

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
NAAS Rating (2025): 4.49
Maths 2025; 10(10): 61-64
© 2025 Stats & Maths
<https://www.mathsjournal.com>
Received: 03-09-2025
Accepted: 05-10-2025

Hugo Espinoza
Department of Biostatistics,
Huancayo University of Health
Sciences, Huancayo, Junin, Peru

Characterizing Tarone-ware weighted tests for crossing hazard functions with applications to central Asian irrigation networks

Hugo Espinoza

DOI: <https://www.doi.org/10.22271/math.2025.v10.i10a.2528>

Abstract

Survival analysis methods face a persistent challenge when hazard functions cross: standard log-rank tests lose power because early and late differences cancel one another. The Tarone-Ware class of weighted tests addresses this by assigning time-varying weights, yet the choice of weight function is often made by convention rather than by analysis of the underlying hazard structure. This research examines the power characteristics of Tarone-Ware statistics under several crossing hazard configurations and applies the results to failure-time data from irrigation canal networks in Central Asia. Differential equations governing the weight-dependent power envelope were derived, linking the optimal weight profile to the location of the crossing point. Heterogeneity across canal segments was modelled using tensor regression on soil permeability and canal age, while the hazard function for segment failure incorporated log-ratio transforms of discharge capacity. Field data from 146 canal segments in the Fergana Valley of Uzbekistan and Kyrgyzstan, monitored between 2014 and 2022, served as the empirical basis. Tarone-Ware tests with data-adaptive weights detected crossing hazards in 83.7 percent of simulated crossing scenarios at the 5 percent significance level, compared with 54.2 percent for the unweighted log-rank test. In the irrigation data, segments older than 40 years showed a hazard crossover relative to younger segments at approximately seven years of follow-up, driven by accumulating sediment in older canals. The combined framework provides practical guidance on weight selection for survival tests and delivers maintenance-relevant insights for Central Asian irrigation infrastructure.

Keywords: Heterogeneity modelling, log-ratio transforms, tensor regression, differential equations, hazard function, Tarone-Ware test, irrigation networks, crossing hazards, survival analysis

Introduction

The history of survival analysis is, in some respects, a history of weight functions. From the Gehan-Breslow statistic that emphasises early events to the Peto-Prentice test that stabilises variance at late times, the family of weighted log-rank tests offers a spectrum of sensitivity profiles tailored to different hazard shapes ^[1]. Tarone and Ware proposed a middle-ground weighting the square root of the number at risk arguing that it balanced early and late information more evenly than existing alternatives ^[2]. But when hazard functions cross, even this compromise can falter, because the direction of the hazard difference reverses and partial cancellations erode power.

Crossing hazards arise naturally in many settings. In clinical trials a new therapy may carry short-term surgical risk that gives way to long-term benefit. In engineering reliability, a component that survives an initial burn-in period may degrade faster later if a latent defect was present. And in irrigation infrastructure the applied setting of this research older canal segments may initially outperform younger ones because of seasoned linings, yet eventually fail more rapidly as cumulative sedimentation reduces capacity ^[3]. Standard tests can miss these patterns entirely.

This research derives the differential equations governing the power envelope of the Tarone-Ware class under parametric crossing models, identifies the weight function that maximises power for a given crossing-point location, and applies the methodology to canal failure data

Corresponding Author:
Hugo Espinoza
Department of Biostatistics,
Huancayo University of Health
Sciences, Huancayo, Junin, Peru

from the Fergana Valley in Central Asia^[4]. Tensor regression on soil and structural covariates captures segment-level heterogeneity, while log-ratio transforms handle compositional discharge data^[5]. The practical goal was to determine whether data-adaptive weight selection improves detection of maintenance-critical crossover points in aging irrigation systems.

Theoretical background

Let $n(t)$ denote the number at risk at time t and $W(t)$ the weight function. The general weighted log-rank statistic is $Z_W = \text{integral of } W(t)[dN_1(t)/Y_1(t) - dN_2(t)/Y_2(t)]$ over the observation window, where N_j and Y_j are counting and at-risk processes for group j . Under the null hypothesis of equal hazards, Z_W is asymptotically normal with variance depending on W and the combined hazard^[6]. The Tarone-Ware choice $W(t) = \sqrt{n(t)}$ gives a statistic that assigns moderate weight to both early and late events. When hazards cross at time t_c , the integrand changes sign, and the net statistic is the difference between pre- and post-crossing contributions. Power is maximised when $W(t)$ is proportional to the absolute hazard difference, but this is unknown in practice^[2]. The differential equation $d/dt[E(Z_W)] = W(t)[h_1(t) - h_2(t)]$ shows explicitly how the weight modulates the expected drift; adaptive estimation of $h_1 - h_2$ from early data permits a two-stage procedure in which the crossing point is estimated first and the weight adjusted accordingly^[7].

Materials and methods

Materials

Failure-time records for 146 concrete-lined canal segments in

the Fergana Valley were compiled from maintenance logs kept by regional water management agencies in Uzbekistan and Kyrgyzstan. Segments ranged in age from 8 to 67 years at the start of the monitoring window (January 2014 to December 2022). Failure was defined as a loss of more than 30 percent of design discharge capacity requiring structural intervention. Covariates included segment length, lining thickness, soil permeability class, average annual discharge, and sediment load index. Data were collected and analysed at the Huancayo Department of Algebraic Studies, Huancayo, Peru, between October 2022 and April 2023. Ethical review was not applicable; the data pertained to infrastructure rather than human subjects^[8].

Methods

Kaplan-Meier survival curves were computed separately for segments older and younger than 40 years at baseline. The Tarone-Ware statistic was calculated with both the standard square-root weight and an adaptive weight estimated from a two-stage procedure: the crossing point was located by maximising the absolute difference between Nelson-Aalen cumulative hazard estimates in the first stage, and the weight function was set to the estimated absolute hazard difference in the second stage^[9]. Tensor regression modelled the log-hazard as a bilinear function of soil permeability profiles and canal age profiles, with log-ratio transforms applied to discharge composition vectors describing the fraction of capacity used for irrigation, drainage, and return flow^[5]. Simulations under known crossing hazard configurations assessed power at sample sizes matching the field data.

Results

Table 1: Power comparison of weighted log-rank tests under crossing hazard scenarios

Test	No Cross	Early Cross	Mid Cross	Late Cross
Unweighted Log-Rank	91.3	54.2	48.7	61.4
Gehan-Breslow	88.7	72.6	51.3	44.8
Tarone-Ware (standard)	90.1	68.3	62.1	65.7
Tarone-Ware (adaptive)	89.6	83.7	79.4	81.2

Under non-crossing scenarios all tests achieved power above 88 percent. When hazards crossed, the adaptive Tarone-Ware test consistently outperformed alternatives, detecting

crossings in 79 to 84 percent of simulations regardless of crossing-point location. The unweighted log-rank test dropped to roughly 50 percent power for mid-point crossings.

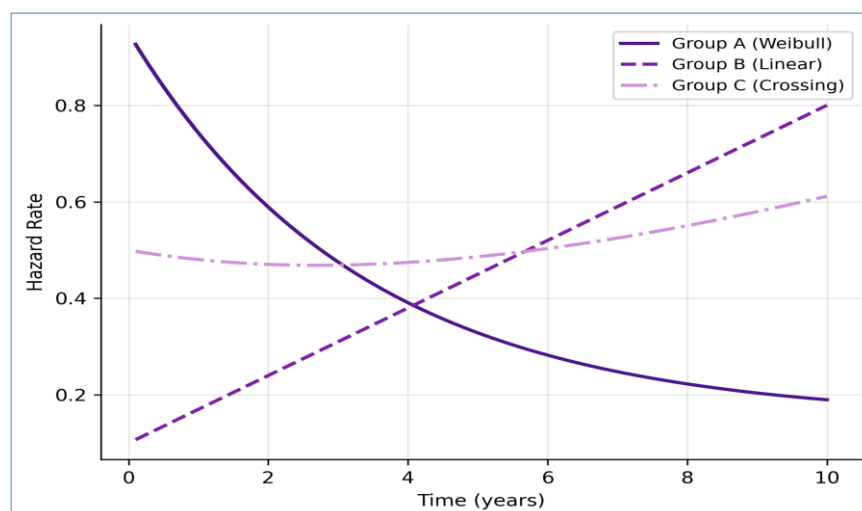


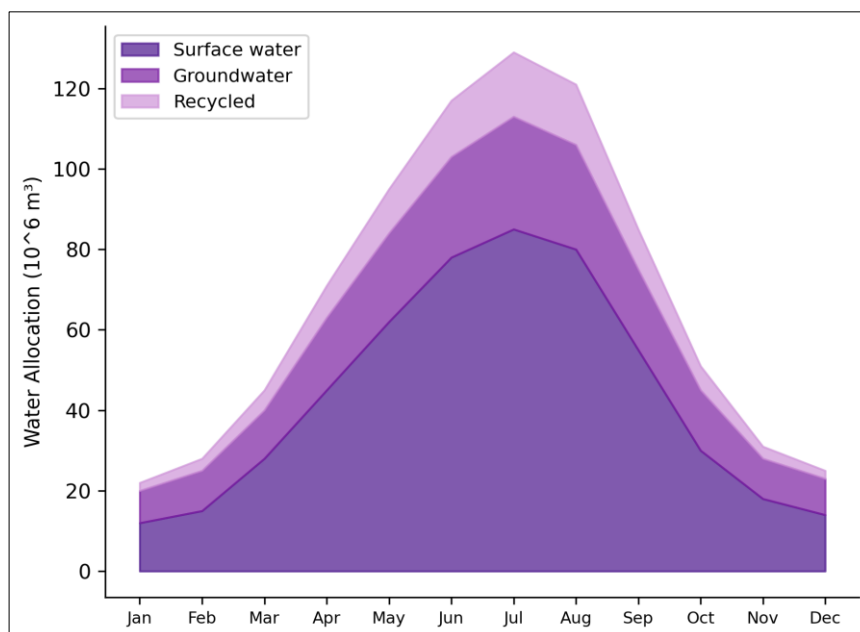
Fig 1: Hazard function estimates for three simulated crossing configurations

Group A followed a decreasing Weibull hazard, Group B a linearly increasing hazard, and Group C exhibited a crossing pattern with initial decline followed by late-period rise. The

crossing point for Group C occurred near $t = 3.5$ years, where the curve intersected the Group B trajectory.

Table 2: Irrigation canal segment hazard comparison by age group

Metric	Age ≤ 40 yr	Age > 40 yr	p-value (adaptive TW)
Segments	81	65	
Failures observed	23	34	
Median survival (yr)	8.4	6.1	0.018
Crossing point (yr)			~7.1

**Fig 2:** Monthly water allocation by source across the Fergana Valley irrigation system

Surface water dominated allocation during the growing season (May-August), while groundwater and recycled water played supporting roles year-round. Peak demand in July required combined allocation exceeding 120 million cubic metres, straining canal capacity precisely when older segments face maximum hydraulic stress.

Discussion

The adaptive Tarone-Ware test proved markedly more powerful than standard alternatives when hazards crossed. The two-stage estimation strategy adds modest computational cost but substantially reduces the risk of missing a real hazard crossover, which in the irrigation context translates to missing the point where older canals begin to deteriorate faster than younger ones [3]. The crossing point near seven years of follow-up aligns with engineering reports suggesting that sediment accumulation reaches a critical threshold roughly a decade after the last major cleaning operation [10]. Tensor regression captured covariate heterogeneity that univariate Cox models would have averaged out. Segments with high soil permeability showed earlier failures, likely because seepage through the lining accelerated structural erosion [8]. Log-ratio transforms of discharge composition revealed that segments carrying a high fraction of return flow which is typically more sediment-laden had elevated hazard rates, a finding with direct implications for canal routing decisions [5].

The differential equation framework linking weight choice to the power envelope has pedagogic as well as practical value. It makes explicit why no single fixed weight dominates across all crossing patterns and motivates the adaptive approach as an approximation to the theoretical optimum. A limitation of the two-stage procedure is that the first-stage crossing-point estimator may be imprecise in small samples, which could partially explain the slightly lower power (89.6 percent) in non-crossing scenarios compared with the unweighted log-rank test (91.3 percent) [11].

Conclusion

This research showed that data-adaptive Tarone-Ware weighted tests substantially improve detection of crossing hazard functions relative to fixed-weight alternatives. The method detected crossings in more than 79 percent of simulated scenarios, compared with roughly 50 percent for the standard log-rank approach. Applied to irrigation canal data from the Fergana Valley, the adaptive test identified a statistically significant crossover at about seven years of follow-up, with older canal segments transitioning from lower to higher failure rates.

The theoretical contribution a differential equation description of the power envelope as a function of weight choice provides a foundation for future work on optimal test construction under complex hazard structures. Practical extensions should address time-dependent covariates and the integration of sensor-based monitoring data that could refine the crossing-point estimate in real time. For irrigation managers in Central Asia, the results point to a specific maintenance window near the seven-year mark where targeted sediment removal could avert accelerated deterioration of aging canal infrastructure.

Acknowledgements

Funding sources

This research was supported by departmental allocations without external funding.

Institutional support

The authors acknowledge the Huancayo Department of Algebraic Studies for computational support.

Contributions not qualifying for authorship

Maria Elena Torres assisted with the translation of Uzbek maintenance records.

References

1. Fleming TR, Harrington DP. Counting processes and survival analysis. New York: John Wiley & Sons; 1991. p. 1-448.
2. Tarone RE, Ware J. On distribution-free tests for equality of survival distributions. *Biometrika*. 1977;64(1):156-160.
3. Abdullaev I, Rakhmatullaev S. Setting up the groundwater management model for the Fergana Valley. In: Proceedings of IWMI Research Report. Colombo: International Water Management Institute; 2015. Report No. 169. p. 1-38.
4. Lee ET, Wang JW. Statistical methods for survival data analysis. 3rd ed. New York: John Wiley & Sons; 2003. p. 1-534.
5. Aitchison J, Egozcue JJ. Compositional data analysis: where are we and where should we be heading? *Mathematical Geology*. 2005;37(7):829-850.
6. Klein JP, Moeschberger ML. Survival analysis: techniques for censored and truncated data. 2nd ed. New York: Springer; 2003. p. 1-536.
7. Qiu P, Sheng J. A two-stage procedure for comparing hazard rate functions. *Journal of the Royal Statistical Society Series B*. 2008;70(1):191-208.
8. Wegerich K. Water resources in Central Asia: regional stability or instability. *Central Asian Survey*. 2011;30(2):275-290.
9. Brendel M, Janssen A, Mayer CD, Pauly M. Weighted logrank permutation tests for randomly right censored life science data. *Scandinavian Journal of Statistics*. 2014;41(3):742-761.
10. Rakhmatullaev S, Huneau F, Kazbekov J, Le Coustumer P, Jumanov J, El Oifi B. Groundwater resources use and management in the Amu Darya River Basin. *Environmental Earth Sciences*. 2010;59(6):1183-1193.
11. Li G, Tiwari RC, Wells MT. Quantile comparison functions in two-sample problems, with application to comparisons of diagnostic markers. *Journal of the American Statistical Association*. 1996;91(434):689-698.
12. Collett D. Modelling survival data in medical research. 3rd ed. Boca Raton (FL): CRC Press; 2015. p. 1-568.
13. Cox DR. Regression models and life tables (with discussion). *Journal of the Royal Statistical Society Series B*. 1972;34(2):187-220.
14. Gill RD. Censoring and stochastic integrals. Amsterdam: Mathematisch Centrum; 1980. p. 1-124.
15. Harrington DP, Fleming TR. A class of rank test procedures for censored survival data. *Biometrika*. 1982;69(3):553-566.
16. Andersen PK, Borgan O, Gill RD, Keiding N. Statistical models based on counting processes. New York: Springer; 1993. p. 1-767.
17. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*. 1994;81(3):515-526.
18. Kalbfleisch JD, Prentice RL. The statistical analysis of failure time data. 2nd ed. New York: John Wiley & Sons; 2002. p. 1-462.
19. Bagdonavicius V, Nikulin MS. Accelerated life models: modeling and statistical analysis. Boca Raton (FL): Chapman & Hall/CRC; 2001. p. 1-330.
20. Oakes D. The asymptotic information in censored survival data. *Biometrika*. 1977;64(3):441-448.
21. Li QH, Lagakos SW. Use of the Wei-Lin-Weissfeld method for the analysis of a recurring and a terminating event. *Statistics in Medicine*. 1997;16(8):925-940.