



Robust doubly censored Weibull modelling of NDVI-based burn-scar persistence in satellite time series

Nora Khalil

Department of Mathematics, Faculty of Science, Capital University (formerly Helwan University), Helwan, Egypt

Correspondence: Nora Khalil (nora_hassan@science.capu.edu.eg)

Received: 4 December 2025 – Revised: 28 March 2026 – Accepted: 30 March 2026 – Published: 7 April 2026

Abstract. Burn-scar persistence in satellite imagery is treated as a lifetime process and modelled within a doubly censored Weibull framework arising from the finite temporal extent of satellite archives. Episodes may appear doubly censored in calendar time because the Landsat record spans a limited analysis window, so that for some pixels neither the true ignition time nor the full recovery are directly observed. On the persistence-time scale used for inference, incomplete episodes contribute as right-censored lifetimes within a unified likelihood formulation. A robust weighted-likelihood estimator with Huber-type influence weights downweights atypically long-lived scars without excluding any observations, thereby limiting the impact of outliers while preserving the central structure of the persistence distribution. The framework is applied to NDVI-based burn-scar persistence derived from Landsat surface-reflectance time series at Monitoring Trends in Burn Severity (MTBS) sampling locations within a large interior-Alaska fire scar, yielding a sample of 90 persistence episodes comprising fully observed and right-censored lifetimes. Weibull parameters are estimated under both maximum likelihood and weighted likelihood, with parametric bootstrap resampling used to obtain confidence intervals. Agreement between model-based and empirical survival behavior is assessed using Kaplan–Meier and Turnbull estimators and a Weibull probability plot. The fitted model indicates a median persistence of approximately 2.7–2.9 years and a 90th-percentile persistence of roughly five to six years, with the Weibull shape parameter consistently exceeding one, indicating an increasing recovery hazard over time. The proposed framework combines interpretability and robustness under censoring and generalizes naturally to other persistence-type environmental processes observed under incomplete monitoring.

1 Introduction

Quantifying how long burned landscapes remain spectrally visible in satellite imagery here termed burn-scar persistence is essential for understanding post-fire recovery and ecological resilience. Even with the continuity of modern Earth-observation archives, incomplete observation windows caused by cloud cover, orbital gaps and sensor outages often prevent full monitoring of fire-affected surfaces. The resulting temporal truncation leads to censoring, where the beginning or end of a persistence episode is unobserved, and thus complicates direct estimation of recovery durations. Similar difficulties arise in other persistence-type phenomena such

as snow-cover longevity (Brown and Mote, 2009), floodwater retention (Sai Bhageerath et al., 2025) and orbital-debris decay (Ashruf et al., 2024), where observational incompleteness can bias lifetime inference (Gross and Clark, 1975). Post-fire vegetation dynamics in remote sensing have typically been described using spectral indices such as the Normalized Difference Vegetation Index (NDVI) and the Normalized Burn Ratio (NBR) derived from MODIS and Landsat imagery. Many studies fit smooth exponential or logistic trajectories to these indices to represent disturbance and recovery (Kennedy et al., 2010; Roy and Boschetti, 2009; Verbesselt et al., 2010), yielding visually intuitive but essentially deterministic descriptions. However, such trajectory-

based models rarely account explicitly for statistical censoring or for radiometric noise in the underlying satellite observations (Kennedy et al., 2010; White et al., 2022) and may therefore underestimate actual recovery durations when recovery extends beyond the observation window. A probabilistic lifetime formulation, by contrast, can incorporate partial observation directly through survival-analysis principles, for example using nonparametric methods such as those of Kaplan and Meier (1958) and Turnbull (1976).

Within this probabilistic setting, the Weibull distribution provides a flexible model for persistence times: its shape parameter characterizes whether recovery accelerates or decelerates, while its scale parameter defines a characteristic duration (Murthy et al., 2004). Standard parameter estimation via maximum likelihood, however, is known to be sensitive to contamination and model deviations (Huber, 1981; Hampel et al., 1986; Huber and Ronchetti, 2009). Satellite data frequently contains outlying or mixed pixels arising from residual cloud, geolocation errors or misclassification, and these anomalies can distort the estimated hazard and bias ecological interpretation. To mitigate such effects, robust estimation methods have been proposed. The weighted-likelihood estimator was developed in the robustness literature to down-weight discordant observations through case-specific likelihood weights (Markatou, 2000), and related minimum-distance and divergence-based approaches have been synthesized by Basu et al. (2011). The underlying principles of bounded influence and adaptive weighting were formalized in the robust M-estimation framework of Huber (1981) and Hampel et al. (1986) and further extended in later treatments (Huber and Ronchetti, 2009; Rousseeuw and Leroy, 1987). Building on these advances, the present study develops a Weibull lifetime model for burn-scar persistence under the doubly censored calendar-time structure induced by a finite observation window, estimated via a Huber-type weighted-likelihood estimator. The framework explicitly accounts for incomplete persistence episodes while moderating the impact of anomalous pixels, thereby improving stability without sacrificing interpretability. Using Landsat surface-reflectance time series at Monitoring Trends in Burn Severity (MTBS) sampling locations within a large interior-Alaska fire scar, with an analysis window from 2010 to 2018, the framework aims to combine statistical rigor with physical realism and to provide a unified tool for extracting meaningful persistence metrics from noisy, incomplete satellite observations.

2 Data and Observational Context

The analysis uses Landsat surface-reflectance time series extracted at Monitoring Trends in Burn Severity (MTBS) sampling locations within a large fire scar in interior Alaska ($\approx 65.5^\circ\text{N}$, $145\text{--}146^\circ\text{W}$; USDA Forest Service and U.S. Geological Survey, 2024, MTBS; see also the MTBS method-

ological documentation in *Fire Ecology*). The MTBS points provide fixed 30 m reference locations inside the mapped burned area, allowing long NDVI histories to be reconstructed without spatial tracking. For each point, all available Landsat 5 TM, Landsat 7 ETM+, and Landsat 8 OLI observations from early 2000 onward were collected and converted to the Normalized Difference Vegetation Index (NDVI). Observations flagged as cloudy, heavily hazy, or snow-covered were discarded, leaving irregular but typically multi-decadal NDVI series for roughly one hundred burned pixels over the period 2000–2020. Burn-scar persistence was defined at each pixel as the time during which vegetation remained in a spectrally degraded post-fire state. To adapt this definition to local conditions, a “healthy vegetation” baseline was computed per pixel as the 80th percentile of its NDVI record. Percentile-based baselines and relative recovery thresholds are commonly used in Landsat time-series studies to define disturbance and recovery in a manner robust to interannual variability (e.g., Pickell et al., 2016; White et al., 2022; Smith-Tripp et al., 2024). A 5-point running median of NDVI was then used to identify disturbance and recovery. A burn episode was declared when the smoothed NDVI fell below 60% of the baseline and remained below this threshold for several consecutive acquisitions; the first such crossing time was taken as the ignition date. Recovery was defined as the first subsequent time when smoothed NDVI returned to at least 85% of the baseline, indicating that pre-fire greenness had effectively re-established. Similar relative-threshold definitions have been employed in post-fire recovery studies using Landsat time series, where recovery is operationally defined as return to a specified fraction of pre-disturbance spectral conditions (Pickell et al., 2016; White et al., 2022).

For each pixel, the dominant episode was taken as the longest contiguous interval below the burn threshold, with onset and recovery dates denoted by S_{true} and E_{true} . To mimic the finite monitoring windows typical of climate and satellite archives (Kennedy et al., 2010; White et al., 2022), a common analysis window was imposed across all pixels. In this study, $[A, B]$ spans 1 January 2010 to 31 December 2018, capturing several post-fire growing seasons while remaining well within the Landsat record. The analysis window (2010–2018) was selected to represent a finite observation period and to ensure that the data exhibit a mixture of fully observed and censored persistence episodes. While the full Landsat record spans 2000–2020, using the entire period would lead to a higher proportion of fully observed episodes and fewer censored cases. In contrast, the selected window provides a more balanced combination of fully observed and right-censored observations, which is consistent with the censored lifetime framework adopted in this study. The choice of window is therefore intended to reflect incomplete monitoring conditions rather than to restrict the available data. Episodes that both ignite and recover within $[A, B]$ contribute to fully observed lifetimes, with persistence $T = E_{\text{true}} - S_{\text{true}}$.

Episodes that ignite within the window but remain spectrally degraded beyond B are right-censored, since only a lower bound on their persistence is observed. A small number of pixels are already in a degraded state at A and have not recovered by B ; these give rise to doubly censored episodes whose ignition and recovery both lie outside the analysis window and whose persistence therefore exceeds the 3286 d span of $[A, B]$. After applying these steps, the final dataset consists of fully observed episodes, right-censored episodes, and a small number of long-lived doubly censored episodes, all expressed as persistence times in days. This configuration reflects key features of satellite-derived impact indicators – finite windows, irregular sampling, and moderate radiometric noise (Kennedy et al., 2010; Verbesselt et al., 2010; White et al., 2022) – and motivates a censored lifetime formulation in which complete and incomplete episodes are estimated jointly rather than discarding partially observed persistence intervals. Although the NDVI thresholds used to define disturbance (60 % of the pixel-specific baseline) and recovery (85 % of the baseline), as well as the 80th-percentile baseline itself, are necessarily heuristic, they are aligned with operational definitions of spectral recovery in Landsat-based disturbance studies. Simple sensitivity checks using alternative thresholds (e.g., 55 %–65 % for disturbance and 80 %–90 % for recovery) yielded qualitatively similar persistence distributions and Weibull parameter estimates, suggesting that the main conclusions are not driven by a single arbitrary choice of cut-off. Additional quantitative sensitivity analyses are provided in Appendix A (Table A2).

3 Conceptual Model of Persistence as a Doubly Censored Lifetime

Satellite archives rarely capture the full life cycle of ecological disturbances. In the case of NDVI-derived burn scars, some pixels are already in a degraded post-fire state when the analysis window opens, while others remain degraded when it closes. As a result, the complete trajectory from ignition to recovery is not always observed. This incomplete visibility naturally leads to censoring, a setting that parallels classical survival and reliability models (Gross and Clark, 1975; Lawless, 2003). We interpret burn-scar persistence as a lifetime variable. For each burned pixel, the persistence lifetime begins at the true onset of spectral degradation and ends when NDVI returns to a near-baseline level. Let S_{true} denote the true onset time of degradation and E_{true} the true recovery time. The persistence time is therefore $T = E_{\text{true}} - S_{\text{true}}$, $T > 0$, measured in days from ignition to recovery. The satellite data are observed only within a fixed temporal window $[A, B]$, defined in Sect. 2. Because this window does not necessarily cover the full interval $[S_{\text{true}}, E_{\text{true}}]$, several censoring patterns arise in calendar time:

- Fully observed episodes: $S_{\text{true}} \geq A$ and $E_{\text{true}} \leq B$. The entire persistence interval lies within the window, and the observed duration equals the true lifetime.
- Episodes beginning before the window: $S_{\text{true}} < A$ and $E_{\text{true}} \leq B$. Degradation has already started when observation begins.
- Episodes ending after the window: $S_{\text{true}} \geq A$ and $E_{\text{true}} > B$. Recovery occurs beyond the observation period.
- Episodes spanning the entire window: $S_{\text{true}} < A$ and $E_{\text{true}} > B$. The pixel remains degraded throughout $[A, B]$.

These censoring configurations are illustrated schematically in Fig. 1. In this paper, the term “doubly censored” refers to the calendar-time structure induced by the finite observation window $[A, B]$, in which both ignition and recovery may lie outside the window. Statistical inference, however, is conducted entirely on the persistence-time scale T . On that scale, incomplete episodes do not generate interval-censored lifetimes; rather, they contribute as right-censored observations at their observed lower bound. The statistical model is formulated entirely in terms of the lifetime variable T , not the absolute calendar dates. When expressed on the T -scale, all incomplete episodes contribute as right-censored observations at their observed lower bound. In particular, episodes degraded throughout the window provide the information that $T > B - A$. Similarly, episodes that begin prior to A are left-censored in calendar time but provide right-censored information on the persistence-time scale. Thus, although the observation window induces patterns that may appear as left- or doubly censored in calendar time, the empirical likelihood for the present data involves only fully observed and right-censored persistence times. No interval-censored lifetimes arise on the T -scale in this application. The broader “doubly censored” perspective serves primarily to clarify how finite satellite archives generate incomplete lifetimes, whereas the statistical analysis itself operates on the persistence-time scale.

To model persistence, we assume that T follows a Weibull distribution with shape parameter $b > 0$ and scale parameter $\theta > 0$. The cumulative distribution function is

$$F(t) = 1 - e^{-\left(\frac{t}{\theta}\right)^b} \tag{1}$$

and the corresponding density is

$$f(t) = \frac{b}{\theta} \left(\frac{t}{\theta}\right)^{b-1} e^{-\left(\frac{t}{\theta}\right)^b} \tag{2}$$

as in standard Weibull lifetime models (Murthy et al., 2004). For a fully observed persistence time t_i , the likelihood contribution is $f(t_i)$. For a right-censored episode observed up to

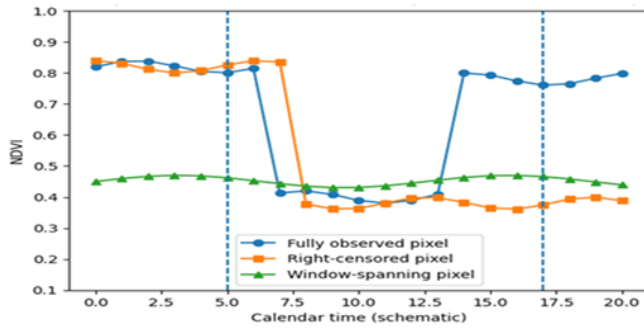


Figure 1. Example NDVI trajectories illustrating calendar-time censoring patterns. The vertical dashed lines mark the opening (*A*) and closing (*B*) dates of the satellite analysis window. Fully observed pixels recover within the window, right-censored pixels remain degraded beyond *B*, and doubly censored pixels are degraded throughout the entire observation window.

time c_i , the contribution is the survival probability $1 - F(c_i)$. Accordingly, the likelihood used in this study is the standard right-censored lifetime likelihood familiar from survival analysis (Lawless, 2003; Turnbull, 1976). Framing burn-scar persistence as a censored lifetime process provides a transparent bridge between satellite data limitations and established statistical tools for incomplete time-to-event data. This formulation also enables principled extensions, including robust estimation and alternative distributional choices, which are developed in the following section. In Fig. 1, schematic NDVI trajectories illustrate fully observed, right-censored, and window-spanning episodes within the observation window $[A, B]$.

In summary, although some episodes are left- or doubly censored in calendar time, the statistical inference is conducted entirely on the persistence-time scale, where incomplete episodes enter the likelihood as right-censored observations. This distinction ensures internal consistency between the conceptual description and the formal likelihood formulation.

4 Methodology

This section describes the estimation procedures used to model burn-scar persistence under censoring and to incorporate robustness through a weighted-likelihood estimator (WLE). The framework has three components: (i) formulation of the Weibull likelihood under general censoring; (ii) introduction of adaptive weighting to achieve robustness against outliers and model deviations; and (iii) quantification of uncertainty using parametric bootstrap resampling.

4.1 Weibull likelihood under general censoring

Let T_i denote the persistence time of the i th burned pixel, assumed to follow a Weibull distribution with shape $b > 0$

and scale $\theta > 0$, with distribution and density given in Eqs. (1)–(2). In general, incomplete lifetime data may be represented by intervals (L_i, U_i) on the persistence-time scale, where $0 \leq L_i < U_i \leq \infty$. Here, (L_i, U_i) denotes the lower and upper bounds within which the true persistence time T_i is known to lie under the censoring mechanism. In other words, the available data implies only that $T_i \in (L_i, U_i)$. When the episode is fully observed, the bounds coincide at the exact lifetime, so that $L_i = U_i = T_i$. Fully observed episodes satisfy $L_i = U_i = T_i$; pure left-censoring arises when $L_i = 0$ and $U_i < \infty$; pure right-censoring corresponds to $L_i < \infty$ and $U_i = \infty$; and interval-censored observations have $0 < L_i < U_i < \infty$ (Lawless, 2003; Turnbull, 1976). In the MTBS application, although some episodes appear doubly censored on the calendar-time axis (Sect. 3), the statistical analysis is conducted on the persistence-time scale. On that scale, incomplete episodes contribute as right-censored observations at their observed duration. In particular, pixels degraded throughout $[A, B]$ provide the information that $T > B - A$ and enter the likelihood as right-censored observations. Let δ_i^{obs} , δ_i^L , δ_i^R , and δ_i^I denote indicators for fully observed, left-censored, right-censored, and interval-censored observations, respectively, satisfying $\delta_i^{\text{obs}} + \delta_i^L + \delta_i^R + \delta_i^I = 1$. The general censored log-likelihood is

$$\ell(b, \theta) = \sum_{i=1}^n [\delta_i^{\text{obs}} \log f(T_i; b, \theta) + \delta_i^L \log F(U_i; b, \theta) + \delta_i^R \log S(L_i; b, \theta) + \delta_i^I \log \{F(U_i; b, \theta) - F(L_i; b, \theta)\}], \tag{3}$$

where F and $S = 1 - F$ denote the Weibull distribution and survival functions. In the empirical analysis below, only fully observed and right-censored persistence times arise on the lifetime scale, so the likelihood reduces to contributions of the form $f(t_i)$ and $S(c_i)$. Consequently, no interval-censored lifetimes arise in the MTBS application, and Eq. (3) simplifies accordingly. Maximizing $\ell(b, \theta)$ yields the ordinary maximum-likelihood estimator (MLE), which is efficient under correct specification but can be sensitive to contamination and model deviations.

4.2 Weighted-likelihood formulation

To improve robustness, we adopt a weighted-likelihood approach following Markatou (2000) and the divergence-based framework of Basu et al. (2011). Each observation is assigned a weight $w_i \in [0, 1]$ that moderates its influence according to its agreement with the current fitted model. The weighted log-likelihood is

$$\ell_W(b, \theta) = \sum_{i=1}^n w_i \ell_i(b, \theta),$$

where $\ell_i(b, \theta)$ denotes the individual contribution in Eq. (3).

Deviance-type residuals

Weights are constructed from standardized deviance-type residuals based on model-implied cumulative probabilities.

Define

$$u_i = \begin{cases} F(T_i; b, \theta), & \text{if fully observed,} \\ U_i, U_i \sim \text{Uniform}(F(L_i; b, \theta), 1), & \text{if right-censored,} \end{cases}$$

so that right-censored observations use randomized quantile residuals (Dunn and Smyth, 1996). The residual is obtained via the normal-score transformation, $r_i^{(0)} = \Phi^{-1}(u_i)$, where Φ^{-1} is the standard normal quantile function. To achieve scale invariance, residuals are standardized using the median absolute deviation (MAD),

$$r_i = \frac{r_i^{(0)}}{\text{MAD}(r^{(0)})}.$$

A Huber-type weight function (Huber, 1981; Huber and Ronchetti, 2009) is then applied

$$w_i = \begin{cases} 1, & |r_i| \leq c, \\ \frac{c}{|r_i|}, & |r_i| > c, \end{cases} \quad (4)$$

where $c > 0$ is a tuning constant controlling the bias-robustness trade-off.

Iterative algorithm

The weighted-likelihood estimator (WLE) $\hat{\theta}_W = (\hat{b}_W, \hat{\theta}_W)$ is obtained by:

- Initializing at the MLE.
- Computing residuals and updating weights.
- Maximizing $\ell_W(b, \theta)$.
- Repeating until successive updates change by less than 10^{-6} .
- When all weights equal one, the procedure reduces to the MLE.

4.3 Data-driven choice of the tuning constant

The tuning constant c governs the strength of downweighting. Rather than fixing it a priori, we adopt a data-driven selection rule. Let \hat{F}_{NP} denote a nonparametric estimator of the persistence-time distribution (Kaplan–Meier or Turnbull, depending on the censoring structure), and let $\hat{F}_{W,c}$ denote the Weibull distribution fitted via WLE with tuning constant c . For candidate values c_1, \dots, c_M , define

$$D(c) = \int_0^\tau \left\{ \hat{F}_{W,c}(t) - \hat{F}_{NP}(t) \right\}^2 dt \quad (5)$$

where τ is a high quantile (e.g., the 95th percentile) of \hat{F}_{NP} . The selected constant is $\hat{c} = \arg \min_{c \in \{c_1, \dots, c_M\}} D(c)$. The selection rule is heuristic rather than likelihood-based, as it does not maximize a parametric criterion. Instead, it aims to balance robustness and fidelity to the empirical survival curve over the well-informed central duration range. The squared-loss functional was chosen for transparency and numerical stability, providing a simple measure of global discrepancy while modestly penalizing larger deviations. The truncation point τ limits the influence of sparsely populated upper-tail regions, where censoring is heavy and nonparametric estimates are inherently variable. As shown in Sect. 5.4, parameter estimates remain stable across a reasonable range of tuning constants, indicating limited sensitivity to this choice. This criterion is analogous to minimum-distance fitting against the nonparametric distribution over its central support, thereby limiting the influence of extreme tail behavior.

4.4 Bootstrap confidence intervals

Uncertainty in (b, θ) is assessed using parametric bootstrap resampling. For each of $B = 1000$ replications, a synthetic dataset is generated from the fitted robust Weibull model $(\hat{b}_W, \hat{\theta}_W)$, preserving the observed censoring pattern. Specifically, lifetimes T_i^* are generated from the fitted model and combined with the original censoring times C_i to form

$$T_i^{\text{obs},*} = \min(T_i^*, C_i), \quad \delta_i^* = \mathbf{1}(T_i^* \leq C_i).$$

The weighted-likelihood algorithm with tuning constant \hat{c} is then applied to each bootstrap sample. Percentile-based 95 % confidence intervals are obtained from the empirical distribution of the bootstrap replicates. This approach captures sampling variability and the impact of censoring while avoiding reliance on information-matrix approximations under robustness adjustments (Efron and Tibshirani, 1993; Huber and Ronchetti, 2009).

4.5 Asymptotic properties

The WLE belongs to the class of divergence-based robust estimators for censored lifetime models (Markatou, 2000; Basu et al., 2011). Under standard regularity conditions for censored parametric models and bounded, smooth weight functions, the solution $\hat{\theta}_W$ to the weighted score equation

$$U_W(\theta) = \sum_{i=1}^n w_i(\theta) u_i(\theta) = 0 \quad (6)$$

is consistent and asymptotically normal:

$$\sqrt{n}(\hat{\theta}_W - \theta_0) \xrightarrow{d} N_2(0, \Sigma_W),$$

where $\Sigma_W = J_W^{-1} K_W J_W^{-1}$ is the usual sandwich covariance matrix based on the limiting sensitivity and variability matrices of the weighted score. When $w_i \equiv 1$, this reduces to the

classical MLE covariance structure. In practice, bootstrap-based inference is preferred here due to censoring and adaptive weighting. In the present application, rather than relying on closed-form expressions for Σ_W under censoring, inference is based on the parametric bootstrap described in Sect. 4.4.

5 Results and Interpretation

Table 1 summarizes the observed and censored burn-scar episodes in the MTBS sample ($n = 90$). In the MTBS sample, 49 episodes (54 %) are fully observed and 35 (39 %) are right-censored on the persistence-time scale. An additional 6 episodes (7 %) remain degraded throughout the analysis window. These latter cases are doubly censored in calendar time but contribute as right-censored observations on the persistence-time scale. Nearly half of all episodes are therefore incomplete, underscoring the need for an explicit censored-lifetime formulation rather than relying on raw observed durations.

5.1 Model fit and parameter estimation

Fitting the Weibull model to the 90 MTBS episodes yields stable parameter estimates under both maximum likelihood (MLE) and weighted likelihood (WLE). Table 2 summarizes numerical results with bootstrap confidence intervals. Under both methods, the Weibull shape parameter exceeds one, indicating an increasing recovery hazard: the longer a scar persists, the more likely it is to recover in subsequent months. The scale parameter corresponds to persistence times on the order of two to three years, consistent with typical recovery trajectories in boreal and sub-arctic ecosystems (White et al., 2022; Smith-Tripp et al., 2024). The WLE primarily reduces sampling variability driven by a small number of very long-lived scars, including the window-spanning episodes discussed in Sect. 3, while leaving the overall recovery timescale essentially unchanged. The robust estimator produces slightly more stable estimates of the median and upper quantiles but preserves the central tendency. Overall, the Weibull model provides a compact and interpretable summary of the persistence distribution.

5.2 Model selection and validation

To assess distributional adequacy, lognormal and gamma models were fitted under the same right-censored likelihood structure used for the Weibull model. In all cases, episodes that are doubly censored in calendar time were represented as right-censored persistence times with $T_i > B - A$, ensuring direct comparability of likelihood-based criteria across distributions. Table 3 reports log-likelihoods and information criteria. The lognormal model attains the smallest AIC and BIC values, followed by the gamma distribution, while the Weibull MLE performs slightly worse un-

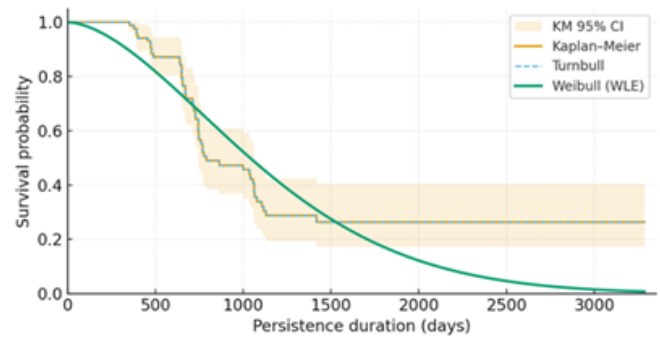


Figure 2. Nonparametric and parametric survival estimates for NDVI-based burn-scar persistence.

der these criteria. The WLE version yields somewhat larger AIC/BIC values due to the effective reduction in influence from downweighted observations, though overall model fit remains comparable. Although the lognormal distribution achieves the smallest information criteria, estimated median persistence times and upper-tail quantiles are broadly comparable across the three models (Table 3). We retain the Weibull family as the primary model because it provides a direct monotone hazard interpretation and integrates naturally with the robust weighted-likelihood framework under right-censoring. The estimated shape parameter consistently exceeds one (Table 2), indicating an increasing recovery hazard over time – a pattern that aligns with ecological expectations of post-fire regrowth. Given the similarity of central and upper-tail persistence estimates across all candidate distributions, the choice of the Weibull model is guided by interpretability and robustness rather than marginal differences in information criteria. Nonparametric survival curves (Kaplan–Meier and Turnbull) were compared with the fitted WLE–Weibull survival curve (Fig. 2). The robust Weibull curve closely follows both empirical estimators over the central duration range of approximately 500–1500 d, with only modest deviations in the tails where data are sparse.

5.3 Robustness diagnostics: bootstrap and influence weighting

Robustness was examined using the bootstrap procedure described in Sect. 4.4. The bootstrap densities for the shape and scale parameters (Fig. 3a–b) exhibit narrower dispersion under the WLE than under the MLE, indicating that the robust procedure reduces variability driven by a small number of long-lived scars. Importantly, the central locations of the bootstrap distributions remain very similar under both estimators, suggesting that the WLE stabilizes inference without materially altering the central tendency. Figure 3c displays the Huber weight function at the selected tuning constant ($\hat{c} \approx 1.3$). Episodes with $|r| \leq 1.3$ receive full weight, while more discordant observations are progressively downweighted. Only a small fraction fall into the strongly down-

Table 1. Summary of observed and censored burn-scar episodes in the MTBS sample ($n = 90$).

Censoring type	Count	Percent	Mean L_{obs} (days)	Median L_{obs} (days)	Min (days)	Max (days)
Window-spanning (right-censored on T -scale)	6	6.7 %	3286.0	3286.0	3286.0	3286.0
Observed	49	54.4 %	745.7	728.0	352.0	1418.0
Right-censored	35	38.9 %	781.8	549.1	132.1	1637.1

Table 2. Weibull parameter estimates and selected quantiles (days).

Method	Shape b	95 % CI for b	Scale θ (days)	95 % CI for θ (days)	$t_{0.1}$ (days)	Median $t_{0.5}$ (days)	$t_{0.9}$ (days)
MLE	1.43	(1.22, 1.76)	1568	(1238, 2068)	323	1212	2815
WLE	1.70	(1.44, 2.07)	1290	(1051, 1594)	343	1039	2108

Table 3. Information criteria for alternative models ($n = 90$).

Model	Estimation	Log-likelihood	AIC	BIC
Weibull	MLE	-408.8	821.5	826.5
Weibull	WLE	-413.3	830.6	835.6
Lognormal	MLE	-396.1	796.2	801.2
Gamma	MLE	-404.5	813.0	818.0

Table 4. Sensitivity of WLE–Weibull estimates to c .

Huber c	\hat{b}	$\hat{\theta}$ (days)	$t_{0.1}$ (days)	$t_{0.5}$ (days)	$t_{0.9}$ (days)
1.0	1.80	1224	351	998	1945
1.3	1.70	1290	343	1039	2108
1.5	1.65	1331	340	1066	2208

weighted region, indicating that most pixels exhibit broadly consistent recovery dynamics and that the robustness gains arise primarily from moderating a limited set of extreme cases.

A Weibull probability plot for uncensored episodes is shown in Fig. 4. The empirical points display approximate linearity over the central quantile range, consistent with the Weibull specification. Departures are more visible in the lower and upper tails, where the number of uncensored observations is small and censoring is relatively heavy. Such deviations are consistent with finite-sample variability under tail sparsity and do not materially affect inference for the central portion of the persistence distribution, which drives the primary ecological interpretation. Taken together, Figs. 3 and 4 support the use of the WLE: it enhances numerical stability and mitigates the influence of extreme observations while preserving the primary structure of the persistence-time distribution.

5.4 Sensitivity to the Huber tuning constant

To assess sensitivity to the Huber constant, the WLE–Weibull model was re-estimated for fixed values $c \in \{1.0, 1.3, 1.5\}$. Table 4 summarizes the results. Across this range, the estimated shape and scale parameters – and the implied 10th, 50th, and 90th percentiles – vary only modestly. Median persistence remains close to approximately 2.7–2.9 years, and the 90th percentile lies between roughly 5.3 and 6.0 years. The data-driven choice \hat{c} falls within this range and yields

estimates very close to those for $c = 1.3$. This stability indicates that substantive conclusions about typical and extreme recovery times are not driven by the precise choice of the robustness parameter.

6 Discussion

The results show that treating NDVI-based burn-scar persistence as a censored lifetime yields a coherent and physically meaningful description of recovery. In the MTBS sample, more than 40 % of episodes are incomplete either right-censored within the analysis window or degraded throughout it so that an approach based solely on raw observed durations would disregard a substantial portion of the available information. Within the Weibull family, the shape parameter summarizes whether the recovery hazard increases or decreases with time since fire, while the scale parameter defines a characteristic persistence horizon. Both quantities behave in an ecologically plausible manner, with median recovery times around 2.7–2.9 years and 90th-percentile durations of roughly five to six years. A key limitation of the present analysis is the modest sample size ($n = 90$ episodes, of which only six span the entire observation window), which limits precision in characterizing the extreme upper tail of the persistence distribution and constrains exploration of more complex model structures. The definition of disturbance and recovery involves several NDVI thresholds: a pixel-specific baseline defined as the 80th percentile of the NDVI record, a 60 % drop below this baseline to flag disturbance, and an

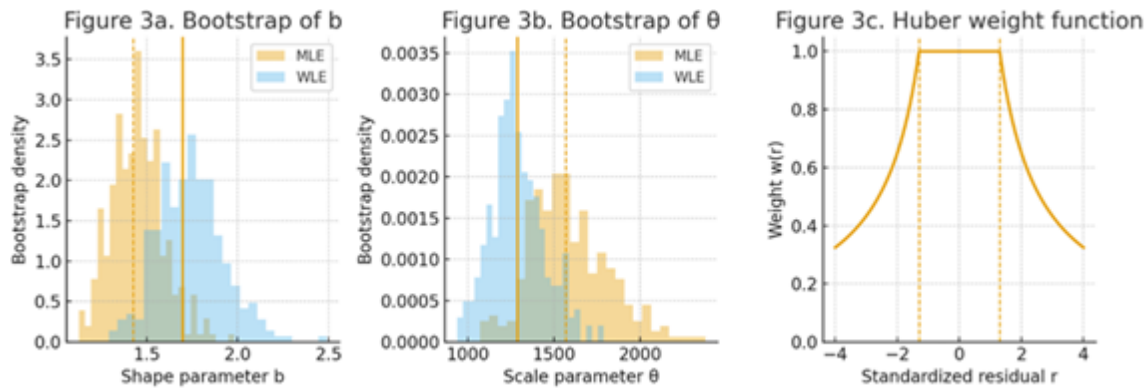


Figure 3. Bootstrap diagnostics and weighting function for the Weibull model: (a) bootstrap densities of the shape parameter b under MLE and WLE; (b) bootstrap densities of the scale parameter θ ; (c) Huber weight function with cutoff at $|r| = 1.3$.

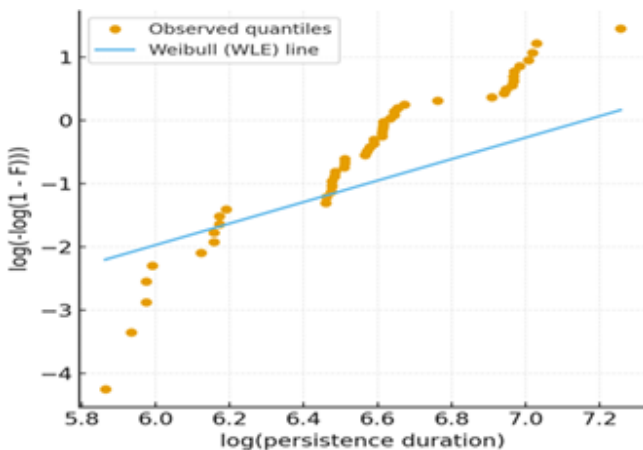


Figure 4. Weibull probability plot for uncensored burn-scar persistence episodes. Points are empirical quantiles of observed durations, and the line is the fitted WLE–Weibull model.

85 % return threshold for recovery. These choices are necessarily heuristic but align with established practice in Landsat-based post-fire recovery studies. Sensitivity analyses with alternative thresholds (e.g., 55 %–65 % for disturbance and 80 %–90 % for recovery) yielded qualitatively similar persistence distributions and Weibull parameter estimates, indicating that the main findings are not driven by a single arbitrary specification. Methodologically, the weighted-likelihood estimator (WLE) provides clear robustness gains relative to standard maximum likelihood. Bootstrap diagnostics show that a small number of very long-lived scars inflate the variance of the MLE. Under the WLE, adaptive Huber weights moderate the influence of these episodes while leaving the bulk of the distribution essentially unchanged. This behavior is consistent with the theoretical properties of bounded-weight estimators, which stabilize likelihood curvature and reduce sensitivity to atypical observations while retaining the large-sample properties of likelihood-based inference. The data-driven selection of the tuning constant further enhances

stability by adapting the robustness level to the observed lifetime distribution rather than fixing it a priori. Comparisons with alternative parametric forms reinforce these conclusions. Although the lognormal model attains the lowest AIC and BIC values and the gamma model performs competitively, neither offers the same combination of monotone hazard interpretation and well-developed robust censored-likelihood methodology available within the Weibull family. The WLE–Weibull fit closely follows the Kaplan–Meier and Turnbull curves over the central range of the data and yields a recovery hazard consistent with established patterns of post-fire regrowth. We therefore retain the Weibull family, estimated via WLE, as the primary modelling framework and regard the lognormal and gamma fits as benchmarks that confirm the overall stability of the substantive conclusions. Taken together, the results indicate that the proposed framework achieves three aims:

- i. It accommodates fully observed and right-censored persistence episodes within a unified lifetime model.
- ii. It embeds robustness directly into likelihood-based inference, addressing radiometric noise, misclassified pixels, and unusually persistent scars; and
- iii. It yields ecologically interpretable summaries of recovery timescales.

These features make the approach well suited to large or heterogeneous satellite datasets, where censoring and measurement noise are unavoidable and direct modelling of raw NDVI trajectories is often impractical. The same conceptual framework can be extended to other disturbance-and-recovery processes observed via remote sensing.

7 Conclusions

This study develops a robust lifetime-based framework for analyzing burn-scar persistence from NDVI time series, explicitly accommodating the censored structures that arise

from the finite temporal extent of satellite archives. By modelling persistence as a Weibull lifetime and incorporating adaptive Huber weighting into the likelihood, the method combines a physically interpretable description of recovery with statistically stable estimation in the presence of noise, misclassified pixels, and unusually persistent scars.

Applied to 90 MTBS persistence episodes, the Weibull model captures the principal features of the recovery distribution, with median recovery times of approximately 2.7–2.9 years and substantially longer durations in the upper tail. The weighted-likelihood estimator stabilizes inference by reducing the influence of a small number of extreme episodes while preserving the central tendency of the distribution. Comparisons with nonparametric survival curves indicate that the robust Weibull fit aligns closely with the empirical lifetime structure over the central range of the data. Although the lognormal model attains slightly smaller AIC/BIC values and the gamma model performs competitively, neither offers the same combination of monotone hazard interpretation and an established robust censored-likelihood framework. In this sense, the Weibull–WLE formulation provides a balanced compromise between interpretability and robustness.

The framework is generic and can be extended to other satellite-based disturbance and restoration processes, such as vegetation browning, drought impacts, land-use transitions, and flood-recovery cycles. In the context of fire, the same methodology could be applied to major events outside interior Alaska – for example, the 2019–2020 Australian bushfires or large Amazonian drought fires – by adapting only the NDVI extraction and baseline definitions. Future work may incorporate spatial pooling across pixels, covariate effects such as climate forcing or fire severity, and hierarchical structures that combine multiple fire events within a unified statistical model. Overall, the proposed approach provides a principled and practical tool for extracting recovery timescales from incomplete and noisy remote-sensing observations, enabling more consistent characterization of post-fire ecosystem dynamics across diverse biomes and fire regimes.

Appendix A: Sensitivity of Parametric Inference

To examine the stability of the Weibull-based results, we carried out additional sensitivity analyses addressing two separate aspects of the modelling framework:

1. the robustness specification and the treatment of window-spanning episodes (Table A1), and
2. the preprocessing choices used to derive burn-scar persistence from the NDVI time series (Table A2).

Unless otherwise stated, all results are based on the same fixed set of 90 MTBS persistence episodes used in the main analysis.

A1 Sensitivity to Robustness Specification

We first evaluated how sensitive the results are to the choice of robustness tuning parameter in the weighted-likelihood estimator (WLE). The model was re-estimated for fixed values of the Huber constant $c \in \{1.0, 1.3, 1.5\}$, covering a reasonable range of robustness levels.

In addition, as a diagnostic comparison, we re-fitted the Weibull model after excluding the six window-spanning episodes (i.e., pixels that remained degraded throughout the analysis window). This alternative specification results in $n = 84$ observations.

The resulting parameter estimates and selected quantiles are reported in Table A1.

As shown in Table A1:

- the Weibull shape parameter remains strictly greater than one under all robustness settings,
- median persistence varies only moderately across specifications,
- upper-tail quantiles remain within the same multi-year range.

Excluding the window-spanning episodes reduces the influence of the extreme upper tail, as expected, but does not change the qualitative conclusion that recovery hazard increases over time. Overall, Table A1 indicates that the principal findings are not sensitive to the choice of robustness parameter nor to the inclusion of the longest-lived scars.

A2 Sensitivity to NDVI Preprocessing Choices

We next evaluated the effect of the four main preprocessing decisions used to define burn-scar persistence:

- baseline percentile (75 %, 80 %, 85 %),
- disturbance threshold (55 %, 60 %, 65 % of the baseline),
- recovery threshold (80 %, 85 %, 90 % of the baseline),
- running-median window length (3 versus 5 observations).

For each alternative specification, persistence episodes were reconstructed from the original NDVI time series using:

- the same analysis window (1 January 2010–31 December 2018),
- the same dominant-episode selection rule, and
- the same fixed set of 90 MTBS sampling locations.

The Weibull model was then re-estimated using the same right-censored likelihood formulation as in the main analysis.

The results are presented in Table A2.

Several patterns are clear from Table A2:

Table A1. Sensitivity of Weibull parameter estimates to the choice of Huber tuning constant and to the treatment of window-spanning episodes ($n = 90$ unless otherwise stated).

Specification	Shape b	Scale θ (days)	$t_{0.1}$ (days)	Median $t_{0.5}$ (days)	$t_{0.9}$ (days)
WLE ($c = 1.0$)	2.22	1102	351	935	1604
WLE ($c = 1.3$)	1.95	1185	343	982	1816
WLE ($c = 1.5$)	1.92	1176	340	971	1816
MLE (exclude window-spanning, $n = 84$)	2.45	1082	331	931	1521

Table A2. Sensitivity of Weibull maximum-likelihood estimates to alternative NDVI baseline percentiles, disturbance and recovery thresholds, and running-median window lengths ($n = 90$).

Specification	n	n_{observed}	Shape b	Scale θ (days)	Median (days)
baseline_75	90	74	3.0409	913.451	809.732
baseline_80	90	63	2.8753	1162.720	1023.566
baseline_85	90	47	2.5331	1654.011	1431.204
disturb_55	90	63	2.8747	1162.190	1023.072
disturb_60	90	63	2.8753	1162.720	1023.566
disturb_65	90	63	2.8702	1165.183	1025.504
recover_80	90	65	2.7620	1088.823	953.514
recover_85	90	63	2.8753	1162.720	1023.566
recover_90	90	58	2.8780	1254.511	1104.504
runmed_3	90	76	3.9984	539.813	492.532
runmed_5	90	63	2.8753	1162.720	1023.566

- The Weibull shape parameter remains greater than one under every specification, indicating a consistently increasing recovery hazard.
- Median persistence remains on the order of several years for the main and nearby specifications.
- No alternative threshold choice produces a qualitatively different interpretation of recovery dynamics.

The disturbance threshold has almost no effect on the parameter estimates. In contrast, changes to the baseline and recovery thresholds produce gradual, monotonic shifts in estimated persistence timescales, which is expected since these parameters directly affect episode duration. The shorter smoothing window ($k = 3$) yields shorter estimated durations, as reduced smoothing allows more rapid threshold crossings, but the increasing hazard interpretation remains unchanged. Taken together, the results in Table A2 confirm that the main conclusions of the study are not driven by any single threshold, percentile choice, or smoothing parameter.

Code availability. The scripts used to construct NDVI trajectories, derive censored persistence episodes, and fit the Weibull models (MLE and WLE) were implemented in Python. The code is available from the corresponding author upon reasonable request.

Data availability. Landsat surface-reflectance data used in this study are publicly available from the USGS Earth Explorer platform. MTBS sampling locations and fire perimeters are available through the MTBS project archive. The dataset generated and analyzed during this study, including NDVI time series and derived burn-scar persistence lifetimes, is publicly available at Zenodo: <https://doi.org/10.5281/zenodo.19340710> (Khalil, 2026).

Competing interests. The author has declared that there are no competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. The authors bear the ultimate responsibility for providing appropriate place names. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Review statement. This paper was edited by Dan Cooley and reviewed by two anonymous referees.

References

- Ashruf, A. M., Bhaskar, A., Vineeth, C., and Pant, T. K.: Deciphering solar cycle influence on long-term orbital deterioration of low-Earth orbiting space debris, arXiv [physics.space-ph], <https://doi.org/10.48550/arXiv.2405.08837>, 2024.
- Basu, A., Shioya, H., and Park, C.: Statistical Inference: The Minimum Distance Approach, Chapman & Hall/CRC, Boca Raton, FL, <https://doi.org/10.1201/b10956>, 2011.
- Brown, R. D. and Mote, P. W.: The response of Northern Hemisphere snow cover to a changing climate, *J. Climate*, 22, 2124–2145, <https://doi.org/10.1175/2008JCLI2665.1>, 2009.
- Dunn, P. K. and Smyth, G. K.: Randomized quantile residuals, *J. Comput. Graph. Stat.*, 5, 236–244, 1996.
- Efron, B. and Tibshirani, R. J.: An Introduction to the Bootstrap, Chapman & Hall/CRC, New York, <https://doi.org/10.1201/9780429246593>, 1993.
- Gross, A. J. and Clark, V. A.: Survival Distributions: Reliability Applications in the Biomedical Sciences, Wiley, New York, ISBN 9780471328179, 1975.
- Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., and Stahel, W. A.: Robust Statistics: The Approach Based on Influence Functions,

- Wiley, New York, <https://doi.org/10.1002/9781118186435>, 1986.
- Huber, P. J.: Robust Statistics, Wiley, New York, <https://doi.org/10.1002/0471725250>, 1981.
- Huber, P. J. and Ronchetti, E. M.: Robust Statistics, 2nd Edn., Wiley, Hoboken, NJ, <https://doi.org/10.1002/9780470434697>, 2009.
- Kaplan, E. L. and Meier, P.: Nonparametric estimation from incomplete observations, *J. Am. Stat. Assoc.*, 53, 457–481, <https://doi.org/10.1080/01621459.1958.10501452>, 1958.
- Kennedy, R. E., Yang, Z., and Cohen, W. B.: Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr – Temporal segmentation algorithms, *Remote Sensing of Environment*, 114, 2897–2910, <https://doi.org/10.1016/j.rse.2010.07.008>, 2010.
- Khalil, N.: NDVI Burn-Scar Persistence Dataset and Sensitivity Analysis (MTBS, Alaska), Zenodo [data set], <https://doi.org/10.5281/zenodo.19340710>, 2026.
- Lawless, J. F.: Statistical Models and Methods for Lifetime Data, 2nd Edn., Wiley, Hoboken, NJ, <https://doi.org/10.1002/9781118033005>, 2003.
- Markatou, M.: Mixture models, robustness, and weighted likelihood methodology, *Biometrics*, 56, 483–486, <https://doi.org/10.1111/j.0006-341X.2000.00483.x>, 2000.
- Murthy, D. N. P., Xie, M., and Jiang, R.: Weibull Models, Wiley, Hoboken, NJ, <https://doi.org/10.1002/047147326X>, 2004.
- Pickell, P. D., Hermosilla, T., Coops, N. C., Wulder, M. A., and White, J. C.: Forest recovery trends derived from Landsat time series for North American boreal forests, *Int. J. Remote Sens.*, 37, 138–149, <https://doi.org/10.1080/2150704X.2015.1126375>, 2016.
- Roy, D. P. and Boschetti, L.: Southern Africa validation of the MODIS, L3JRC, and GlobCarbon burned-area products, *IEEE T. Geosci. Remote Sens.*, 47, 1032–1044, 2009.
- Rousseeuw, P. J. and Leroy, A. M.: Robust Regression and Outlier Detection, Wiley, New York, <https://doi.org/10.1002/0471725382>, 1987.
- Sai Bhageerath, Y. V., Suresh Babu, A. V., Durga Rao, K. H. V., Sreenivas, K., and Chauhan, P.: Flood period estimation using multi-sensor satellite data: Case study on Punjab floods 2023, *J. Earth Syst. Sci.*, 134, 43, <https://doi.org/10.1007/s12040-024-02499-6>, 2025.
- Smith-Tripp, A., White, J. C., Wulder, M. A., and Coops, N. C.: Landsat assessment of variable spectral recovery linked to post-fire forest structure in dry sub-boreal forests, *ISPRS J. Photogramm.*, 208, 121–135, 2024.
- Turnbull, B. W.: The empirical distribution function with arbitrarily grouped, censored and truncated data, *J. R. Stat. Soc. B*, 38, 290–295, 1976.
- USDA Forest Service and U.S. Geological Survey: Monitoring Trends in Burn Severity (MTBS), <https://www.mtbs.gov>, last access: 30 November 2024.
- Verbesselt, J., Hyndman, R., Newnham, G., and Culvenor, D.: Detecting trend and seasonal changes in satellite image time series, *Remote Sens. Environ.*, 114, 106–115, <https://doi.org/10.1016/j.rse.2009.08.014>, 2010.
- White, J. C., Wulder, M. A., Hermosilla, T., Coops, N. C., and Hobart, G. W.: A Landsat time-series approach for characterizing post-fire forest recovery, *Remote Sens. Environ.*, 268, 112761, <https://doi.org/10.1016/j.rse.2021.112761>, 2022.