the sample is observed may no longer be reliable when only the first few failures are available. The simulation evaluates two prespecified censoring and target designs, while the real data applications report predictions over their stated future index ranges. Results should therefore be interpreted within these examined settings, rather than as a general ranking across all censoring levels or prediction horizons.

4.4. Practical method selection guide

Table 4 converts the preceding discussion into conditional suggestions for common prediction settings. It is an evidence-informed guide based on the reviewed assumptions and the comparisons reported here, not a separately validated decision rule. The recommendations, therefore, identify candidate methods and checks rather than guarantee that one procedure will perform best.

The guide establishes neither outlier robustness nor a minimum sample size threshold for stable performance. Computational requirements are discussed qualitatively rather than ranked by measured runtime.

4.5. When methods fail: Diagnostic patterns

Understanding when a prediction method is likely to fail is as important as knowing when it works. Based on the simulation results, the literature synthesis, and real-world data, we identify the following failure patterns.

Pivotal Methods (U and V) *Failure mode:* Low empirical coverage when the assumed distribution F is misspecified, particularly in the tail.

Diagnostic indicators:

- Poor fit of the distribution in the upper tail (e.g., QQ plot deviation).
- Sensitivity of interval width to small changes in estimated parameters.
- Coverage probability significantly below nominal in simulation.

Remedies: Use goodness of fit tests; consider distribution-free alternatives; apply bootstrap calibration.

Table 4. Conditional guidance for selecting a prediction method.

Prediction setting	Candidate method	Checks and limitations
Point prediction under fixed Type-II censoring; a parametric model is reasonably supported and the target is local or intermediate.	Near-MMSE exponential or uniform conditional expectation predictor.	Check model fit and report bias and point prediction error. Use caution with far-tail targets, where point predictions may be sensitive to the fitted tail.
An interval with a stated nominal coverage is required.	U- or V-pivotal prediction interval.	Report nominal and empirical coverage, along with the average width. If parameters or F are estimated, treat the result as a plug-in and assess the effect of model fitting.
Parameters must be estimated from a Type-II censored prefix as part of the prediction.	Cumulative hazard MWLS-PWLS.	Check that the chosen hazard form is suitable; report starting values, convergence, and fit diagnostics. Its inclusion does not imply uniform superiority over other procedures.
The number of units is random, and its count distribution is substantively supported.	Random sample size prediction procedure.	State and assess the assumed distribution of N, and ensure the target order statistic exists under the modeled sample size.
There is credible evidence of population heterogeneity.	Mixture-based prediction.	Check component identifiability, convergence, and sensitivity to starting values; small censored samples may not support stable mixture estimation.
The model is uncertain, or the target is far into the tail.	Compare plausible combinations of model methods; do not select from a single overall ranking.	Report the sensitivity of predictions and uncertainty to model choice and horizon. Large disagreement is a warning that the available data do not support a precise ranking.

Near-MMSE conditional expectation Predictors *Failure mode:* The point prediction may be pulled toward the center of the distribution, leading to underprediction of extreme future order statistics.

Diagnostic indicators:

- Systematic negative bias for high s (far tail prediction).
- RMSE increasing sharply as $s \to n$.
- Sensitivity to the choice of distribution family.

Remedies: For far tail prediction, consider median-based or quantile-based predictors; use bootstrap or Bayesian averaging to account for model uncertainty.

Random Sample Size Methods *Failure mode:* Inflated interval width and poor coverage if the assumed distribution for N is misspecified.

Diagnostic indicators:

- Discrepancy between the assumed and empirical distribution of sample size.
- Overdispersion in observed counts relative to the assumed model.

Remedies: Use empirical distribution of N when available; conduct sensitivity analysis over plausible count distributions.

Mixture Based Methods *Failure mode:* Numerical instability and non-identifiability, especially with small samples or overlapping components.

Diagnostic indicators:

- High correlation between estimated component parameters.
- Convergence failures in the estimation algorithm.

- Sensitivity of predictions to starting values.

Remedies: Use penalized or Bayesian estimation; impose identifiability constraints; consider profile likelihood.

Cumulative Hazard Least Squares *Failure mode:* Suboptimal performance when the hazard function is not well approximated by the assumed parametric form.

Diagnostic indicators:

- Poor linearity of the cumulative hazard plot.
- Large residuals in the weighted least squares criterion.
- Instability of predictions under small perturbations of the observed data.

Remedies: Consider flexible hazard models; use cross validation; check residual plots.

General Principle The most important diagnostic is to compare predictions under multiple plausible models. If different methods or model choices produce substantially different predictions, this indicates that the data are insufficient to support strong conclusions and that uncertainty quantification should be increased.

5. Other Recent Prediction Methods

In addition to the pivotal, conditional expectation, mixture-based, and hazard-oriented procedures discussed earlier, several other important prediction paradigms have gained traction over the past decade. These developments have been driven by increasingly complex censoring schemes, heterogeneous populations, computational advances, and a growing need for uncertainty quantification in reliability and survival studies. Below, we review some of the most active research directions that have emerged between 2015 and 2026.

5.1. Bayesian prediction methods

Bayesian predictive inference has become a cornerstone of modern prediction theory because it naturally incorporates parameter uncertainty into future predictions. Unlike classical plug-in approaches, which replace unknown parameters with point estimates, Bayesian methods average over the posterior distribution of the parameters, thereby producing predictive distributions that reflect estimation uncertainty.

For lifetime models, the Bayesian predictive density of a future order statistic $Y_{s:m}$ given observed data $\mathbf{x}$ is

$$p(y_{s:m} \mid \mathbf{x}) = \int p(y_{s:m} \mid \Theta)\, \pi(\Theta \mid \mathbf{x})\, d\Theta,$$

where $\pi(\Theta \mid \mathbf{x})$ denotes the posterior distribution of the model parameters.

Recent contributions have focused on Bayesian prediction under progressively censored samples, competing-risk models, and record-value schemes. For instance, Al-Hussaini and colleagues developed Bayesian prediction intervals for future order statistics under progressively Type-II-censored competing-risk data and under progressive stress models. Similar extensions have been made to Weibull Geometric models, modified Weibull distributions, and record-based sampling schemes.

More recently, Bayesian prediction for Weibull current records has been investigated, with applications to industrial reliability data; see Ahmadi and Doostparast [6], Aldallal [14], Al-Hussaini et al. [15, 16], Dey et al. [43], Jaheen and Ali [51].

A typical Bayesian prediction interval is obtained from posterior predictive quantiles satisfying

$$P(L_\alpha \le Y_{s:m} \le U_\alpha \mid \mathbf{x}) = 1 - \alpha.$$

The main strength of Bayesian prediction is its ability to combine prior information with limited failure data. However, posterior computation often requires simulation-based methods, especially in complex censoring scenarios.

5.2. Predictive distributions under censoring

An important development over the last decade has been the direct construction of predictive distributions under censoring. Rather than producing only point predictions or interval estimates, these methods derive the complete conditional distribution of future observations.

Let

$$X_{1:n} \leq \cdots \leq X_{r:n}$$

denote the observed failures under Type-II censoring. The predictive distribution of a future order statistic $X_{s:n}$, with $s > r$, is often expressed as

$$P(X_{s:n} \leq y \mid X_{1:n}, \ldots, X_{r:n}).$$

Recent studies have investigated predictive distributions for progressive censoring, competing risks, and generalized lifetime distributions. Exact predictive distributions are available for some parametric families, while numerical and simulation-based procedures are often required in more complicated settings; see Al-Hussaini et al. [15, 16], Asgharzadeh et al. [20], Ahmed [62].

The key advantage of predictive distribution methods is that they provide complete quantification of uncertainty. Point predictors, prediction intervals, and risk measures can all be derived directly from the predictive distribution.

5.3. Progressive censoring schemes

Progressive censoring has become one of the most widely studied censoring mechanisms in reliability experiments. In progressively Type-II censored tests, surviving units may be removed during the experiment according to a predetermined censoring plan.

Recent research has focused on prediction procedures for progressively censored order statistics, generalized progressive hybrid censoring schemes, and competing-risk settings. Exact prediction intervals, Bayesian prediction methods, and inferential procedures for progressively censored samples have been developed for many lifetime models; see Abu-Moussa et al. [1], Al-Hussaini et al. [15, 16], Balakrishnan and Aggarwala [21], Ahmed [62].

A typical prediction problem involves estimating a future censored order statistic

$$Y_{s:m:R},$$

where R denotes the progressive censoring scheme. The primary challenge arises from the additional randomness introduced by progressive removals. Nevertheless, progressive censoring more closely reflects practical life testing experiments than ordinary Type-II censoring, which explains its growing importance in modern reliability studies.

5.4. Prediction based on record values

Record statistics continue to play an important role in prediction because records arise naturally in meteorology, sports statistics, hydrology, and reliability studies.

Recent work has examined prediction based on upper records, current records, and k-record values. Bayesian prediction procedures and exact prediction intervals have been proposed for a variety of lifetime distributions, including Weibull models and their extensions; see Ahmadi and Doostparast [6], Aldallal [14], Bdair and Raqab [37].

Suppose

$$R_1 < R_2 < \cdots < R_n$$

denotes a sequence of upper record values. The prediction target is often a future record value R_{n+k}. Record-based prediction is particularly attractive because useful inference can be performed even when only record-breaking observations are available.

5.5. *Nonparametric prediction intervals*

Nonparametric predictive inference, conformal prediction, and model-free predictive procedures provide alternatives to a fully specified parametric family, with validity subject to their respective assumptions on exchangeability, sampling, or resampling; see Coolen [40], Vovk et al. [66], Wang and Politis [67].

The primary objective is to construct intervals

$$[L(X), U(X)]$$

satisfying

$$P(Y_{\text{new}} \in [L(X), U(X)]) \approx 1 - \alpha.$$

Unlike parametric intervals, these procedures remain valid under broad model classes, making them attractive when the underlying distribution is difficult to specify.

5.6. *Bootstrap-based predictive procedures*

Bootstrap prediction has grown substantially due to increased computational power. The bootstrap is particularly useful when exact predictive distributions are unavailable. Politis's transformation-based framework develops model-free point and interval prediction, as described by Politis [63], while Wang and Politis [67] compares model-free bootstrap and conformal prediction in regression. These are broader methodological contexts, not procedures evaluated in the present order statistic benchmark.

The basic predictive bootstrap algorithm repeatedly generates pseudo-samples

$$X_1^*, \ldots, X_n^*,$$

and computes future predictions for each resample. Prediction intervals are then obtained from empirical predictive quantiles. A particularly important development is the predictive bootstrap framework, which explicitly targets future observations rather than parameter estimation. Such procedures are especially useful in reliability applications where analytical predictive distributions are unavailable.

5.7. *Recent developments in reliability prediction*

Foundational and recent sources cover complementary aspects of uncertainty quantification, progressive censoring, competing risks, record-based prediction, and nonparametric predictive inference; see Abu-Moussa et al. [1], Aldallal [14], Balakrishnan and Aggarwala [21], Coolen [40].

An important trend is the shift from reporting only point predictions toward providing complete predictive distributions and interval estimates. This shift allows practitioners to quantify predictive uncertainty and make more informed reliability decisions. Looking ahead, future research is expected to focus on more flexible predictive models, increasingly complex censoring structures, and scalable computational algorithms suitable for large reliability databases.

The modern methods discussed in this section broaden the literature review beyond the numerical benchmark. The simulation below is deliberately narrower: it compares six selected procedures under common fixed-size Type-II censoring schemes and future order-statistic targets. Point predictors are evaluated using prediction error measures, while interval procedures are assessed by coverage and width. Bayesian, progressive censoring, record-based, nonparametric/conformal, and bootstrap procedures remain part of the review and taxonomy but are not included in this benchmark. As represented in the cited literature, they may involve different observation schemes, predictive targets, assumptions, or calibration choices; the present design does not provide a common implementation protocol for a defensible numerical comparison. Their omission reflects scope, not inferior performance; a broader benchmark is left for future work.

6. Simulation Study

6.1. Simulation design

To compare the predictive performance of the competing procedures under controlled conditions, a Monte Carlo simulation study was conducted. Within each replication, all methods use the same observed order statistics and are evaluated against the same realized future target, giving a paired comparison. In each Monte Carlo replication, a complete sample of size n is generated from a known lifetime distribution and then sorted to obtain

$$X_{1:n} < X_{2:n} < \cdots < X_{n:n}.$$

Only the first r order statistics are treated as observed Type-II censored data, and the task is to predict a later order statistic $X_{s:n}$, where $r < s \leq n$.

The simulation used $B = 250$ Monte Carlo replications in the known-model setting. This setting is common in methodological simulation tables because it separates the quality of the prediction formula from the additional error caused by parameter estimation. Four lifetime models were considered: exponential (rate 1), Weibull (shape 1.6, scale 1), gamma (shape 2, rate 1), and a two component exponential mixture with weights $0.5, 0.5$ and rates $2, 4$. The R code uses the default random seed `20260610`. For each model, two censoring designs were used:

$$(n, r, s) = (30, 15, 24) \quad \text{and} \quad (n, r, s) = (50, 35, 45).$$

The first design observes 50% of the ordered sample (50% censored) and predicts a later unobserved order statistic. The second design observes 70% of the sample (30% censored) and predicts a high future order statistic. Together, these designs test both moderate and heavier extrapolation beyond the observed failures. Here, n is fixed at 30 or 50 in every replication; its appearance in the (U)-, (V)-, and conditional expectation formulas does not make the sample size random. The simulation does not draw N from a count distribution, so random-sample-size prediction is reviewed theoretically but remains outside this fixed-size benchmark. The benchmark is correctly specified: the generating distributional family matches the family assumed by model-based procedures, while each method retains its prescribed parameter treatment, including any estimation step. We do not deliberately fit a wrong distributional family; thus, this simulation compares methods under model adequacy and does not assess robustness to model misspecification. Exactness of the pivotal interval construction is conditional on a fully specified F. Replacing F by an estimate $\widehat{F}$ gives a plug-in procedure; its finite sample coverage is not thereby guaranteed to remain exact. All prediction intervals in the simulation use a nominal 95% prediction level ($1 - \alpha = 0.95$, $\alpha = 0.05$). The coverage values reported below are empirical coverage proportions over the $B = 250$ Monte Carlo replications.

6.2. Compared Prediction Methods

The numerical comparison uses six procedures or implementation variants based on published methods and legacy R code, evaluated under the same fixed size Type-II censoring designs and future order statistic targets. The ratio third summary is identified separately below as a legacy code-based variant. This is not a benchmark of every method family discussed in the review.

1. U-pivotal S1 midpoint: the fixed sample interval based on U, with its midpoint used as the point prediction. The pivotal construction is due to Barakat et al. [27]; the midpoint summary is Method M1 of Khaled et al. [54].
2. V-pivotal harmonic: the fixed sample interval based on V, with the harmonic mean of its endpoints used as the point prediction. The pivotal construction is due to Barakat et al. [27]; the harmonic summary is Method M2 of Khaled et al. [54].
3. V-pivotal ratio third: the same V-pivotal interval with the legacy R code endpoint summary defined below. We identify it as an implementation variant, not as a separate pivotal statistic.
4. Near-MMSE exp-CE: the conditional expectation predictor based on exponential spacings from Barakat et al. [34].

5. Near-MMSE uniform-CE: the conditional expectation predictor based on the uniform order statistic scale from Barakat et al. [34].
6. Hazard MWLS-PWLS: the cumulative hazard modified weighted least squares estimator and predictive weighted least squares predictor related to Aly et al. [18] and the comparative workflow of Harpy et al. [49].

To state the point summaries exactly, let $[L_U, U_U]$ and $[L_V, U_V]$ denote the intervals in Eqs. (7) and (8), respectively. The two summaries corresponding to Methods M1 and M2 of Khaled et al. [54] are

$$\widehat{X}_{s:n}^{U,\mathrm{mid}} = \frac{L_U + U_U}{2}, \tag{17}$$

$$\widehat{X}_{s:n}^{V,\mathrm{harm}} = \frac{2L_V U_V}{L_V + U_V}. \tag{18}$$

In the R implementation, the harmonic rule is used when both endpoints are positive; otherwise, the code falls back to their midpoint. The endpoints in the lifetime applications considered here are positive. The ratio third rule used in `R code` is the endpoint formula implemented in the earlier `statistics v 2023.R` code:

$$\widehat{X}_{s:n}^{V,\mathrm{ratio}} = L_V + \frac{L_V}{3}\left(1 - \frac{L_V}{U_V}\right). \tag{19}$$

This is not the one third interpolation $L_V + (U_V - L_V)/3$, nor is it the geometric mean Method M3 of Khaled et al. [54]. The citation to Barakat et al. [27] supports the underlying U- and V-pivotal interval constructions; Khaled et al. [54] supports the midpoint and harmonic point summaries. The ratio third expression is reported transparently as a legacy code-based summary rather than attributed to either published method.

The first three methods produce prediction intervals and therefore allow comparisons of coverage and interval width. The near-MMSE and hazard-based methods are treated here as point predictors, so their main comparison is through prediction error.

Numerical implementation of MWLS-PWLS The benchmark uses a two-step implementation, rather than joint minimization over Θ and x_s: it first minimizes (15) to obtain $\widehat{\Theta}$, then minimizes (16) over $x_s > x_r$ with $\widehat{\Theta}$ fixed. Parameter fitting uses R's `optim` with `Nelder-Mead` on log-transformed parameters, ensuring positivity, and a maximum of 800 iterations. Let m and v be the observed prefix mean and sample variance, and $\varepsilon = 10^{-10}$. The single starting point is rate $r/\max\{\sum_{i=1}^{r} x_i + (n-r)x_r, \varepsilon\}$ for the exponential model; shape 1.4 and scale $\max(m, \varepsilon)$ for Weibull; and shape $k_0 = \max(m^2/v, 0.2)$, rate $\max\{k_0/\max(m, \varepsilon), 0.2\}$ for gamma, with $k_0 = 1$ if v is nonpositive or nonfinite.

Prediction uses `optimise` on $[x_r + \varepsilon, \max\{3c, 2x_r, x_r + 1\}]$, where $c = \widehat{F}^{-1}(1 - \exp\{-\mu_{s:n}^*\})$. The code retains the initialized model after a fitting error or nonfinite fitted objective, and uses $\max(c, x_r)$ after a prediction optimization error. These are implementation settings; no systematic assessment of convergence or sensitivity to alternative starting values was conducted.

6.3. Performance measures

For each replication $b = 1, \ldots, B$, let $\widehat{X}_{s:n}^{(b)}$ be the predicted value and $X_{s:n}^{(b)}$ the true simulated future order statistic. Point prediction accuracy is summarized by

$$\mathrm{MSE}(\widehat{X}_{s:n}) = \frac{1}{B}\sum_{b=1}^{B}\left(\widehat{X}_{s:n}^{(b)} - X_{s:n}^{(b)}\right)^2, \tag{20}$$

$$\mathrm{RMSE}(\widehat{X}_{s:n}) = \sqrt{\mathrm{MSE}(\widehat{X}_{s:n})}, \tag{21}$$

and

$$\text{Bias}(\widehat{X}_{s:n}) = \frac{1}{B}\sum_{b=1}^{B}\left(\widehat{X}_{s:n}^{(b)} - X_{s:n}^{(b)}\right). \tag{22}$$

For interval prediction, if $L_{s:n}^{(b)}$ and $U_{s:n}^{(b)}$ are the lower and upper prediction limits, empirical coverage and average width are computed as

$$\widehat{\text{CP}} = \frac{1}{B}\sum_{b=1}^{B}\mathbf{1}\left\{L_{s:n}^{(b)} \le X_{s:n}^{(b)} \le U_{s:n}^{(b)}\right\}, \tag{23}$$

Here the nominal level is 95% ($\alpha = 0.05$); $\widehat{\text{CP}}$ is the resulting empirical coverage, not the nominal level itself. and

$$\widehat{\text{AW}} = \frac{1}{B}\sum_{b=1}^{B}\left(U_{s:n}^{(b)} - L_{s:n}^{(b)}\right). \tag{24}$$

The best method in each scenario is selected based on the smallest RMSE, since the primary simulation objective is point prediction. Coverage and width are interpreted separately for the interval producing methods.

6.4. Best method by scenario

Table 5. Best performing published method in each simulation scenario, selected by RMSE.

Scenario	Observed data	Best method	RMSE	Bias
Exponential	50% observed	Near-MMSE exp-CE	0.3068	-0.0198
Exponential	70% observed	Near-MMSE exp-CE	0.3517	-0.0325
Gamma	50% observed	Near-MMSE uniform-CE	0.4092	0.0121
Gamma	70% observed	Near-MMSE exp-CE	0.4054	0.0220
Mixture exponential	50% observed	Near-MMSE uniform-CE	0.1185	0.0019
Mixture exponential	70% observed	Near-MMSE exp-CE	0.1372	0.0087
Weibull	50% observed	Near-MMSE uniform-CE	0.1533	-0.0025
Weibull	70% observed	Near-MMSE exp-CE	0.1608	-0.0061

Table 5 gives the best performing method in each scenario. The near-MMSE predictors achieve the smallest RMSE in all eight scenario design combinations. The exponential-spacing version is best in five settings: both exponential cases, the 70% observed gamma, mixture exponential, and Weibull cases. The uniform-scale version is best in three settings: the 50% observed-gamma, mixture-exponential, and Weibull cases.

These are empirical rankings for the examined settings. Although the exact conditional mean minimizes expected squared error loss when second moments exist, the near-MMSE rules back transform conditional means from other scales and need not equal that exact predictor. Conditional mean optimality, therefore, motivates these constructions but does not prove their observed ranking or universal superiority; see Sections 4.1 and 6.7.

6.5. Full method level results

Figures 1-2 give a visual summary of the full method level results, while Tables 6-9 report the numerical values by distributional scenario.

Across the scenario-design combinations available for each method, the average RMSEs are 0.2564 for near-MMSE exp-CE, 0.2579 for near-MMSE uniform-CE, 0.4260 for hazard MWLS-PWLS, 0.4619 for V pivotal harmonic, 0.5543 for U pivotal S1 midpoint, and 0.6450 for V pivotal ratio-third. Hazard MWLS-PWLS is evaluated in six combinations; its mixture exponential implementation is not included. Thus, these averages describe the evaluated settings for each method and do not imply identical scenario coverage for all six procedures.

The V pivotal harmonic summary has the smallest RMSE among the pivotal point summaries in these settings. The harmonic and ratio third rules use the same V-pivotal interval, so their empirical coverage and average width

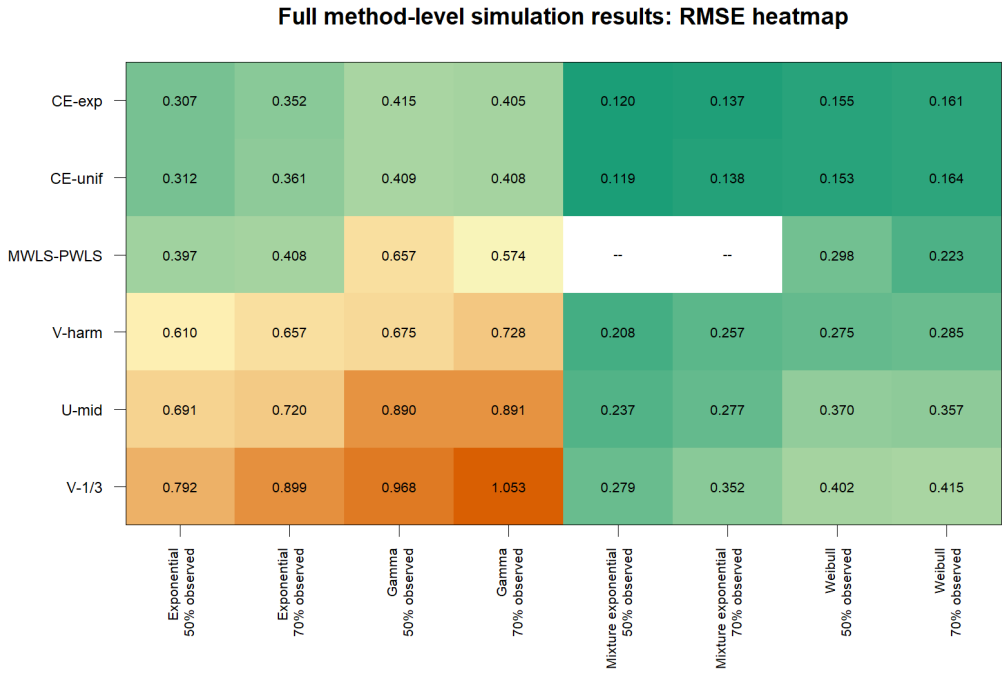

Figure 1. Heatmap version of the full method level simulation results. Darker green cells indicate smaller RMSE values; orange cells indicate larger RMSE values.

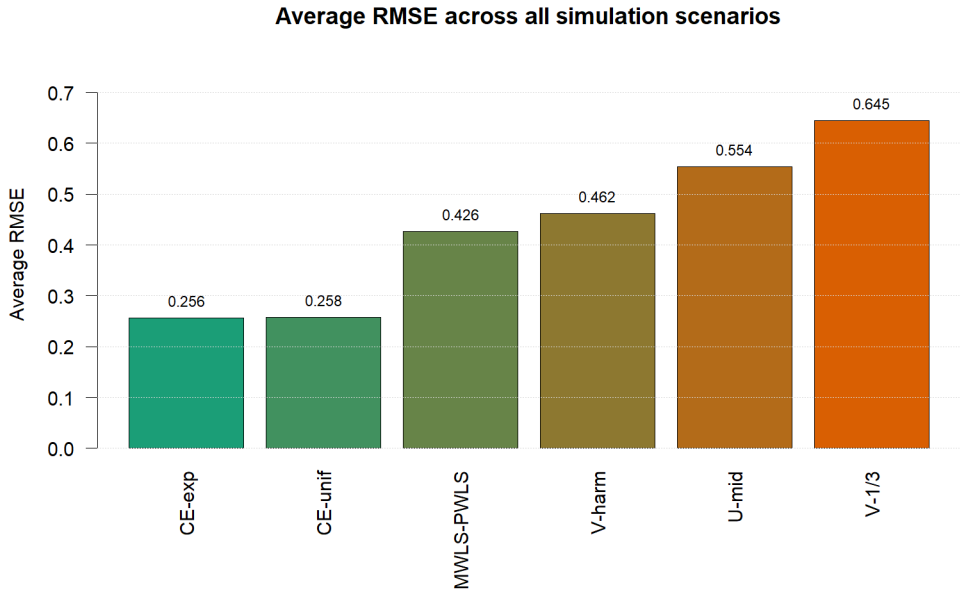

Figure 2. Average RMSE by method across all simulation scenario design combinations.

Table 6. Full method level simulation results for the exponential model.

Observed data	Method	RMSE	Bias	Coverage	Avg. width
50% observed	Near-MMSE exp-CE	0.3068	-0.0198	–	–
50% observed	Near-MMSE uniform-CE	0.3123	-0.0613	–	–
50% observed	Hazard MWLS-PWLS	0.3970	0.0095	–	–
50% observed	V pivotal harmonic	0.6101	-0.5132	0.936	1.7055
50% observed	U pivotal S1 midpoint	0.6913	-0.6118	0.168	0.5525
50% observed	V pivotal ratio-third	0.7924	-0.7278	0.936	1.7055
70% observed	Near-MMSE exp-CE	0.3517	-0.0325	–	–
70% observed	Near-MMSE uniform-CE	0.3607	-0.0866	–	–
70% observed	Hazard MWLS-PWLS	0.4076	0.0004	–	–
70% observed	V pivotal harmonic	0.6574	-0.5503	0.948	1.8231
70% observed	U pivotal S1 midpoint	0.7205	-0.6256	0.352	0.8837
70% observed	V pivotal ratio-third	0.8985	-0.8267	0.948	1.8231

Table 7. Full method level simulation results for the gamma model.

Observed data	Method	RMSE	Bias	Coverage	Avg. width
50% observed	Near-MMSE uniform-CE	0.4092	0.0121	–	–
50% observed	Near-MMSE exp-CE	0.4146	0.0682	–	–
50% observed	Hazard MWLS-PWLS	0.6566	0.1291	–	–
50% observed	V pivotal harmonic	0.6751	-0.5056	0.956	2.3638
50% observed	U pivotal S1 midpoint	0.8899	-0.7779	0.208	0.8307
50% observed	V pivotal ratio-third	0.9679	-0.8699	0.956	2.3638
70% observed	Near-MMSE exp-CE	0.4054	0.0220	–	–
70% observed	Near-MMSE uniform-CE	0.4076	-0.0466	–	–
70% observed	Hazard MWLS-PWLS	0.5738	0.0633	–	–
70% observed	V pivotal harmonic	0.7280	-0.5752	0.940	2.3185
70% observed	U pivotal S1 midpoint	0.8914	-0.7753	0.340	1.1612
70% observed	V pivotal ratio-third	1.0534	-0.9635	0.940	2.3185

Table 8. Full method level simulation results for the mixture exponential model.

Observed data	Method	RMSE	Bias	Coverage	Avg. width
50% observed	Near-MMSE uniform-CE	0.1185	0.0019	–	–
50% observed	Near-MMSE exp-CE	0.1199	0.0186	–	–
50% observed	V pivotal harmonic	0.2076	-0.1706	0.984	0.7074
50% observed	U pivotal S1 midpoint	0.2373	-0.2062	0.200	0.2132
50% observed	V pivotal ratio-third	0.2793	-0.2534	0.984	0.7074
70% observed	Near-MMSE exp-CE	0.1372	0.0087	–	–
70% observed	Near-MMSE uniform-CE	0.1377	-0.0146	–	–
70% observed	V pivotal harmonic	0.2565	-0.2144	0.952	0.7630
70% observed	U pivotal S1 midpoint	0.2769	-0.2384	0.368	0.3510
70% observed	V pivotal ratio-third	0.3517	-0.3238	0.952	0.7630

are identical under the paired comparison; coverage ranges from 0.936 to 0.984 in this run. Their point predictions differ, with the ratio third rule giving a larger RMSE and more negative bias. The U pivotal S1 midpoint method produces narrower intervals, but its empirical coverage is too low in these simulations, especially when only 50% of the sample is observed. A narrow interval is not valuable if it misses the future order statistic too often.

The hazard MWLS-PWLS method is competitive but not dominant in this known model run. It performs better than the pivotal point predictors in several settings, especially for the exponential and Weibull cases, but it does not outperform the near-MMSE predictors. Its value is nevertheless important in practice because it links parameter estimation and prediction through a cumulative-hazard least-squares framework. In real applications, where the model parameters are unknown and may be estimated from censored data, this framework can be attractive even if it is not the RMSE winner in the ideal, known-model simulation.

Table 9. Full method level simulation results for the Weibull model.

Observed data	Method	RMSE	Bias	Coverage	Avg. width
50% observed	Near-MMSE uniform-CE	0.1533	-0.0025	–	–
50% observed	Near-MMSE exp-CE	0.1546	0.0197	–	–
50% observed	V pivotal harmonic	0.2749	-0.2163	0.968	0.9464
50% observed	Hazard MWLS-PWLS	0.2980	0.0745	–	–
50% observed	U pivotal S1 midpoint	0.3698	-0.3322	0.180	0.3584
50% observed	V pivotal ratio-third	0.4019	-0.3685	0.968	0.9464
70% observed	Near-MMSE exp-CE	0.1608	-0.0061	–	–
70% observed	Near-MMSE uniform-CE	0.1636	-0.0314	–	–
70% observed	Hazard MWLS-PWLS	0.2231	0.0102	–	–
70% observed	V pivotal harmonic	0.2853	-0.2323	0.948	0.8690
70% observed	U pivotal S1 midpoint	0.3569	-0.3164	0.308	0.4548
70% observed	V pivotal ratio-third	0.4152	-0.3820	0.948	0.8690

6.6. Statistical interpretation

For point prediction of $X_{s:n}$ under squared error loss, the near-MMSE methods have the lowest RMSE in the scenarios studied here. The exponential spacing version leads in both exponential designs. For gamma, Weibull, and mixture exponential data, the uniform-scale version leads in the 50% observed design, whereas the exponential-spacing version leads in the 70% observed design. These differences are specific to the present Monte Carlo run and do not establish a general preference for one variant across a distributional family.

Bias helps interpret RMSE differences without identifying their causes. For gamma data with 50% observed, the U-midpoint bias is -0.7779, compared with 0.0121 for uniform near-MMSE (Table 7), indicating systematic underprediction by the former. This pattern alone cannot distinguish the effects of skewness, tails, or the prediction rule. The pivotal and near-MMSE procedures use supplied parameters, so their differences are not due to parameter estimation; MWLS-PWLS estimates parameters. The designs vary n, r, s together and do not isolate individual effects. This four-model benchmark is not a systematic parameter sensitivity or misspecification study; broader robustness assessment remains future work.

If the goal is interval prediction with reliable empirical coverage, the pivotal methods remain relevant. Among them, the V pivotal harmonic method gives the most balanced behavior: high coverage with smaller RMSE than the V ratio third implementation. However, the practitioner pays for coverage through wider intervals. The U pivotal midpoint implementation is not recommended as the primary interval procedure in these settings because its coverage is substantially below the nominal target.

Therefore, the simulation supports a clear practical recommendation. Use near MMSE conditional expectation when the paper's objective is point prediction and the model F is trusted or well estimated. Use V-based pivotal intervals when the objective is uncertainty quantification and empirical coverage is more important than point accuracy. Treat hazard-based least squares as a flexible bridge between estimation and prediction, especially for censored or generalized ordered data where a cumulative hazard representation is natural.

6.7. Interpretation and limitations of the comparative study

The comparative results reported in Table 5 and Tables 6-9 should be interpreted in light of the design of the simulation experiment. In particular, the simulation scenarios were generated from fully specified parametric models and evaluated primarily through squared error criteria such as MSE and RMSE. The observed near-MMSE rankings are empirical findings under these design choices, rather than consequences of an exact MMSE guarantee for the implemented approximations.

The near-MMSE predictors are derived from conditional expectation principles. The exact conditional mean minimizes squared error loss when second moments exist, but the predictors in Equations (9)-(10) back transform conditional means on transformed scales and need not equal that exact mean (Section 4.1). Their favorable RMSEs provide empirical evidence for the studied scenarios, not proof of exact MMSE optimality or universal dominance.

It is also important to recognize that different practical objectives may lead to different conclusions. The present study focuses primarily on point prediction accuracy. However, many reliability and survival applications require interval prediction and uncertainty quantification rather than solely minimizing prediction error. In such settings, empirical coverage probability and interval width become equally important performance measures. The pivotal procedures considered in this review remain valuable because they directly provide prediction intervals and often achieve satisfactory coverage properties, even when their point predictions are less accurate than those of conditional expectation methods.

Another limitation is that the simulation assumes the underlying lifetime distribution is correctly specified. In real applications, the true distribution is unknown and must be estimated from the available data. Model misspecification may affect different prediction procedures in different ways. Methods that perform well under ideal model assumptions may lose part of their advantage when the fitted model does not accurately represent the underlying failure mechanism. Consequently, the relative ranking of prediction procedures may differ under distributional misspecification, heterogeneous populations, or complex censoring structures.

The pivotal and near-MMSE procedures use known generating parameters, whereas MWLS-PWLS includes parameter estimation from the observed prefix. The benchmark, therefore, does not quantify estimation uncertainty across all procedures on a common fitted-parameter basis. This issue becomes particularly important for small samples, heavily censored experiments, and multi parameter lifetime models. Future studies may therefore benefit from incorporating bootstrap-based or Bayesian uncertainty assessments in order to evaluate the combined effects of parameter estimation and future prediction.

Finally, the conclusions reported here are specific to the distributions, censoring schemes, and prediction targets considered in the simulation design. Different results may arise with alternative lifetime models, more severe censoring levels, far-tail prediction targets, or other loss functions. For these reasons, the present findings should be interpreted as comparative evidence within a common benchmark framework rather than as proof that a single prediction method is uniformly superior in all practical situations.

7. Applications

The reviewed methods are most naturally applied in reliability and life testing experiments. If a test begins with n components but is stopped after the first r failures, predicting $X_{s:n}$ helps decide when future failures are likely to occur and whether testing should continue. Pivotal methods are especially suitable when the fitted lifetime distribution passes goodness-of-fit checks.

In reliability studies, the future order statistic has a direct engineering interpretation. For example, $X_{r+1:n}$ may represent the next expected failure after a censored experiment, while $X_{n:n}$ may represent the time by which all units are expected to have failed. The first target is useful for maintenance scheduling and warranty planning; the second is for assessing the complete service life profile of a batch. This distinction explains why both short-horizon and long-horizon predictions must be reported. A method that accurately predicts the next failure may still be overly optimistic about the maximum future failure time.

Industrial examples also show why the fitted distribution matters. Weibull-type models are common in failure analysis because they can represent increasing, decreasing, and approximately constant hazard rates. Modified lifetime distributions, including modified Kies exponential and modified Muth models, provide additional flexibility when the empirical failure pattern does not follow a simple Weibull shape. Mixture models become relevant when the observed data may come from different subpopulations, such as materials produced under slightly different manufacturing conditions or patients receiving different treatment paths.

Record-value applications provide another way to understand the practical importance of the same problem. In many real systems, analysts do not keep the full data stream; they retain only record highs, record lows, failure extremes, or current ranges. This occurs in industrial monitoring, climate records, sports performance, and financial risk tracking; see Aldallal and Barakat [9], Brock et al. [38], Glick [48]. Although the sampling scheme differs from ordinary order statistics, the inferential question is similar: what can be said about a future extreme or future ordered event from a limited ordered history?

Manufacturing and quality-control problems motivate random-sample-size models. In many production processes, the number of usable units observed in advance is not fixed. Modeling N as binomial, Poisson, or negative binomial can be more realistic than pretending that the sample size is fixed. The reviewed papers of varying sizes show that this choice affects both coverage and interval width.

Medical and biological settings often involve censoring, incomplete follow-up, and heterogeneous populations. Mixture gamma and beta distributions are relevant when recovery rates, survival times, or event risks are not well described by one component, as discussed by Khaled et al. [54]. Near-MMSE conditional expectation prediction is useful when the priority is a stable point prediction for future ordered observations under Type-II censoring, as discussed by Barakat et al. [34]. Cumulative hazard least squares is relevant for censored survival type data and generalized ordered statistics, as discussed by Aly et al. [18]. Sequential prediction methods may be useful when practitioners want a predicted trajectory rather than a single future observation, as discussed by Harpy et al. [49].

The windscreen failure data application below is included for this reason. It is small enough to audit by hand, but rich enough to expose the behavior of several prediction rules. The observed data are sorted failure measurements, the prediction starts after a fixed observed order statistic, and each method is asked to predict the same future targets. This makes the application a useful bridge between the theoretical review and the simulation study: the same family of methods is compared, but errors are now computed against real, ordered observations rather than the simulated truth. For both real data applications, the pivotal prediction intervals use the same nominal 95% prediction level ($\alpha = 0.05$) as in the simulation.

7.1. Aircraft windscreen failure data

The aircraft windscreen failure data contain 84 ordered failure times. The first $r = 34$ ordered observations are treated as the observed Type-II censored sample, and the remaining future order statistics $s = 35, \ldots, 84$ are predicted. The analysis uses the modified Kies-exponential (MKE) distribution.

$$G(x; a, b) = 1 - \exp\left[-\{\exp(bx) - 1\}^a\right], \qquad a = 1.783644, \quad b = 0.2366933.$$

The comparison uses the same representative methods as the simulation section: the U-pivotal midpoint rule, two V-pivotal point rules, the two near-MMSE conditional expectation predictors, and the cumulative hazard MWLS-PWLS method fitted from the first observed failures.

Table 10. Real data comparison for the aircraft windscreen failure data under the MKE model.

Method	Bias	MAE	MSE	RMSE	CV	Coverage	Avg. width
Near-MMSE uniform-CE	-0.0542	0.1123	0.0229	0.1512	0.2235	–	–
Near-MMSE exp-CE	-0.0333	0.1105	0.0244	0.1561	0.2301	–	–
Hazard MWLS-PWLS fitted	-0.3700	0.3804	0.2227	0.4719	0.1754	–	–
V pivotal harmonic	-0.6414	0.6573	0.7054	0.8399	0.0909	1.000	1.4290
V pivotal ratio-third	-0.8804	0.8827	1.2287	1.1085	0.0445	1.000	1.4290
U pivotal S1 midpoint	-0.9135	0.9536	1.4114	1.1880	0.0464	0.220	0.4538

Table 11. Point predictions for the aircraft windscreen data under the MKE model: $s = 35, \ldots, 59$.

s	Actual	U-mid	V-harm	V-1/3	CE-exp	CE-unif	MWLS-PWLS
35	2.154	2.372	2.191	2.171	2.174	2.174	2.167
36	2.190	2.272	2.221	2.190	2.213	2.212	2.195
37	2.194	2.497	2.248	2.207	2.252	2.251	2.222
38	2.223	2.201	2.272	2.221	2.290	2.289	2.249
39	2.224	2.281	2.295	2.234	2.328	2.326	2.276
40	2.229	2.469	2.316	2.246	2.366	2.364	2.303
41	2.300	2.402	2.336	2.258	2.404	2.402	2.329
42	2.324	2.279	2.356	2.269	2.442	2.439	2.356
43	2.385	2.164	2.375	2.279	2.480	2.477	2.382
44	2.481	2.295	2.393	2.288	2.518	2.514	2.409
45	2.610	2.440	2.411	2.298	2.556	2.552	2.435
46	2.625	2.158	2.428	2.307	2.594	2.589	2.462
47	2.632	2.397	2.445	2.315	2.632	2.627	2.488
48	2.646	2.135	2.462	2.324	2.670	2.665	2.515
49	2.661	2.350	2.478	2.332	2.708	2.702	2.541
50	2.688	2.484	2.494	2.340	2.746	2.740	2.568
51	2.823	2.268	2.509	2.347	2.785	2.778	2.594
52	2.890	2.407	2.524	2.355	2.823	2.817	2.621
53	2.902	2.315	2.540	2.362	2.862	2.855	2.648
54	2.934	2.486	2.555	2.369	2.902	2.894	2.675
55	2.962	2.289	2.569	2.376	2.941	2.933	2.703
56	2.964	2.139	2.584	2.382	2.981	2.973	2.730
57	3.000	2.218	2.598	2.389	3.022	3.013	2.758
58	3.103	2.454	2.613	2.395	3.063	3.053	2.786
59	3.114	2.251	2.627	2.402	3.104	3.094	2.815

Table 12. Point predictions for the aircraft windscreen data under the MKE model: $s = 60, \ldots, 84$.

s	Actual	U-mid	V-harm	V-1/3	CE-exp	CE-unif	MWLS-PWLS
60	3.117	2.500	2.641	2.408	3.146	3.136	2.844
61	3.166	2.437	2.656	2.414	3.189	3.178	2.873
62	3.344	2.336	2.670	2.420	3.233	3.221	2.903
63	3.376	2.438	2.684	2.426	3.277	3.264	2.933
64	3.443	2.493	2.698	2.432	3.322	3.309	2.964
65	3.467	2.299	2.712	2.438	3.369	3.355	2.996
66	3.478	2.435	2.726	2.444	3.416	3.401	3.028
67	3.578	2.279	2.741	2.450	3.465	3.449	3.062
68	3.595	2.536	2.755	2.455	3.515	3.498	3.096
69	3.699	2.499	2.770	2.461	3.567	3.549	3.131
70	3.779	2.466	2.785	2.467	3.620	3.601	3.167
71	3.924	2.350	2.800	2.473	3.676	3.656	3.205
72	4.035	2.498	2.815	2.479	3.734	3.712	3.245
73	4.121	2.477	2.831	2.485	3.795	3.772	3.286
74	4.167	2.292	2.848	2.491	3.859	3.834	3.330
75	4.240	2.303	2.865	2.498	3.927	3.900	3.376
76	4.255	2.431	2.883	2.504	4.000	3.970	3.425
77	4.278	2.332	2.901	2.511	4.079	4.045	3.478
78	4.305	2.487	2.922	2.518	4.165	4.127	3.536
79	4.376	2.462	2.944	2.526	4.261	4.218	3.600
80	4.449	2.418	2.968	2.534	4.369	4.319	3.673
81	4.485	2.476	2.996	2.544	4.497	4.437	3.759
82	4.570	2.299	3.029	2.555	4.655	4.578	3.865
83	4.602	2.297	3.074	2.570	4.870	4.761	4.009
84	4.663	2.232	3.149	2.593	5.241	5.040	4.257

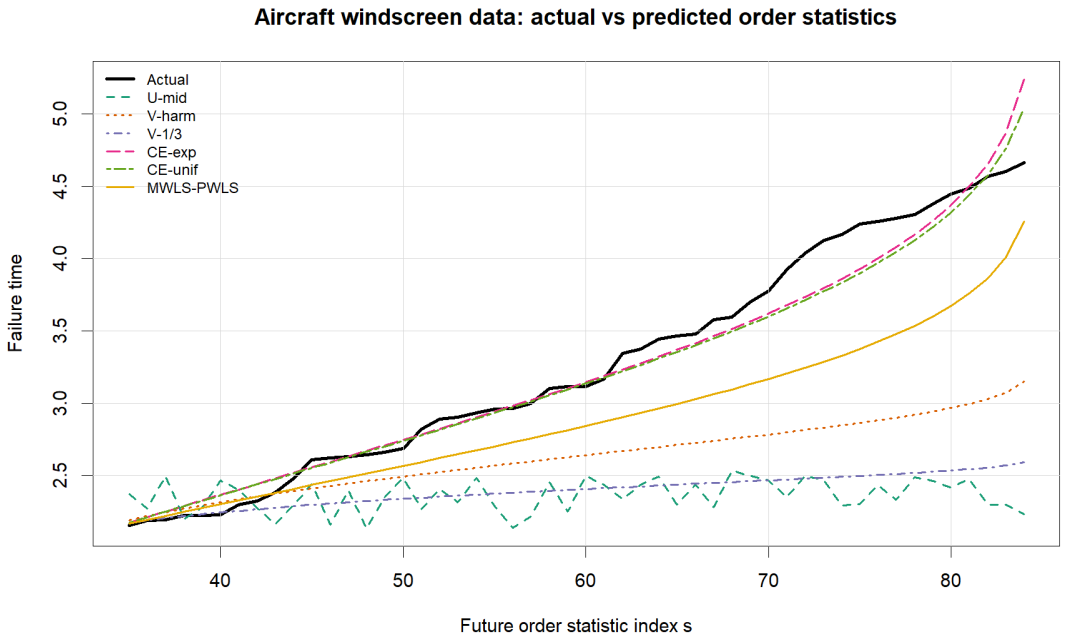

Figure 3. Actual aircraft windscreen failure order statistics and point predictions from the compared methods.

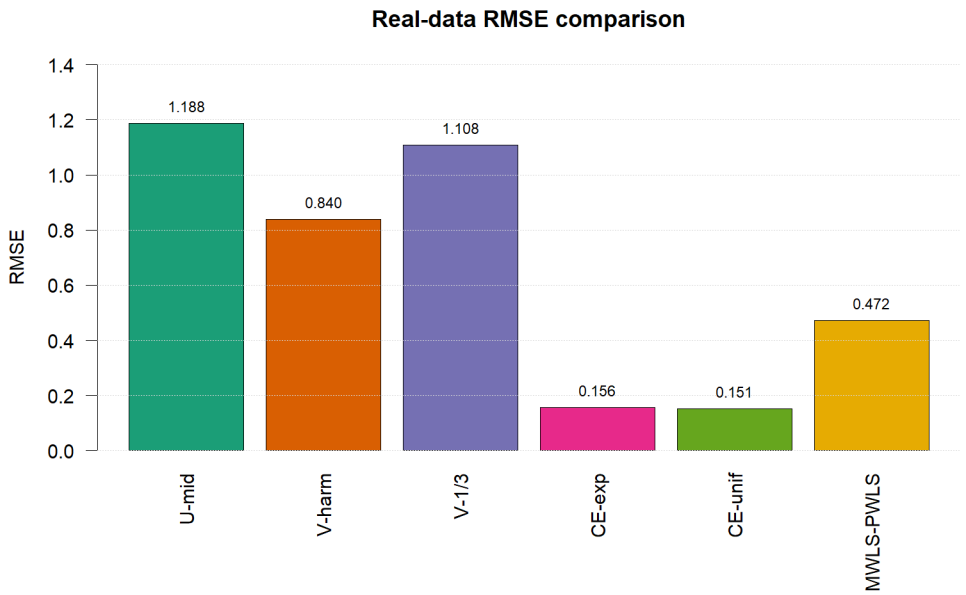

Figure 4. RMSE comparison for the aircraft windscreen real data application.

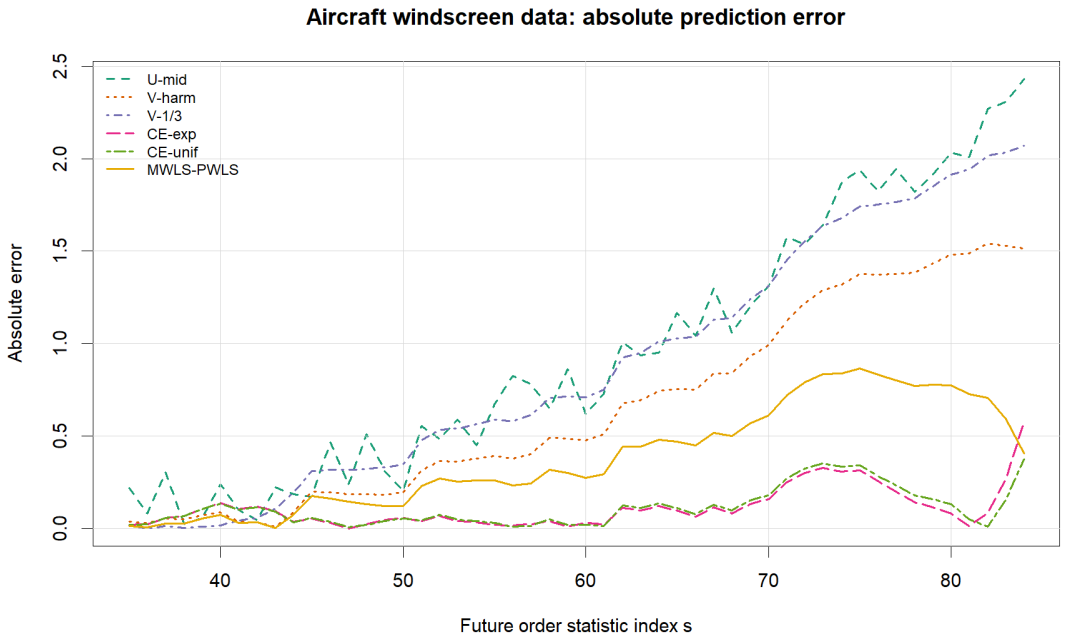

Figure 5. Absolute prediction error across future order statistic targets.

For this real data split, the smallest RMSE is obtained by Near-MMSE uniform-CE, with RMSE 0.1512 and MAE 0.1123. Figures 3-5 visualize the same comparison: the first figure shows how closely each method follows the observed order statistics, the second ranks the methods by RMSE, and the third shows where errors increase along the tail. The table also keeps the interval information for the pivotal methods; near-MMSE and MWLS-PWLS are treated as point predictors in the method comparison.

At the first unobserved rank $s = 35$, predictions are relatively close to the observation; near $s = 84$, the pivotal point summaries substantially underpredict, while near-MMSE remains closer than the pivotal summaries (Tables 11-12 and Figure 3). This describes target-specific behavior, not separate short and long-horizon RMSE estimates.

7.2. Voltage stress data

We consider the voltage stress failure data reported by El-Adll et al. [46]; the measurements were originally given by **?**]p. 189]lawlss. The experiment recorded the failure voltages, in kilovolts per millimeter, of 20 electrical cable insulation specimens subjected to increasing stress. The observations and the two-parameter Weibull model used in the source are shown below.

Table 13. Voltage stress failure data (kV/mm), reproduced from El-Adll et al. [46].

32.0	35.4	36.2	39.8	41.2	43.3	45.5	46.0	46.2	46.4
46.5	46.8	47.3	47.3	47.6	49.2	50.4	50.9	52.4	56.3

For a common 50% holdout comparison, we treat the first $r = 10$ ordered observations as observed and predict $X_{s:n}$ for $s = 11, \ldots, 20$. This split is the design of the present illustration; it differs from the several observed sample sizes examined in the source paper. We use

$$F(x; k, \sigma) = 1 - \exp\left[-(x/\sigma)^k\right], \qquad k = 9.1973, \quad \sigma = 47.7383,$$

using the reported Weibull fit. The pivotal and near-MMSE predictions use these supplied parameter values, whereas MWLS-PWLS estimates the Weibull parameters from the first ten observations; the resulting estimates are $\widehat{k} = 6.8045$ and $\widehat{\sigma} = 49.6587$. Thus, the latter also reflects parameter fitting from a substantially smaller observed subset.

Table 14 summarizes point prediction errors and, for pivotal procedures, the empirical coverage and average width of their prediction intervals on the ten held out observations. Table 15 reports every target specific prediction. These are descriptive results for one small dataset and one split, not evidence of a universally best procedure. In this split, Near-MMSE exp-CE has the lowest RMSE (1.1618), closely followed by Near-MMSE uniform-CE (1.1666); the fitted MWLS-PWLS predictor has the largest RMSE (3.4724).

Near $s = 11$, all predictions are relatively close to the observation; at $s = 20$, the pivotal summaries underpredict and fitted MWLS-PWLS overpredicts (Table 15 and Figure 6). Both near-MMSE predictions also underpredict the maximum but have smaller absolute errors than the pivotal summaries.

Table 14. Real data comparison for the voltage stress data under the Weibull model.

Method	Bias	MAE	MSE	RMSE	CV	Coverage	Avg. width
Near-MMSE exp-CE	0.8309	1.0841	1.3497	1.1618	0.0482	–	–
Near-MMSE uniform-CE	0.5659	1.0031	1.3611	1.1666	0.0436	–	–
U pivotal S1 midpoint	0.5752	1.2263	2.1538	1.4676	0.0349	1.000	7.2903
V pivotal harmonic	0.2568	1.2488	2.6076	1.6148	0.0299	1.000	7.2880
V pivotal ratio-third	-1.0262	1.6457	5.7057	2.3887	0.0177	1.000	7.2880
Hazard MWLS–PWLS fitted	3.3083	3.3083	12.0577	3.4724	0.0696	–	–

Table 15. Observed and predicted voltage stress order statistics for $s = 11, \ldots, 20$.

s	Actual	U-mid	V-harm	V-1/3	CE-exp	CE-unif	MWLS–PWLS
11	46.500	47.489	47.465	47.094	47.020	46.993	47.776
12	46.800	48.092	48.033	47.452	47.639	47.583	48.733
13	47.300	48.619	48.518	47.750	48.264	48.180	49.694
14	47.300	49.118	48.967	48.022	48.906	48.790	50.675
15	47.600	49.612	49.403	48.280	49.576	49.424	51.694
16	49.200	50.120	49.843	48.537	50.294	50.097	52.782
17	50.400	50.667	50.306	48.802	51.087	50.832	53.980
18	50.900	51.291	50.822	49.092	52.008	51.666	55.371
19	52.400	52.081	51.459	49.441	53.180	52.681	57.138
20	56.300	53.363	52.453	49.970	55.034	54.114	59.939

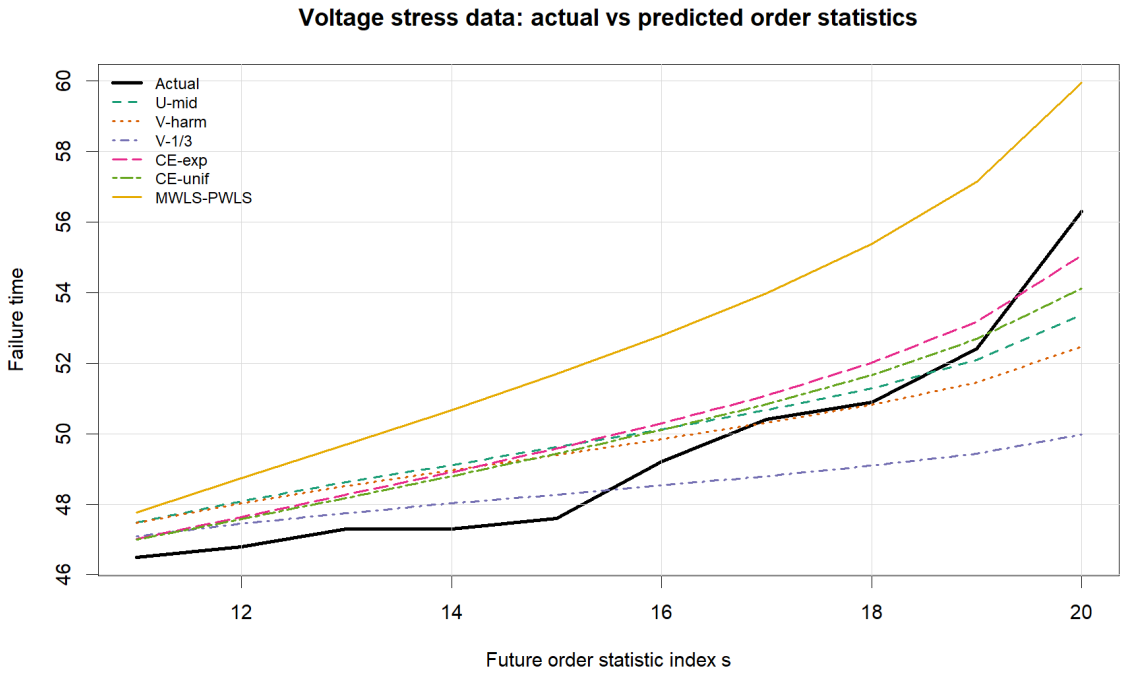

Figure 6. Observed voltage stress order statistics and point predictions from the compared methods.

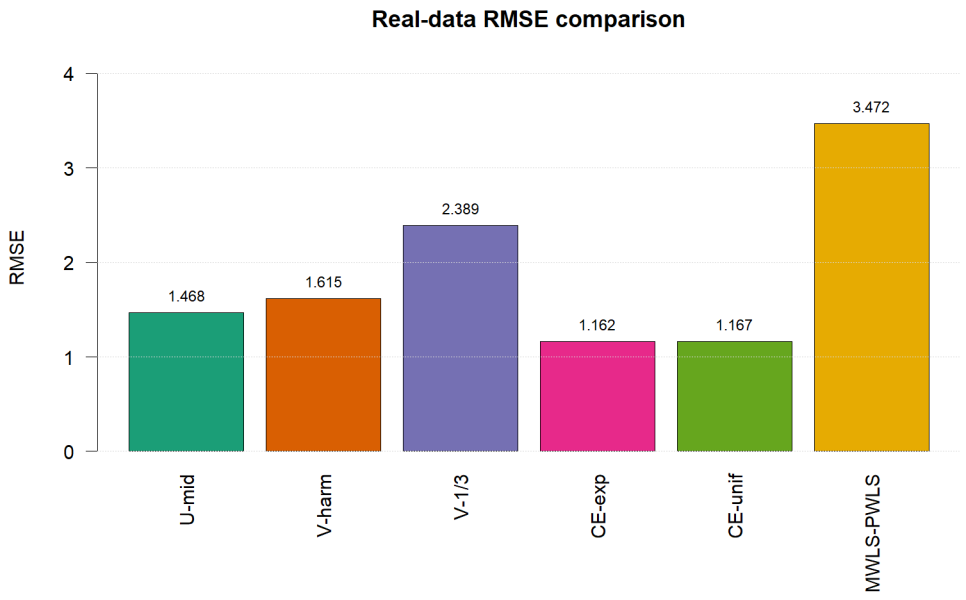

Figure 7. RMSE comparison for the voltage stress data application.

Both applications retain the favorable near-MMSE point prediction performance observed in the simulation, whereas fitted MWLS-PWLS is less competitive in the voltage example. The divergence at higher ranks cautions

against transferring complete method rankings across settings or using a single-point prediction as an operational failure threshold.

8. Limitations and Research Gaps

Three limitations recur throughout the reviewed literature. First, many methods assume that the fitted lifetime model is correct. When the model is wrong, the interval may have poor empirical coverage even if the mathematical derivation is exact under the assumed model. Second, mixture methods are flexible but can be unstable with small sample sizes, particularly when component distributions overlap. Third, random-sample-size models require a credible distribution for N; choosing a binomial, Poisson, or negative binomial distribution without substantive justification can distort the assessment of uncertainty.

8.1. Model misspecification

The strongest theoretical results in this area are usually conditional on the assumed model. The hazard transform, normalized spacings, and pivotal quantities work cleanly when the distributional assumptions are correct. In practice, however, the fitted model is an approximation. If the tail of the fitted distribution is too light, far future order statistics will be underpredicted; if it is too heavy, intervals may become unnecessarily wide. This is not a minor issue, because most prediction targets in censored life testing lie beyond the observed part of the sample.

This limitation suggests that future studies should report sensitivity analyses. Instead of presenting predictions under a single distribution, a stronger applied paper can compare Weibull, gamma, modified Kies exponential, modified Muth, and mixture alternatives when they are appropriate for the data. The goal is not to automatically choose the most complicated model, but to check whether the ranking of prediction methods is stable under plausible distributional choices. Accordingly, the present simulation's conclusions are restricted to its correctly specified benchmark. A systematic study in which the generating and fitted distributional families differ is a separate robustness analysis and is left for future work.

8.2. Parameter estimation uncertainty

Many formulas are written as if the distribution F is known. In applications, F is replaced by $F(x;\widehat{\Theta})$. This substitution introduces additional uncertainty that is not always reflected in the reported prediction interval. The issue becomes more serious when r is small, because only a limited number of observed failures is available for estimating Θ. It is also important for multi-parameter lifetime models and mixtures, where standard errors may be large, and parameters can be highly correlated.

A natural research direction is to combine the reviewed prediction formulas with bootstrap or Bayesian uncertainty quantification. For example, one may resample or draw parameter values from an estimated distribution, recompute the future prediction for each draw, and then combine the results. Such an approach would preserve the structure of pivotal or near-MMSE prediction while acknowledging that the fitted model itself is uncertain. This could be especially useful for real data examples, where the number of observed failures is fixed and cannot be increased by simulation.

8.3. Prediction far into the tail

Prediction becomes fragile as s approaches n. The upper order statistics are governed by the tail of the fitted distribution, and small errors in the fitted survival function can produce large differences after applying F^{-1}. This explains why some methods may perform similarly for $s = r + 1$ but diverge for larger s. It also explains why interval width should be reported together with coverage: an interval may achieve high coverage simply because it is too wide to be useful.

The near-MMSE approach is attractive for point prediction, but it should still be assessed carefully in far tail settings. Conditional expectations can be pulled toward the center relative to high quantiles, while pivotal upper bounds may become conservative. Hazard least-squares methods may perform well when the hazard model is

appropriate, but they still depend on numerical optimization and the chosen hazard family. Therefore, a balanced review should avoid a universal claim that one method dominates in all tail regimes.

8.4. Mixture identifiability and numerical stability

Mixture lifetime models are appealing because they represent heterogeneous populations, but they introduce identifiability and numerical challenges. Two different sets of mixture weights and component parameters may produce very similar distribution functions over the observed range but different tail predictions. This is especially problematic under Type-II censoring, because the censored part of the sample is exactly where the mixture components may separate.

For this reason, mixture-based prediction should be accompanied by clear numerical reporting: starting values, parameter constraints, convergence checks, and the method used for quantile inversion. Without this information, it is difficult to decide whether a reported improvement is due to a genuinely better predictive model or to a favorable numerical run. The simulation section partially addresses this by using common performance criteria across methods, but mixture prediction remains an area where more reproducible comparative work is needed.

8.5. Software and reproducibility

There is also a gap between theoretical formulas and reproducible software. Some papers derive exact distribution functions, but practical users still need numerical root-finding, quantile inversion, and model-selection routines. The accompanying R script addresses this gap by comparing pivotal, near-MMSE, mixture, and hazard-based methods under the same simulation scenarios.

Future review papers in this area should treat code as part of the methodology. A table of formulas is useful, but a working implementation reveals details that formulas may hide: boundary corrections for probabilities near 0 or 1, optimization failures, missing upper confidence bounds, and sensitivity to starting values. The windscreen application in this paper illustrates this point. The same observed data can be evaluated by several published prediction ideas, but a clear comparison requires that all methods use the same n, r, future targets, fitted model information, and error measures.

Another research direction is hybrid modeling. Parsimonious lifetime distributions, such as the modified Muth distribution, can serve as interpretable baseline models, as described by Barakat et al. [33], while mixture or machine-learning components can capture heterogeneity. However, such hybrid methods should be evaluated against simple baselines; otherwise, added flexibility may only increase variance without improving prediction.

9. Conclusion

Prediction of future ordered observations is best understood as a combination of three decisions: how to model the lifetime distribution, how to represent the observed ordered information, and how to measure predictive performance. Pivotal methods yield clean interval construction via transformed spacings; near-MMSE conditional expectation methods provide direct point predictors under squared-error loss; random sample-size methods incorporate count uncertainty; mixture models handle heterogeneity; and cumulative hazard least squares provides a useful estimation and prediction framework for generalized ordered and censored data.

This paper provides a critical comparative synthesis and a reproducible, scope-limited assessment of published methods for predicting future order statistics. It examines pivotal prediction procedures, conditional expectation predictors, mixture-based methods, and techniques based on generalized order statistics and cumulative hazard transformations under the settings considered here.

The comparative discussion and simulation results indicate that no single method is uniformly superior across all settings. The performance of a prediction procedure depends on the prediction objective, the censoring structure, the assumed lifetime model, and the characteristics of the target order statistic. Consequently, meaningful evaluation requires multiple performance criteria, including prediction accuracy, bias, interval coverage probability, and interval width.

The study also highlights the importance of aligning the prediction methodology with the underlying data structure. conditional expectation approaches provide competitive point predictions under many Type-II censoring scenarios, whereas pivotal methods remain fundamental when interval estimation and uncertainty quantification are the primary objectives. More flexible models, including heterogeneous and mixture populations, may offer improved performance when their assumptions are adequately supported by the data.

Overall, the prediction of future order statistics should be viewed as a complete statistical workflow that involves model specification, parameter estimation, validation of assumptions, construction of prediction procedures, and objective performance assessment. Future research may focus on robust prediction methods, adaptive procedures for heterogeneous populations, and prediction under increasingly complex censoring mechanisms.

Acknowledgment

This study is supported via funding from Prince sattam bin Abdulaziz University project number (PSAU/2026/R/1448)

REFERENCES

[1] Abu-Moussa, M. H., Almetwally, E. M., & Abd El-Raheem, A. M (2026). Classical and Bayesian inference for progressively censored competing risks data under the Gompertz-Lindley model. *AIMS Mathematics*. 11(4). 12064–12093. doi: https://doi.org/10.3934/math.2026495

[2] Aggarwala, R. (2001). Progressive censoring: A review. *Handbook of Statistics*. 20. 373–429. doi: https://doi.org/10.1016/S0169-7161(01)20015-7.

[3] Ahmadi, J. & Balakrishnan, N (2005). Preservation of Some Reliability Properties by Certain Record Statistics. *Statistics*. 39(4). 347–354. doi: https://doi.org/10.1080/02331880500178752

[4] Raqab, M.Z. & Balakrishnan, N (2008). Prediction Intervals for Future Records. *Statistics & Probability Letters*. 78(13). 1955–1963. doi: https://doi.org/10.1016/j.spl.2008.01.064

[5] Ahmadi, J. & Balakrishnan, N (2011). Distribution-Free Prediction Intervals for Order Statistics Based on Record Coverage. *Journal of the Korean Statistical Society*. 40(2). 181–192. doi: https://doi.org/10.1016/j.jkss.2010.09.003

[6] Ahmadi, M. V. & Doostparast, M. (2018). Bayesian prediction of progressively first-failure-censored order statistics based on k-record values from Weibull distribution. *Istatistik: Journal of the Turkish Statistical Association*. 11(1–2). 12–28.

[7] Ahsanullah, M. (1988). *Introduction to Record Statistics*. Ginn Press. Needham Heights, Massachusetts.

[8] Ahsanullah, M. (1995). *Record Statistics*. Nova Science Publishers. Commack, New York.

[9] Aldallal, R. A. & Barakat, H. M (2026). A Comprehensive Review: Current Record Values and Their Applications. *Applied Mathematics & Information Sciences*. 20(2). 401–423. doi: https://doi.org/10.18576/amis/200208

[10] Aldallal, R. A. (2013). *Study on Some Topics of Records*. Doctoral dissertation, Zagazig University.

[11] Aldallal, R. (2022). Best Linear Unbiased Estimation and Prediction of Record Values Based on Kumaraswamy Distributed Data. *Statistics, Optimization & Information Computing*. 10(4). 1250–1266. doi: https://doi.org/10.19139/soic-2310-5070-1397.

[12] Aldallal, R. A (2023). Moment Generating Function of Current Records Based on Generalized Exponential Distribution with Some Recurrence Relations. *European Journal of Pure and Applied Mathematics*. 16(1). 97–111. doi: https://doi.org/10.29020/nybg.ejpam.v16i1.4641

[13] Aldallal, R. A (2026a). Rayleigh Current Record Distributed Data: Estimation and Prediction with Application on Saudi Arabia Industrial Data. *Journal of Statistical Theory and Applications*. 25(1). doi: https://doi.org/10.1007/s44199-025-00150-x.

[14] Aldallal, R. A (2026b). Bayesian Estimation and Prediction of Weibull Current Records with Application to Saudi Industrial Data. *AIMS Mathematics*. 11(2). 4369–4394. doi: https://doi.org/10.3934/math.2026175

[15] Al-Hussaini, E. K., Abdel-Hamid, A. H., & Hashem, A. F (2015a). Bayesian prediction intervals of order statistics based on progressively Type-II censored competing risks data from the half-logistic distribution. *Journal of the Egyptian Mathematical Society*. 23(1). 190–196. doi: https://doi.org/10.1016/j.joems.2014.01.008

[16] Al-Hussaini, E. K., Abdel-Hamid, A. H., & Hashem, A. F (2015b). One-sample Bayesian prediction intervals based on progressively Type-II censored data from the half-logistic distribution under progressive stress model. *Metrika*. 78(7). 771–783. doi: https://doi.org/10.1007/s00184-014-0526-4.

[17] Al-Olaimat, N. M., Migdadi, H. S., Bayoud, H. A., & Raqab, M. Z. (2026). Prediction Methods for Future Record Values from Two-Parameter Kies Distribution. *Statistics, Optimization & Information Computing*. 15(5). 3381–3400. doi: https://doi.org/10.19139/soic-2310-5070-2123.

[18] Aly, A. E., El-Adll, M. E., Barakat, H. M., & Aldallal, R. A (2023). A new least squares method for estimation and prediction based on the cumulative Hazard function. *AIMS Mathematics*. 8(9). 21968–21992. doi: https://doi.org/10.3934/math.20231120

[19] Arnold, B. C., Balakrishnan, N., & Nagaraja, H. N. (1998). *Records*. Wiley. New York.

[20] Asgharzadeh, A., Marchand, E., & Saadati Nik, A. (2024). Bayesian prediction regions and density estimation with Type-II censored data. *IEEE Transactions on Reliability*. 74(2). 2827–2836. doi: https://doi.org/10.1109/tr.2024.3438189.

[21] Balakrishnan, N. & Aggarwala, R. (2000). *Progressive Censoring: Theory, Methods, and Applications*. Birkhäuser. Boston. doi: https://doi.org/10.1007/978-1-4612-1334-5.

[22] Balakrishnan, N. & Cohen, A. C. (1991). *Order Statistics and Inference: Estimation Methods*. Academic Press. San Diego.

[23] Balakrishnan, N. & Stepanov, A (2004). Two Characterizations Based on Order Statistics and Records. *Journal of Statistical Planning and Inference*. 124(2). 273–287. doi: https://doi.org/10.1016/j.jspi.2003.07.001

[24] Balakrishnan, N., & Cramer, E. (2014). *The Art of Progressive Censoring: Applications to Reliability and Quality*. Birkh"auser. doi: https://doi.org/10.1007/978-0-8176-4807-7.

[25] Barakat, H. M., El-Adll, M. E., & Aly, A. E (2011). Exact prediction intervals for future exponential lifetime based on random generalized order statistics. *Computers & Mathematics with Applications*. 61(5). 1366–1378. doi: https://doi.org/10.1016/j.camwa.2011.01.002.

[26] Barakat, H. M., Nigm, E. M., & Aldallal, R. A. (2013). Current Record Range from Exponential Distribution. In *Alexandria International Conference for Statistics*.

[27] Barakat, H. M., El-Adl, M. E., & Aly, A. E (2014a). Prediction intervals of future observations for a sample of random size. *Mathematics and Computers in Simulation*. 97. 1–13. doi: https://doi.org/10.1016/j.matcom.2013.06.007

[28] Barakat, H. M., Nigm, E. M., & Aldallal, R. A (2014b). Current Records and Record Range with Some Applications. *Journal of the Korean Statistical Society*. 43(2). 263–273. doi: https://doi.org/10.1016/j.jkss.2013.09.004

[29] Barakat, H. M., Nigm, E. M., & Aldallal, R. A. (2014c). Exact Prediction Intervals for Future Current Records and Record Ranges from Any Continuous Distribution. *SORT*. 38(2). 251–270.

[30] Barakat, H. M., Nigm, E. M., El-Adll, M. E., & Yusuf, M. (2018). Prediction of future generalized order statistics based on exponential distribution with random sample size. *Statistical Papers*. 59(2). 605–631. doi: https://doi.org/10.1007/s00362-016-0779-2.

[31] Barakat, H., Khaled, O., & Ghonem, H (2020a). New method for prediction of future order statistics. *Quality Technology & Quantitative Management*. 18(1). 101–116. doi: https://doi.org/10.1080/16843703.2020.1782087

[32] Barakat, H., Khaled, O., & Ghonem, H (2020b). Predicting future lifetime for mixture exponential distribution. *Communications in Statistics - Simulation and Computation*. 51(7). 3533–3552. doi: https://doi.org/10.1080/03610918.2020.1715434

[33] Barakat, H. M., Bakouch, H. S., Elgawad, M. A. A., Semary, H., Alawady, M. A., Husseiny, I. A., Enany, M. G., & Taher, T. S (2026a). The modified Muth distribution: Statistical properties, entropy measures, and parameter estimation. *AIMS Mathematics*. 11(3). 6269–6296. doi: https://doi.org/10.3934/math.2026259

[34] Barakat, H. M., Khaled, O. M., Khalil, K. S., & Harpy, M. H. (2026b). Optimal prediction of future order statistics using conditional expectation under type II censoring for continuous distributions. *Filomat*. 40(1). 333–354. doi: https://doi.org/10.2298/FIL2601333B.

[35] Basak, P. (2000). An Application of Record Range and Some Characterization Results. In *Advances on Theoretical and Methodological Aspects of Probability and Statistics*. pp. 83–95. CRC Press.

[36] Basiri, E., & Ahmadi, J. (2015). Prediction intervals for generalized-order statistics with random sample size. *Journal of Statistical Computation and Simulation*. 85(9). 1725–1741. doi: https://doi.org/10.1080/00949655.2014.896362.

[37] Bdair, O. M. & Raqab, M. Z (2016). One-sequence and two-sequence prediction for future Weibull records. *Journal of Statistical Theory and Applications*. 15(4). 345. doi: https://doi.org/10.2991/jsta.2016.15.4.3

[38] Brock, W., Lakonishok, J., & LeBaron, B (1992). Simple Technical Trading Rules and the Stochastic Properties of Stock Returns. *The Journal of Finance*. 47(5). 1731–1764. doi: https://doi.org/10.1111/j.1540-6261.1992.tb04681.x.

[39] Chandler, K. N (1952). The Distribution and Frequency of Record Values. *Journal of the Royal Statistical Society Series B: Statistical Methodology*. 14(2). 220–228. doi: https://doi.org/10.1111/j.2517-6161.1952.tb00115.x.

[40] Coolen, F. P. A. (2011). Nonparametric Predictive Inference. In M. Lovric (ed.), *International Encyclopedia of Statistical Science*. Springer, Berlin, Heidelberg, pp. 968–970. doi: https://doi.org/10.1007/978-3-642-04898-2_416.

[41] Cramer, E. (2023). Ordered and censored lifetime data in reliability: An illustrative review. *Wiley Interdisciplinary Reviews: Computational Statistics*. 15(2). e1571. doi: https://doi.org/10.1002/wics.1571.

[42] David, H. A. (1973). Concomitants of Order Statistics. *Bulletin of the International Statistical Institute*. 45. 295–300.

[43] Dey, A., Jha, A., & Dey, S. (2016). Bayesian analysis of modified Weibull distribution under progressively censored competing risk model. arXiv preprint arXiv:1605.06585.

[44] Dunsmore, I. R (1983). The Future Occurrence of Records. *Annals of the Institute of Statistical Mathematics*. 35(2). 267–277. doi: https://doi.org/10.1007/bf02480982

[45] Dziubdziela, W. & Kopocinski, B (1976). Limiting Properties of the k-th Record Values. *Applicationes Mathematicae*. 15(2). 187–190. doi: https://doi.org/10.4064/am-15-2-187-190.

[46] El-Adll, M. E., Ateya, S. F., & Rizk, M. M (2012). Prediction intervals for future lifetime of three parameters Weibull observations based on generalized order statistics. *Arabian Journal of Mathematics*. 1(3). 295–304. doi: https://doi.org/10.1007/s40065-012-0004-7.

[47] Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics*. 2 ed. Krieger. Malabar, Florida.

[48] Glick, N (1978). Breaking Records and Breaking Boards. *The American Mathematical Monthly*. 85(1). 2–26. doi: https://doi.org/10.1080/00029890.1978.11994501

[49] Harpy, M. H., Khaled, O. M., El-Morshedy, M., & Khalil, K. S (2025). An improvement in predictive modeling techniques with application to pivotal quantity and least square method. *AIMS Mathematics*. 10(11). 25639–25666. doi: https://doi.org/10.3934/math.20251136

[50] Houchens, R. L. (1984). *Record Value, Theory and Inference*. Doctoral dissertation, University of California, Riverside.

[51] Jaheen, Z. F. & Ali, S. M. A. M (2017). Bayesian prediction for progressive censored data from the Weibull-Geometric model. *American Journal of Mathematical and Management Sciences*. 36(3). 247–258. doi: https://doi.org/10.1080/01966324.2017.1334603

[52] Johnson, N. L. & Kotz, S. (1970). *Continuous Univariate Distributions–1*. Houghton Mifflin. Boston.

[53] Kaminsky, K. S., & Nelson, P. I. (1998). Prediction of order statistics. In *Handbook of Statistics–17: Order Statistics: Applications*. pp. 431–450. Elsevier. doi: https://doi.org/10.1016/S0169-7161(98)17017-7.

[54] Khaled, O. M., Khalil, K. S., & Harpy, M. H. (2023). Predicting future data from gamma-mixture and beta-mixture distributions and application to the recovery rate of COVID-19. *Advances and Applications in Statistics*. 90(1). 1–34. doi: https://doi.org/10.17654/0972361723061.

[55] Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*. John Wiley & Sons. New York.

[56] Lawless, J. F., & Fredette, M. (2005). Frequentist prediction intervals and predictive distributions. *Biometrika*. 92(3). 529–542. doi: https://doi.org/10.1093/biomet/92.3.529.

[57] Muhammad, N., Coolen-Maturi, T., & Coolen, F. P. A. (2018). Nonparametric predictive inference with parametric copulas for combining bivariate diagnostic tests. *Statistics, Optimization & Information Computing*. 6(3). 398–408. doi: https://doi.org/10.19139/soic.v6i3.579.

[58] Nagaraja, H. N (1988). Record Values and Related Statistics-A Review. *Communications in Statistics - Theory and Methods*. 17(7). 2223–2238. doi: https://doi.org/10.1080/03610928808829743

[59] Nevzorov, V. B. & Balakrishnan, N. (1998). A Record of Records. In *Handbook of Statistics–16: Order Statistics: Theory and Methods*. pp. 515–570. North-Holland. Amsterdam.

[60] Nevzorov, V. B. (1987). Records. *Theory of Probability and Its Applications*. 32(2). 201–228. doi: https://doi.org/10.1137/1132032

[61] Nevzorov, V. B. (2001). *Records: Mathematical Theory*. American Mathematical Society. Providence, Rhode Island.

[62] Ahmed, E. A. (2017). Estimation and prediction for the generalized inverted exponential distribution based on progressively first-failure-censored data with application. *Journal of Applied Statistics*. 44(9). 1576–1608. doi: https://doi.org/10.1080/02664763.2016.1214692.

[63] Politis, D. N. (2015). *Model-Free Prediction and Regression: A Transformation-Based Approach to Inference*. Springer, Cham. doi: https://doi.org/10.1007/978-3-319-21347-7.

[64] Raqab, M. R (2009). Distribution-Free Prediction Intervals for the Future Current Record Statistics. *Statistical Papers*. 50(2). 429–439. doi: https://doi.org/10.1007/s00362-007-0082-3.

[65] Resnick, S. I (1973). Limit Laws for Record Values. *Stochastic Processes and their Applications*. 1(1). 67–82. doi: https://doi.org/10.1016/0304-4149(73)90033-1.

[66] Vovk, V., Gammerman, A., & Shafer, G. (2022). *Algorithmic Learning in a Random World*. 2nd ed. Springer, Cham. doi: https://doi.org/10.1007/978-3-031-06649-8.

[67] Wang, Y. & Politis, D. N. (2025). Model-free bootstrap and conformal prediction in regression: conditionality, conjecture testing, and pertinent prediction intervals. *Journal of Nonparametric Statistics*. Advance online publication. doi: https://doi.org/10.1080/10485252.2025.2603398.

[68] Wergen, G., Bogner, M., & Krug, J. (2011). Record statistics for biased random walks, with an application to financial data. *Physical Review E*. 83. 051109. doi: https://doi.org/10.1103/PhysRevE.83.051109.