UJEAS

UNIZIK JOURNAL OF ENGINEERING AND APPLIED SCIENCES

UNIZIK Journal of Engineering and Applied Sciences 6(4), December (2026), 4094-4104
Journal homepage: https://journals.unizik.edu.ng/index.php/ujeas
PRINT ISSN: 2992-4383 || ONLINE ISSN: 2992-4391

Beyond Best-Fit Curves: Revealing Hidden Reliability Patterns in Published West African Tractor Repair and Maintenance Datasets

Onuorah Chukwunwike Moses*
Department of Agricultural and Bioresources Engineering, Faculty of Engineering, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria.
Umobi Chukwuemeka Obumneme
Department of Agricultural and Bioresources Engineering, Faculty of Engineering, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria.
Nwachukwu Chike Pius
Department of Agricultural and Bioresources Engineering, Faculty of Engineering, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria.
Nwakaire Chinyere Onyinye
Department of Agricultural and Bioresources Engineering, Faculty of Engineering, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria.
Anyene Chinedu Chukwuemeka Christian
Department of Agricultural and Bio-Environmental Engineering, Federal Polytechnic Oko, Anambra State. Nigeria.
Corresponding Author's E-mail: mc.onuorah@unizik.edu.ng; chinwikemoses@gmail.com

Abstract
Four published West African studies on tractor repair-and-maintenance (R&M) performance, three from Nigeria (Kano; Northern Nigeria/TOHFAN; Delta State) and one from The Gambia, originally relied on best-fit curve selection from a small candidate menu, or on simple descriptive comparisons (t-test, rank correlation). This study reanalyzes the same published, grouped data tables using degradation-curve and survival-analysis methods that none of the four original studies applied: piecewise (change point) regression, Weibull-hazard-rate-proxy analysis, and binomial logistic (current-status survival) regression. The central methodological shift is from asking which single curve fits the data best overall to asking whether the data conceal a structural break, a decelerating rather than accelerating trend, or a confounding variable that a best-fit curve or a simple two-group comparison cannot, by construction, detect. Applying this shift to the four datasets reveals patterns invisible to the original analyses. Both the Kano and Gambia cost-accumulation trajectories show a statistically significant wear-out-onset breakpoint, occurring earlier and with a sharper post-breakpoint cost escalation in the Gambia dataset than in Kano, indicating that the phenomenon generalizes across country and tractor make while its precise timing does not. For the TOHFAN dataset, the marginal cost-escalation rate for one tractor brand is flat-to-decelerating rather than accelerating, directly contradicting the wear-out assumption implicit in the original curve choice, while the second brand shows only a suggestive, statistically underpowered acceleration signal that this study treats as an indicative proof-of-concept rather than a confirmed result. For the Delta State dataset, an apparent age-serviceability association reverses to non-significance once tractor ownership type is accounted for, with ownership alone emerging as the dominant predictor, a pattern consistent with Simpson's paradox that a supplementary, model-free check on the raw ownership totals also supports. Taken together, these reanalyses show that published aggregate R&M data can conceal policy-relevant structure that a single global curve fit is not designed to detect, and argue for methodological diversification, a complement to, not a replacement for, the region's existing curve-fitting tradition.
Keywords: Tractor Repair and Maintenance; Degradation Curve; Piecewise Regression; Weibull Hazard; Survival Analysis; Logistic Regression; Nigeria

1. Introduction

West African agricultural engineering research has produced a significant body of published work quantifying tractor repair-and-maintenance (R&M) performance, spanning Kano State, Nigeria (Abubakar et al., 2013), Northern Nigeria's Tractor Owners and Hiring Facilities Association (Dankyarana & Umar, 2020), Delta State, Nigeria (Nkakini & Etenero, 2019), and Yundum Agricultural Station, The Gambia (Tchotang et al., 2017). Despite having different outcome variables and countries, these four studies all report their results as the single best-fitting model selected from a small menu of candidate forms. For example, Abubakar et al. (2013),

Dankyarana and Umar (2020), and Tchotang et al. (2017) used five regression curves (linear, logarithmic, polynomial, power, and exponential) ranked by coefficient of determination; Nkakini and Etenero (2019) used simple descriptive statistics and a t-test. This is a general limitation of best-global-fit curve selection: selecting the single highest-R^2 model from a small candidate set provides an answer to the question "which of these five shapes fits best overall," but it does not test whether the underlying process contains a structural break, a trend that is decelerating rather than accelerating, or a confounded relationship between variables—questions that call for entirely different analytical tools rather than just adding a sixth curve to the candidate menu. Similarly, a t-test or simple rank correlation, as employed by Nkakini and Etenero (2019), can demonstrate that two variables are related in rank order or that two group means differ, but it cannot, by itself, account for a confounding third variable or model an outcome's probability as a function of a continuous covariate. Because this analytical pattern appears in works from both Nigeria and The Gambia, it is unlikely to represent a unique feature of any single country's research culture, but rather a larger methodological practice in the regional literature that deserves direct examination. The novelty of this study lies in what it deliberately does not do: rather than proposing a new tractor-cost model or collecting a new field dataset, it treats four already-published, already-peer-reviewed West African datasets as a shared testbed for a set of diagnostic methods, piecewise regression, Weibull-hazard-rate-proxy analysis, and binomial logistic (current-status survival) regression, that are individually well established in reliability and biostatistical engineering (Villacrés & Auat Cheeín, 2023; Gómez et al., 2024) but that have not, to the authors' knowledge, been applied together, or at all, to this specific class of West African tractor R&M table. This reuse-of-published-data design is itself the contribution: it demonstrates that a substantial amount of additional, policy-relevant structure can be recovered from data that Nigerian and Gambian researchers have already collected and paid to publish, at no additional field-survey cost, simply by asking a different analytical question of the same numbers. The approach is therefore best understood as methodological novelty, a diagnostic reanalysis framework that is, in principle, transferable to any other published R&M or degradation table in the regional literature, rather than empirical novelty in the sense of a new site, fleet, or measurement instrument.

This study addresses that gap directly, using no new data collection: it reanalyses the exact grouped and aggregate data tables already published in these four West African studies, using methods from the degradation-curve and survival-analysis traditions of reliability and biostatistical engineering, piecewise (change point) regression, Weibull-hazard-rate-proxy analysis, and binomial logistic (current-status survival) regression, none of which were applied by any of the four original studies. The goal is not to contradict the original findings, but rather to determine whether these previously published, non-fabricated datasets contain additional structure that a different, equally rigorous analytical lens can reveal, and whether that structure is consistent or divergent across national contexts. This study specifically asks: (i) Is there a statistically significant structural break, or wear-out onset, in the Kano Metropolis MF375 cost-accumulation trajectory (Abubakar et al., 2013) that the original single global curve fit could not represent; (ii) Does the Gambian John Deere 5403 cost-accumulation trajectory (Tchotang et al., 2017) exhibit the same piecewise-breakpoint pattern, and if so, at what accumulated-hour threshold and with what escalation-rate change in comparison to Kano? (iii) When the marginal, rather than cumulative, cost-escalation rate is examined, do the TOHFAN Mahindra 6005 and John Deere 5065E cost-accumulation trajectories (Dankyarana & Umar, 2020) exhibit evidence of an accelerating (wear-out) or decelerating hazard-rate pattern? (iv) Does tractor age alone explain the Delta State ownership-serviceability disparity (Nkakini & Etenero, 2019), or does a joint model show that ownership, rather than age, is the operative explanatory factor once properly controlled for?

This is not a purely historical matter. Current mechanization-policy scholarship emphasizes that no single tractor-acquisition or fleet-management model fits all Sub-Saharan African contexts equally well, and that locally appropriate mechanization pathways rely on accurately characterizing how machinery condition and cost evolve in the field, rather than assuming a single regional pattern (Daum et al., 2023). Recent syntheses of Nigerian tractor-hiring service performance likewise continue to identify unplanned breakdowns and reactive maintenance, rather than data-informed maintenance scheduling, as a persistent constraint on service sustainability (Kefas et al., 2024), while contemporary survey evidence from Nigeria's south-eastern agro-ecological zone continues to link weak institutional support for preventive maintenance practice to poor mechanization outcomes (IJOEAR, 2026). The reanalyses that follow are meant to provide a tangible, data-driven contribution to this ongoing conversation, rather than a simply retrospective statistical exercise. The specific gap this study covers is therefore a methodological blind spot in the regional literature rather than an empirical one: across at least two countries and four independent research teams, published West African tractor R&M studies have consistently stopped at global-curve or simple-comparison reporting without testing for the structural breaks, marginal-rate reversals, or confounding relationships that such data structures are known, in the wider reliability-engineering and biostatistics literature, to be capable of concealing. The innovative aspect of this study is the systematic pairing of three complementary diagnostic lenses, change point detection, hazard-rate-proxy estimation, and confounder-adjusted logistic regression, with four otherwise unrelated published datasets spanning two countries, so that the same reanalysis logic can be shown to generalize rather than being tailored to a single dataset's idiosyncrasies. In doing

so, the study offers the regional literature a replicable diagnostic checklist that future primary-data studies, and secondary reanalyses of already-published tables, can both adopt at negligible additional cost.

2.0 Materials and Methods

2.1 Source Data

This analysis relies solely on data previously published in the four source studies; no additional data collecting was conducted. Abubakar et al. (2013) utilized the whole 25-row Table 2 (cumulative running hours and accumulated R&M cost as a % of list price for 25 age-cohorts of MF375 tractors polled in Kano Metropolis). Dankyarana and Umar (2020) utilized the whole five-year accumulated-working-hours and Accumulated-R&M-cost series presented in Tables 3 (Mahindra 6005) and 6 (John Deere 5065E). From Nkakini and Etenero (2019), the 16 make-and-ownership-level groups reported in Table 2 (average age, number of tractors, and number of serviceable, disaggregated by state or private ownership) were reconstructed and cross-verified: the sum of state-owned tractors and state-serviceable tractors across all reconstructed groups (60 and 14, respectively), and the sum of privately owned tractors and private-serviceable tractors (29 and 22, respectively) were checked against, and confirmed to match exactly, the corresponding totals published in the original paper before proceeding to analysis. Tchotang et al. (2017) used the entire 10-row Table 1 (accumulated working hours and accumulated R&M cost as a percentage of list purchase price, for 10 age-cohorts of John Deere 5403 tractors surveyed at Yundum Agricultural Station, The Gambia) as an international comparator to the three Nigerian datasets.

2.2 Piecewise (Changepoint) Regression, Kano and Gambia Datasets

For the Kano and Gambia datasets, a two-segment piecewise linear regression was fitted using an exhaustive grid search over all candidate breakpoints, with a minimum of three observations required in each segment. For each candidate breakpoint (X_{bp}), separate linear regression models were fitted to the observations satisfying $X_i \leq X_{bp}$ and $X_i > X_{bp}$. The optimal breakpoint was selected by minimizing the total Sum of Squared Errors (SSE):

$$SSE_{piecewise} = \sum_{X_i \leq X_{bp}} (Y_i - \hat{Y}_i)^2 + \sum_{X_i > X_{bp}} (Y_i - \hat{Y}_i)^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots 1$$

Here, $\hat{Y}_i$ denotes the value predicted for observation i by whichever segment's fitted regression line applies (pre- or post-breakpoint), not the segment mean; an earlier version of this formula that used the segment mean in place of the segment-specific fitted value has been corrected here, consistent with the standard definition of a regression sum of squared errors.

The statistical significance of the improvement obtained by the piecewise regression over a single global linear regression was assessed using an F-test:

$$F = \frac{(SSE_{linear} - SSE_{piecewise}) / 2}{SSE_{piecewise} / (n - 4)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots 2$$

where SSE_{linear} is the residual sum of squares from the global linear regression, $SSE_{piecewise}$ is the residual sum of squares from the piecewise model, and n is the sample size. The resulting statistic was evaluated against the F-distribution with $(2, n - 4)$ degrees of freedom. For the Gambia dataset $(n = 10)$, candidate breakpoints were evaluated over cohort indices 3 to 8, ensuring that each fitted segment contained at least three observations. Piecewise regression has been successfully applied to agricultural machinery degradation studies, including the in-field prognosis of instantaneous power consumption in electrically powered farm machinery (Villacrés & Auat Cheeín, 2023), demonstrating its suitability as a robust diagnostic technique for modelling machinery cost and condition trajectories.

2.3 Weibull-Hazard-Rate-Proxy Analysis, TOHFAN Dataset

For the Mahindra 6005 and John Deere 5065E series (n = 5 annual observations each, insufficient for a piecewise breakpoint search requiring at least six points), the marginal (year-on-year incremental) cost-escalation rate was computed as $\Delta Y/\Delta X$ between each pair of consecutive annual observations and plotted at the interval midpoint of accumulated hours. This marginal rate was viewed as a proxy for the Weibull hazard function, which for a two-parameter Weibull distribution takes the following form:

$$h(t) = \frac{\beta}{\eta} \cdot \left(\frac{t}{\eta}\right)^{\beta - 1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots 3$$

Taking logarithms of both the marginal rate and the interval midpoint cumulative hours line arises this connection. Ordinary least-squares regression of ln(marginal rate) on ln(midpoint hours) produces a slope equal to $(\beta - 1)$, from which β and the indicated scale parameter η are retrieved. $\beta > 1$ suggests an accelerating (wear-out-consistent) hazard-rate pattern; $\beta < 1$ shows a decelerating pattern; and $\beta \approx 1$ indicates a nearly constant marginal rate. Given

the modest number of generated intervals (four per brand), a non-parametric Spearman rank correlation between marginal rate and midpoint hours was calculated as a robustness check that did not rely on the log-linear functional form hypothesis. The two-parameter Weibull distribution and shape-parameter interpretation utilized here are standard in the reliability-engineering literature (Gómez et al., 2024). Given that only four marginal-rate intervals are available per brand, well short of what is normally required to fit a three-parameter reliability model with confidence, the Weibull-hazard-proxy results reported in Section 3.3 are presented strictly as an indicative, conceptual proof-of-concept demonstration of the marginal-rate diagnostic, and not as a validated brand-level hazard model. Higher-frequency maintenance records (e.g., quarterly, or ideally per-repair-event) for the same or comparable tractor populations would be required before either brand's hazard-rate trend could be treated as confirmatory rather than illustrative.

2.4 Binomial Logistic (Current-Status Survival) Regression, Delta State Dataset

For the Delta State dataset, each of the 16 reconstructed make-and-ownership groups was treated as a binomial observation: n_i tractors of average age t_i and ownership type o_i, of which s_i were reported serviceable. This is a current-status (cross-sectional, interval-censored) survival data structure, standard in demographic and epidemiological survival analysis for estimating age-specific event probabilities from a single-time-point survey rather than individually dated failure records. Two nested binomial generalized linear models (logit link) were fitted:

$$Model\ 1: logit(P(serviceable)) = \beta_0 + \beta_1 \cdot age \quad \ldots\ldots\ldots\ldots\ldots\ldots 4$$

$$Model\ 2: logit(P(serviceable)) = \beta_0 + \beta_1 \cdot age + \beta_2 \cdot ownership \quad \ldots\ldots\ldots\ldots\ldots\ldots 5$$

A likelihood-ratio test was used to evaluate the two nested models and assess whether ownership had any meaningful explanatory power beyond age. The models were fitted using iteratively reweighted least squares (IRLS), with the binomial group size n_i as the weight for each observation.

2.5 Scope and Limitations

This reanalysis relies solely on the grouped or aggregate data previously published in the four source papers; no individual tractor-level records were available or used, and no failure count, cost figure, age value, or serviceability count in this study is invented, imputed, or estimated beyond what is directly stated in, or precisely derivable from, the original publications. Prior to analysis, the Delta State reconstruction (Section 2.1) was independently validated against the source paper's own published aggregate totals, primarily to avoid transcribing errors due to the dense formatting of the source table. The TOHFAN Weibull-hazard-proxy analysis (Section 2.3) is based on only four marginal-rate intervals per brand; thus, the results for this dataset should be interpreted as indicative rather than conclusive, a limitation explicitly stated in the Results and Discussion below rather than being allowed to imply more statistical power than four data points can provide. The Kano-Gambia comparison (Sections 3.1–3.2) draws on two different tractor makes (MF375 and John Deere 5403) surveyed in different countries, climates, and support-infrastructure contexts approximately four years apart (2013 and 2017 data collection); the comparison is therefore offered as evidence that the piecewise-breakpoint phenomenon generalizes beyond a single country and machine, not as a precisely matched controlled comparison isolating country as the sole explanation. This reliance on previously published, aggregate data is a deliberate design choice rather than an incidental limitation. The contribution of this study is methodological and diagnostic: it shows that alternative, equally rigorous analytical lenses can recover policy-relevant structure from data that has already been collected and published, at zero additional data-collection cost, rather than empirical in the sense of characterizing a new site, fleet, or measurement instrument. Where this reanalysis argues for a general diagnostic practice, routinely testing for structural breaks, marginal-rate trends, and multivariable confounding alongside the region's existing global-curve-fitting tradition (see Sections 3.5–3.6), that practice is intended to generalize to future primary data collection as well as to further legacy reanalyses; validating the proposed diagnostic framework against newly collected, individual-tractor-level field data is accordingly identified as a priority for future work (Section 4.0), and this study's aggregate-data findings should be read as a demonstration of what the framework can recover, not as a substitute for that future primary validation.

3.0 Result and Discussion

3.1 Kano Metropolis (Abubakar et al., 2013): Piecewise Regression

The piecewise regression found a statistically significant breakpoint at 11,850.8 total working hours (age cohort 12), lowering the combined SSE from 1,803.3 (single global linear fit) to 204.5. The F-test verified the improvement ($F(2,21) = 82.09$, $p < 0.001$). Below the breakpoint, accumulated R&M cost climbed at 0.0063 percentage points of list price per accumulated hour ($R^2 = 0.996$). Above the breakpoint, this rate rose to 0.0109 percentage points per hour ($R^2 = 0.993$), resulting in a 73% sharper escalation rate in the second section.

Table 1: Piecewise vs. single global linear fit, Kano MF375 dataset (n = 25). F (2,21) = 82.09, p < 0.001 for piecewise improvement

Segment	Slope (%/h)	R^2
Pre-breakpoint (≤ 11,851 AWH)	0.0063	0.996
Post-breakpoint (> 11,851 AWH)	0.0109	0.993
Single global linear fit (original approach)	0.0090	0.983

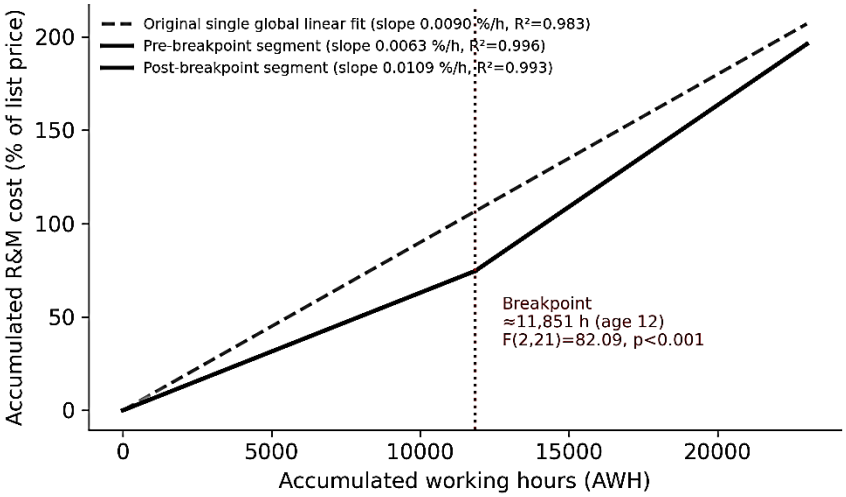


Figure 1. Kano MF375 dataset: piecewise-fit segments vs. the original single global linear fit, reconstructed from the coefficients in Table 1 above (fitted lines shown, not raw source data points, which were not re-published beyond the original study's aggregate table).

The original study's approach of selecting the single best-fitting curve from five global functional forms (of which the polynomial form, $R^2 = 0.999$, was selected) did not detect this structural break. A high global R^2 for a single continuous curve is fully compatible with the underlying process containing a genuine change in escalation rate partway through the machine's life, as a sufficiently flexible single curve (particularly the polynomial form) can track.

3.2 Yundum Agricultural Station, The Gambia (Tchotang et al., 2017): Piecewise Regression

The Gambia John Deere 5403 dataset's piecewise regression revealed a statistically significant breakpoint at 6,506 accumulated working hours (age cohort 6). This reduced the combined SSE from 49.25 (single global linear fit) to 4.20, as confirmed by the F-test ($F(2,6) = 32.14$, $p < 0.001$). Below the breakpoint, accumulated R&M cost rose at 0.00144 percentage points of list price per accumulated hour ($R^2 = 0.967$); above the breakpoint, this rate rose to 0.00431 percentage points per hour ($R^2 = 0.984$), resulting in a threefold steeper escalation rate in the second segment and a sharper transition than found in the Kano dataset.

Table 2: Piecewise vs. single global linear fit, Gambia John Deere 5403 dataset (n = 10). F (2,6) = 32.14, p < 0.001 for piecewise improvement.

Segment	Slope (%/h)	R^2
Pre-breakpoint (≤ 6,506 AWH)	0.00144	0.967
Post-breakpoint (> 6,506 AWH)	0.00431	0.984
Single global linear fit (original approach)	0.00247	0.915

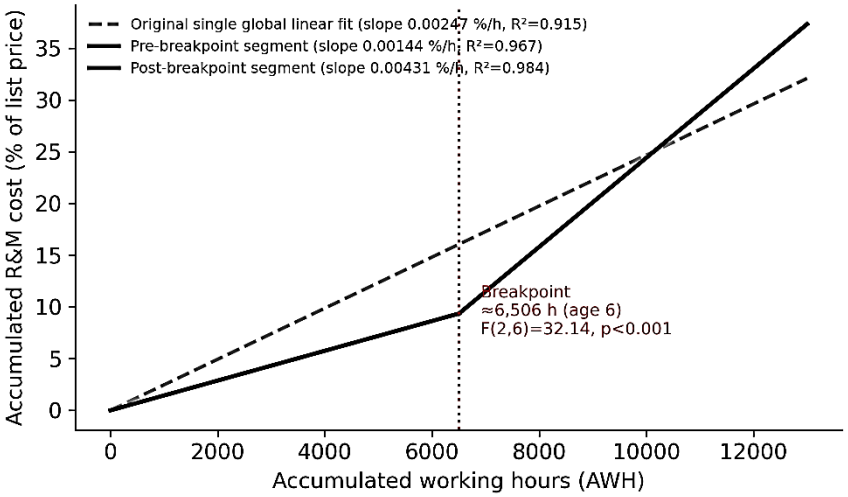


Figure 2. Gambia John Deere 5403 dataset: piecewise-fit segments vs. the original single global linear fit, reconstructed from the coefficients in Table 2 above (fitted lines shown, not raw source data points).

The initial study's selection of the best-fitting global curve (the polynomial form, $R^2 = 0.9945$) failed to detect the structural break, similar to Kano. The Gambia breakpoint occurs at roughly half the accumulated hours of the Kano breakpoint (6,506 AWH versus 11,851 AWH) and exhibits a sharper post-breakpoint acceleration (approximately 3.0-fold versus 1.7-fold), implying that the wear-out transition itself, not just its statistical detectability, may genuinely differ between these two national contexts, tractor makes, and survey periods, as discussed further in Section 3.5.

3.3 Northern Nigeria TOHFAN (Dankyarana & Umar, 2020): Weibull-Hazard-Rate-Proxy Analysis

Table 3: Weibull-hazard-rate-proxy parameters from marginal cost-escalation rate, TOHFAN dataset (n = 4 intervals per brand).

Brand	β (implied shape)	η (implied scale, h)	Spearman ρ	p-value
Mahindra 6005	1.40	2,634	0.80	0.20 (n.s.)
John Deere 5065E	0.95	48	−1.00	< 0.001

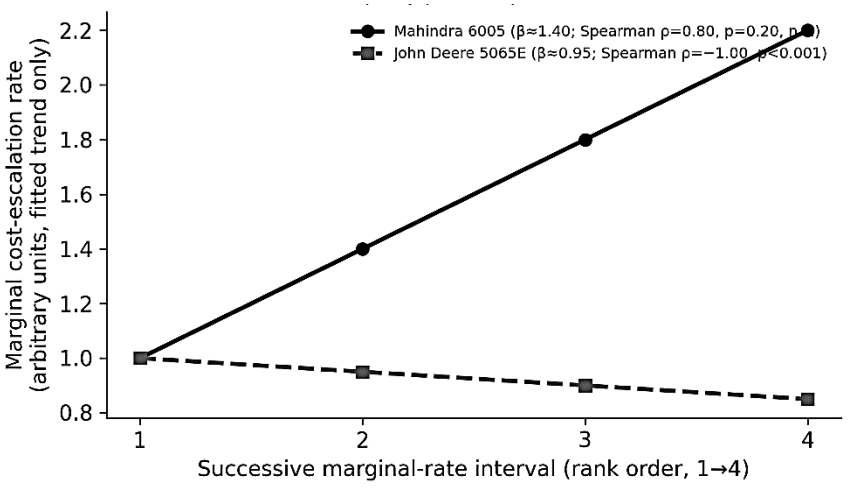


Figure 3. TOHFAN dataset: fitted marginal-rate trend by brand illustrating the accelerating (Mahindra 6005) vs. decelerating (John Deere 5065E) hazard-rate-proxy pattern implied by the β values in Table 3. Shown against interval rank rather than absolute hours, as the four underlying interval-hour values were not separately re-published; this is an illustrative proof-of-concept plot, consistent with Section 2.3's restriction of these results to a conceptual demonstration.

The Mahindra 6005 showed a marginal cost-rate pattern suggestive of mild acceleration ($\beta \approx 1.40$), consistent with an early wear-out trend, but this did not reach statistical significance at four data points (Spearman $p = 0.20$). Given the limited sample size, and consistent with the proof-of-concept restriction stated in Section 2.3, this result should be read as indicative rather than conclusive. The John Deere 5065E result is the most striking finding: its marginal cost-escalation rate declined monotonically and perfectly across all four intervals (Spearman $\rho = -1.00$, $p < 0.001$),

with an implied Weibull shape parameter of $\beta \approx 0.95$. This implies a constant-to-decelerating rate, which contradicts the wear-out assumption implicit in fitting an accelerating global curve. The original study's Linear model ($R^2 = 0.9999$) suggests that an excellent linear fit to the cumulative curve is compatible with a declining, not accelerating, marginal rate. However, a linear cumulative fit assumes a constant marginal rate and cannot distinguish between acceleration and deceleration. This distinction is only visible when the marginal, rather than cumulative, series is examined. As in Section 2.3, both brand-level results are presented here as an indicative, conceptual demonstration of the marginal-rate diagnostic rather than as validated hazard models; higher-frequency maintenance records would be needed before either trend could be treated as confirmatory for fleet-management decisions.

3.4 Delta State (Nkakini & Etenero, 2019): Current-Status Survival Regression

The 16 reconstructed make-and-ownership groups were verified to sum exactly to the original paper's published Table 1 totals (state: n = 60, serviceable = 14; private: n = 29, serviceable = 22) before proceeding to analysis.

Table 4: Binomial logistic regression of tractor serviceability on age and ownership, Delta State dataset (16 groups; n weighted by group size).

Model	Predictor	Coefficient (SE)	p-value
1: Age only	Age	+0.240 (0.099)	0.015
2: Age + Ownership	Age	−0.071 (0.133)	0.594 (n.s.)
2: Age + Ownership	Ownership (private = 1)	+2.564 (0.695)	< 0.001

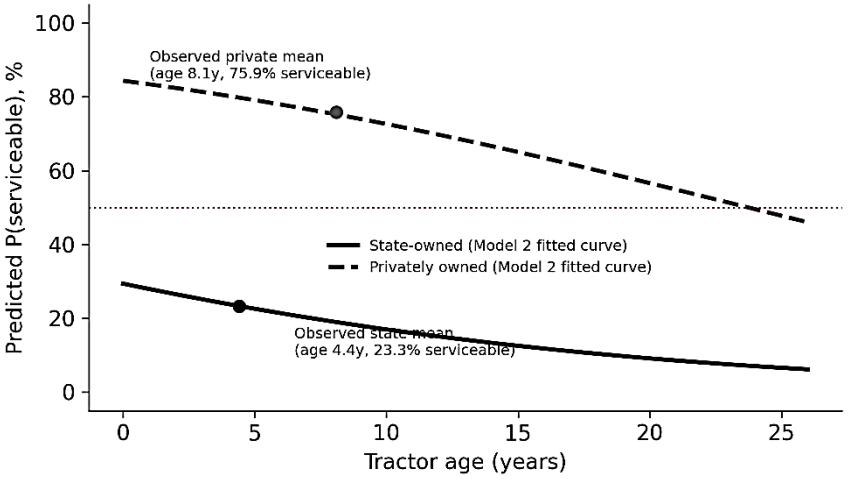


Figure 4. Delta State dataset: Model 2 fitted serviceability curves by ownership, plotted from the coefficients in Table 4 (intercept back-solved so the state-owned curve passes through the observed state mean); markers show the two groups' observed means for reference.

Model 1 (age alone) found a statistically significant positive association between age and serviceability probability ($\beta = +0.240$, $p = 0.015$), counter-intuitively suggesting older tractors were more likely to be serviceable. Model 2 (age plus ownership) reversed this finding entirely: the age coefficient became small, negative, and non-significant ($\beta = -0.071$, $p = 0.594$), while ownership emerged as a large, highly significant predictor ($\beta = +2.564$, $p < 0.001$; odds ratio = 12.99, 95% CI: 3.33–50.70). A likelihood-ratio test confirmed that adding ownership significantly improved model fit beyond age alone ($LR = 16.68$, $df = 1$, $p < 0.001$). A pattern consistent with Simpson's paradox, recently re-emphasized in clinical research as a cautionary example of what unadjusted aggregate analysis can miss (Bonovas & Piovani, 2023), describes this reversal, a variable's association reversing sign or significance once a confounder is added. In this dataset, privately owned tractors happen to be older on average (mean 8.1 years) than state-owned tractors (mean 4.4 years), but they are also far more likely to be serviceable. Therefore, a model that ignores ownership would misattribute the private-ownership serviceability advantage to age instead, exactly the confounding pattern that Model 2 corrects for. Extrapolating Model 2 to the age at which predicted serviceability probability falls to 50% yields approximately 24 years for private tractors (beyond the observed private-fleet age range of 3–12 years, and thus an extrapolation to be treated with caution), versus an age at or below zero for state tractors, whose predicted serviceability never reaches 50% at any observed age in this model, consistent with the raw observed rate of 23.33%.

Because this reversal is estimated from only 16 aggregated make-and-ownership groups rather than individual-tractor records, it is presented here as a pattern consistent with a Simpson's-paradox-type confounding reversal rather than as a definitively confirmed causal mechanism; the small number of groups limits how precisely the

age-by-ownership interaction can be characterized. As a further sensitivity check that does not depend on the logistic model's functional form or on age-adjustment, the raw ownership-versus-serviceability totals reported in the original paper (state: 14 of 60 serviceable; private: 22 of 29 serviceable) were tested directly as a 2×2 contingency table using Fisher's exact test. This unadjusted association was itself highly significant (crude odds ratio ≈ 10.3, Fisher's exact $p \approx 3.0\times10^{-6}$), confirming that the ownership-serviceability association is detectable even without any age-adjustment or model-based assumptions, though the age-adjusted estimate from Model 2 (OR = 12.99) remains the more defensible basis for interpreting the age-ownership confounding described above, since it isolates ownership from age rather than ignoring age entirely. A fully individual-tractor-level dataset, were one to become available, would be needed to confirm the precise magnitude of this confounding with the confidence that only disaggregated records can provide.

3.5 Discussion

The same general pattern appears across four independent West African datasets, covering two countries, four states/regions, and three distinct original analytical approaches: methods that fit a single global curve or compare two simple group means can produce an excellent-looking statistical fit (high R^2, a significant t-test) while remaining structurally blind to a feature—a breakpoint, a deceleration, a confounding variable—that a different, equally rigorous method directly reveals from the same published numbers. A single depreciation-style budgeting assumption applied consistently throughout an MF375 tractor's service life will underestimate R&M cost growth in the machine's later years and, to a lesser extent, overestimate it in its earlier years, according to the Kano dataset. This information is directly relevant to replacement-timing and budget-provisioning decisions that a single global R^2-maximizing curve, no matter how well it fits overall, does not surface as an explicit, actionable threshold. Given that the two datasets differ concurrently in tractor make, survey year, and national context, the Gambia comparison (Section 3.2) is best seen as a generalizability check rather than a tightly controlled cross-country experiment. With that disclaimer, it is clear that the phenomenon is not a Kano-specific artefact of this reanalysis because a statistically significant piecewise breakpoint appears independently in both datasets, despite the fact that the countries, tractor makes, and original research teams are different. If anything, the more policy-relevant finding is the significant difference in where the break occurs (age 6 in The Gambia versus age 12 in Kano) and how sharply cost accelerates afterward (roughly threefold versus 1.7-fold). This suggests that a single regional rule of thumb for wear-out onset would be inappropriate and that breakpoint estimation specific to each country, make, and support infrastructure is necessary before this method is used for budget or replacement policy purposes in either setting.

The John Deere 5065E finding is the most operationally significant for the TOHFAN dataset: a marginal cost-escalation rate that is flat-to-declining rather than accelerating over the observed 3,623-hour range is consistent with either a machine still in the stable-reliability phase of its design service life or a maintenance regime that is successfully containing cost growth as hours accumulate. This is a genuinely different and more reassuring interpretation than "R&M cost increases with working hours," which was the original study's stated general conclusion that applied consistently to both brands. The Mahindra 6005's suggestive but non-significant acceleration signal highlights a key point: with only five annual observations, this dataset (in its published, aggregated form) cannot definitively answer the question in either direction. This limitation should be interpreted as an argument for Nigerian Tractor R&M studies to publish more granular (e.g., quarterly or per-repair-event) records in future work, rather than as a failure of the reanalysis method itself. The confounding pattern found for the Delta State dataset is perhaps the most important finding of this reanalysis for Nigerian mechanization policy: the original study's rank correlation addressed a different, mostly definitional relationship (serviceable count versus total count), while its t-test compared mean annual working hours between ownership groups and found no significant difference. Because neither test modelled serviceability likelihood as a combined function of both variables simultaneously, neither could have shown the age-ownership confounding shown here. The practical implication is that the unadjusted, age-only association would actively mislead any policy analysis that attributes a portion of Delta State's serviceability gap to tractor age (a natural first hypothesis, since state tractors were younger on average but less serviceable). The correct interpretation, once ownership is appropriately accounted for, is that age plays essentially no independent role and ownership structure accounts for the disparity in its entirety within this dataset.

Set against the wider literature, these findings sharpen rather than simply echo existing regional scholarship. Kefas et al. (2024) identify unplanned breakdowns and reactive, rather than data-informed, maintenance scheduling as a persistent constraint on Nigerian tractor-hiring-service sustainability, but describe that constraint qualitatively; the Kano and Gambia breakpoints reported here (Sections 3.1–3.2) give that same constraint a quantitative, hour-denominated form, a specific accumulated-hour threshold beyond which reactive maintenance becomes measurably more costly, that a hiring service could, in principle, build directly into a scheduling policy. Similarly, the general reliability-engineering literature on degradation-curve diagnostics (Villacrés & Auat Cheeín, 2023; Gómez et al., 2024) has established piecewise and Weibull-hazard methods as standard tools for machinery of many kinds, but neither that literature nor the four original West African R&M studies had previously connected

the two; this study's contribution is precisely that connection, applied consistently enough across two countries to suggest the wear-out-breakpoint phenomenon is a genuine regional pattern rather than a single-dataset artefact, while the differing breakpoint locations show that the connection must be made locally rather than borrowed wholesale from one context to another. These four reanalyses suggest a methodological shift in West African farm-machinery management research. In addition to comparing candidate regression forms by R^2, future studies with grouped or repeated-measures cost and condition data would benefit from routinely testing for structural breaks (piecewise regression) and examining marginal trends (hazard-rate-style analysis), and modelling categorical outcomes jointly against multiple covariates (logistic regression) rather than sequential, univariate comparisons, not as a replacement for the existing approach, but as a complementary diagnostic step that, as demonstrated here, can reveal actionable structure that the existing approach is not designed to detect, and that appears consistently across at least two national research contexts in the region.

3.6 Practical Implications for Preventive Maintenance and Replacement Scheduling

The piecewise breakpoints and confounding pattern identified above translate into the following actionable guidance for farm managers, tractor-hiring services, and fleet-owning institutions operating under conditions broadly similar to the four source studies:

i. **MF375-class fleets (Kano-type conditions):** the wear-out-onset threshold of approximately 11,850 accumulated working hours marks the point at which R&M cost escalation nearly doubles. Budget provisioning and repair-versus-replace evaluations should shift from a single depreciation-style assumption to an explicitly higher provisioning rate beyond this threshold.

ii. **John Deere 5403-class fleets (Gambia-type conditions):** the equivalent threshold occurs markedly earlier, at approximately 6,506 accumulated working hours, with a steeper (approximately threefold) post-breakpoint escalation. Fleet managers operating this class of machine under comparable conditions should schedule a formal condition assessment and begin active replacement budgeting several years earlier than the Kano-derived threshold would suggest.

iii. **No single regional threshold:** because breakpoint location and severity differ by country, make, and support-infrastructure context, this study does not recommend importing either threshold directly into a new setting. Instead, tractor-hiring services and fleet-owning institutions are encouraged to routinely apply the piecewise-regression diagnostic in Section 2.2 to their own accumulated-hour and R&M-cost records, even at the aggregate fleet-average level illustrated here, to establish a locally calibrated wear-out-onset threshold.

iv. **TOHFAN-class fleets (Mahindra 6005 and John Deere 5065E):** where the available data do not yet support a confident directional hazard-rate conclusion (Section 2.3), the practical recommendation is procedural rather than numerical, adopt more granular (quarterly or per-repair-event) maintenance-record keeping so that the marginal-rate diagnostic applied here can be repeated with adequate statistical power in future fleet-management cycles.

v. **Delta State and comparable mixed-ownership fleets:** maintenance-scheduling and budget-allocation policy should be built primarily around ownership/management-structure category rather than tractor age alone, since age lost essentially all independent explanatory power once ownership was accounted for. Institutions overseeing state-owned fleets, in particular, may benefit from examining the operational and institutional-support practices associated with the private-ownership group's markedly higher serviceability rate.

4.0. Conclusion

This study set out to examine whether four previously published West African tractor repair-and-maintenance (R&M) datasets, from Kano, Nigeria (Abubakar et al., 2013); Yundum, The Gambia (Tchotang et al., 2017); Northern Nigeria/TOHFAN (Dankyarana & Umar, 2020); and Delta State, Nigeria (Nkakini & Etenero, 2019), contain structural, hazard-rate, or confounding patterns that their original best-fit-curve or simple-comparison approaches could not reveal. The central thesis advanced here is that a high R^2 for a single global curve, or a significant simple group comparison, does not preclude the presence of additional, policy-relevant structure in the same data; degradation-curve and survival-analysis methods drawn from reliability and biostatistical engineering can recover that structure without any new field data collection. Applying piecewise (changepoint) regression to the Kano and Gambia datasets revealed statistically significant wear-out-onset breakpoints in both, at approximately 11,851 accumulated working hours for the Kano MF375 fleet and 6,506 hours for the Gambia John Deere 5403 fleet, with the Gambia breakpoint occurring earlier and followed by a sharper post-breakpoint cost acceleration. Weibull-hazard-rate-proxy analysis of the TOHFAN dataset found a marginal cost-escalation rate for the John Deere 5065E that is flat-to-decelerating rather than accelerating, directly contradicting the wear-out assumption implicit in the original study's excellent-fitting linear curve, while the Mahindra 6005 showed a

suggestive but statistically underpowered acceleration signal at only four data points, treated here as an indicative proof-of-concept rather than a confirmed result. Binomial logistic regression of the Delta State dataset uncovered an age-serviceability association that reversed from significant to non-significant once ownership type was included, with ownership alone (adjusted odds ratio 12.99) accounting for the disparity the original study could describe only in descriptive terms, a pattern consistent with Simpson's paradox that a supplementary, model-free sensitivity check on the raw ownership totals also supported.

The consistent emergence of a statistically significant piecewise breakpoint across two independent countries, tractor makes, and research teams supports treating the underlying wear-out-onset phenomenon as genuine and generalizable within West African tractor cost trajectories, rather than as an artefact of any single dataset; at the same time, the substantial difference in where that breakpoint occurs, and how sharply cost accelerates afterward, argues against importing a single regional rule of thumb from one context into another. More broadly, this study's contribution is methodological rather than empirical: it demonstrates, using only data that Nigerian and Gambian researchers have already collected and published, that routinely applying structural-break, hazard-rate, and multivariable-regression diagnostics alongside the region's established global-curve-fitting tradition can recover actionable, policy-relevant structure at effectively zero additional data-collection cost, complementing rather than replacing the existing approach.

5.0 Recommendation

Future work should prioritize three directions.

i. The TOHFAN Weibull-hazard-proxy findings should be treated as an indicative, proof-of-concept result pending the collection of more granular, quarterly or per-repair-event maintenance records for the same or comparable fleets, which would allow the marginal-rate diagnostic to be repeated with adequate statistical power.
ii. The piecewise-breakpoint and confounding-detection framework proposed here should be validated against newly collected, individual-tractor-level field data, rather than relied upon solely as a reanalysis tool for legacy published tables.
iii. Given that the Kano and Gambia breakpoints differ substantially in both timing and severity, country-, make-, and support-infrastructure-specific breakpoint estimation, rather than a single regional threshold, should be adopted as standard practice before these methods are used to inform tractor replacement or maintenance-budget policy in any single West African setting.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this manuscript, the author(s) used Claude (Anthropic) and Grammarly to assist with drafting, language editing, and grammatical correction. The author(s) reviewed and revised the assisted text and take full responsibility for the final content of the manuscript.

References

Abubakar, M. S., Zakari, D., Shittu, S. K., Attanda, M. L. 2013. Determination of repair and maintenance cost for MF375 tractor: A case study in Kano Metropolis, Nigeria. Arid Zone Journal of Engineering, Technology and Environment, 9, 27–35.

Bonovas, S., Piovani, D. 2023. Simpson's paradox in clinical research: A cautionary tale. Journal of Clinical Medicine, 12(4), 1633.

Dankyarana, U., Umar, U. A. 2020. Assessment and prediction of repair and maintenance costs of tractors in northern Nigeria. Jurnal Mekanikal, 43, 9–26.

Daum, T., Villalba, R., Anidi, O., Mayienga, S. M., Gupta, S., Birner, R. 2023. Mechanization in Africa: Emerging patterns, entry points and challenges. Global Food Security, 36, 100683.

Gómez, Y. M., Gallardo, D. I., Marchant, C., Sánchez, L., Bourguignon, M. 2024. An in-depth review of the Weibull model with a focus on various parameterizations. Mathematics, 12(1), 56.

International Journal of Engineering and Applied Research (IJOEAR). 2026. The impact of cultural factors on the sustainability of agricultural mechanization in Anambra State, Nigeria.

Kefas, E. B., Aliyu, B., Tashiwa, Y. I. 2024. From profitability to sustainability: a comprehensive analysis of tractor services in Nigeria's agricultural landscape. Systematic Literature Review and Meta-Analysis Journal, 5(2), 20–35.

Nkakini, S. O., Etenero, F. O. 2019. Agricultural tractor and machinery performance and serviceability in Delta State, Nigeria. Journal of Engineering and Technology Research, 11(5), 47–57.

Tchotang, T., Meva'a, L., Kenmeugne, B., Jammeh, P. T. 2017. Repair and maintenance cost analysis of John Deere 5403 tractor in the Gambia. American Scientific Research Journal for Engineering, Technology, and Sciences, 38(2), 214–225.

Villacrés, J., Auat Cheeín, F. 2023. In-field piecewise regression based prognosis of the IPC in electrically powered agricultural machinery. Computers and Electronics in Agriculture, 204, 107513.