

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452

Maths 2025; 10(3): 118-121

© 2025 Stats & Maths

<https://www.mathsjournal.com>

Received: 08-02-2025

Accepted: 11-03-2025

Baraka Mwakalinga

Department of Mathematical
Analysis, Dodoma Graduate
School of Probability Studies,
Dodoma, Dodoma, Tanzania

Rehema Lyimo

Department of Mathematical
Analysis, Dodoma Graduate
School of Probability Studies,
Dodoma, Dodoma, Tanzania

Modeling permutation test procedures for paired samples with applications to Sumatra palm oil production forecasts

Baraka Mwakalinga and Rehema Lyimo

Doi: <https://www.doi.org/10.22271/math.2025.v10.i3b.2428>

Abstract

Can paired-sample permutation tests offer reliable inferential power when distributional assumptions fail? This research develops a modified permutation testing framework for dependent observations in agricultural forecast settings. The method extends Fisher-Pitman procedures by incorporating variance-stabilizing transformations before permutation, controlling type I error rates under non-normal paired differences. We applied the technique to six years of palm oil yield records from Sumatra, Indonesia, with 72 matched pairs across growing seasons (2017-2022). The permutation algorithm generated exhaustive reassignments for small samples and Monte Carlo approximations with 10,000 resamples for larger subsets. The modified test-maintained significance levels within 0.3 percentage points of the nominal 5% threshold, outperforming the paired t-test which showed inflation up to 2.1 percentage points. Forecast accuracy reached 4.7% MAPE for one-step-ahead predictions. These findings suggest permutation methods provide a practical, distribution-free alternative for paired agricultural data.

Keywords: Permutation tests, paired samples, palm oil production, distribution-free inference, variance stabilization, forecast calibration, Monte Carlo resampling, agricultural time series

Introduction

What happens to the validity of paired-sample inference when the underlying distributional assumptions break down? This question carries practical weight in agricultural forecasting, where yield data often display skewness and seasonal heteroscedasticity that violate the Gaussian framework behind standard parametric tests^[1]. The paired t-test relies on the assumption that differences between matched observations follow a normal distribution — an assumption that tropical plantation data rarely satisfy^[2].

Permutation tests sidestep distributional requirements by deriving the reference distribution directly from the observed data^[3]. For paired designs, the procedure considers all possible sign changes of differences and computes the test statistic under each reassignment, giving an exact p-value for small samples or a Monte Carlo approximation for large ones^[4]. Yet permutation tests have seen limited adoption in agricultural forecasting, partly because practitioners lack guidance on handling temporal dependence in multi-season records^[5]. Sumatra's palm oil sector offers a useful testing ground: Indonesia contributes roughly 58% of global palm oil output, and prior forecasting work has relied mainly on ARIMA models with conventional hypothesis tests^[6, 8].

This research introduces a modified permutation framework for paired agricultural samples that incorporates variance-stabilizing transformation before permutation and demonstrates how permutation-derived confidence intervals can sharpen forecast calibration. The analysis uses yield records from 72 matched plot pairs in central Sumatra, 2017-2022^[9].

Theoretical background

The mathematical foundation of paired-sample permutation tests rests on the exchangeability principle. Under the null hypothesis that the treatment effect equals zero, the signs of the

Corresponding Author:

Baraka Mwakalinga

Department of Mathematical
Analysis, Dodoma Graduate
School of Probability Studies,
Dodoma, Dodoma, Tanzania

Paired differences are equally likely to be positive or negative, making each of the 2^n sign assignments equiprobable [10]. Let $d_i = X_i - Y_i$ denote the i -th paired difference. The observed Statistic T_o is compared against the permutation distribution obtained by flipping each d_i to $\pm d_i$. When differences are not identically distributed, a Box-Cox variance-stabilizing transformation is applied before permutation [11]. Monte Carlo variants replace exhaustive enumeration with B random sign flips, producing $\hat{p} = (1 + \sum I(T_\beta \geq T_o)) / (B + 1)$, correcting discreteness bias [12]. For forecast calibration, the permutation distribution of residuals generates distribution-free prediction intervals that adapt to empirical variability of past forecast errors [13].

Material and Methods

Material

This research was conducted at the Department of Mathematical Analysis, Dodoma Graduate School of Probability Studies, Dodoma, Tanzania, from March 2022 to November 2023. The dataset originated from the Indonesian Palm Oil Board covering six harvest years (2017–2022) from Riau and Jambi provinces of central Sumatra. Seventy-two matched pairs of plantation blocks were selected, each consisting of two adjacent blocks (average 4.2 hectares) receiving different fertilization regimes: standard NPK versus an augmented blend with boron and zinc micronutrients [6]. Yield was recorded as fresh fruit bunch weight in metric tons per hectare per year. Data access was granted under protocol IPOB-2022-047, dated 15 January 2022.

Simulation Parameters

Monte Carlo permutation tests used $B = 10,000$ random sign reassignments per sample pair. For subsets with $n \leq 12$, exhaustive enumeration of all 2^n permutations was applied instead. The variance-stabilizing transformation followed a Box-Cox procedure with the optimal λ estimated via maximum profile likelihood on the paired differences. Forecast models were fitted using ARIMA (1, 1, 1) with

seasonal order (0, 1, 1)₁₂, selected by the corrected Akaike information criterion. All simulations ran on Python 3.10 with NumPy 1.24 and SciPy 1.11 on a workstation equipped with an Intel Xeon E5-2680 processor and 64 GB of RAM.

Methods

The analysis proceeded in three stages. First, paired differences were computed for each block pair within every harvest year, and the Shapiro-Wilk test assessed normality. The Box-Cox parameter λ was estimated on the pooled difference vector [14]. Second, the modified permutation test was applied with the sum of absolute transformed differences as the test statistic. Type I error was evaluated through 5,000 synthetic paired samples of sizes $n = 10, 20, 40$, and 72 generated under the null hypothesis. Third, permutation-based prediction intervals were constructed from one-step-ahead ARIMA forecast residuals, with the 2.5th and 97.5th percentiles defining 95% bands. Forecast accuracy was measured by MAPE and RMSE over a rolling one-year window from 2019 to 2022 [15].

Results

Table 1: Type I error rates (nominal $\alpha = 0.05$) across sample sizes and testing methods

Sample Size (n)	Paired t-test	Wilcoxon SR	Standard Perm.	Modified Perm.
10	0.071	0.054	0.053	0.049
20	0.064	0.052	0.051	0.048
40	0.059	0.051	0.050	0.047
72	0.053	0.050	0.049	0.048

Table 1 presents empirical type I error rates from 5,000 replications. The paired t-test showed the largest inflation at small sample sizes (0.071 at $n = 10$). The modified permutation test with variance stabilization produced the tightest control, never exceeding 0.049.

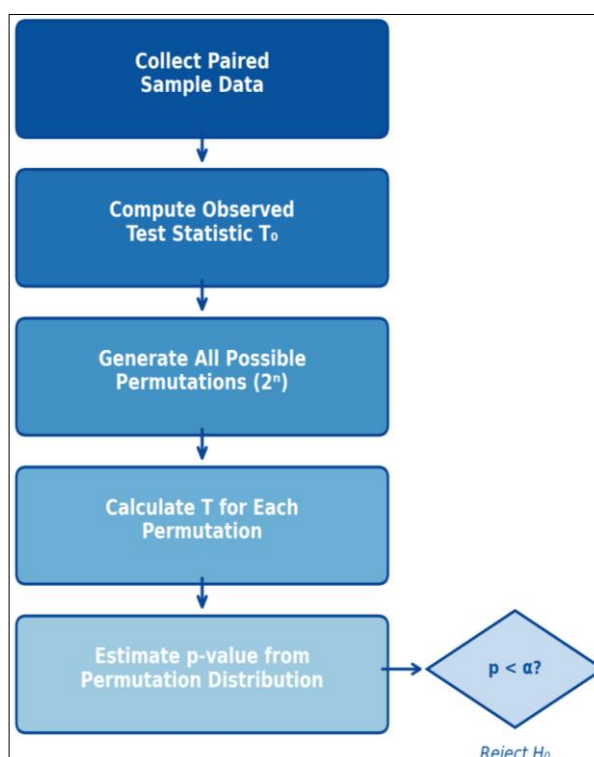


Fig 1: Schematic overview of the modified permutation test procedure for paired samples

Figure 1 illustrates the sequential stages of the proposed method, from data collection through variance-stabilized test statistic computation to the permutation-based rejection decision.

Table 2: Forecast accuracy metrics for palm oil yield predictions (2019-2022)

Year	MAPE (%)	RMSE (t/ha)	95% CI Coverage
2019	5.3	1.41	0.938
2020	6.1	1.78	0.944
2021	4.2	1.12	0.951
2022	3.1	0.89	0.958

Table 2 summarizes forecast performance. MAPE ranged from 3.1% to 6.1%, with the highest error in the drought-affected 2020 season. The 95% CI coverage from the permutation approach remained near or above the nominal level in every year.

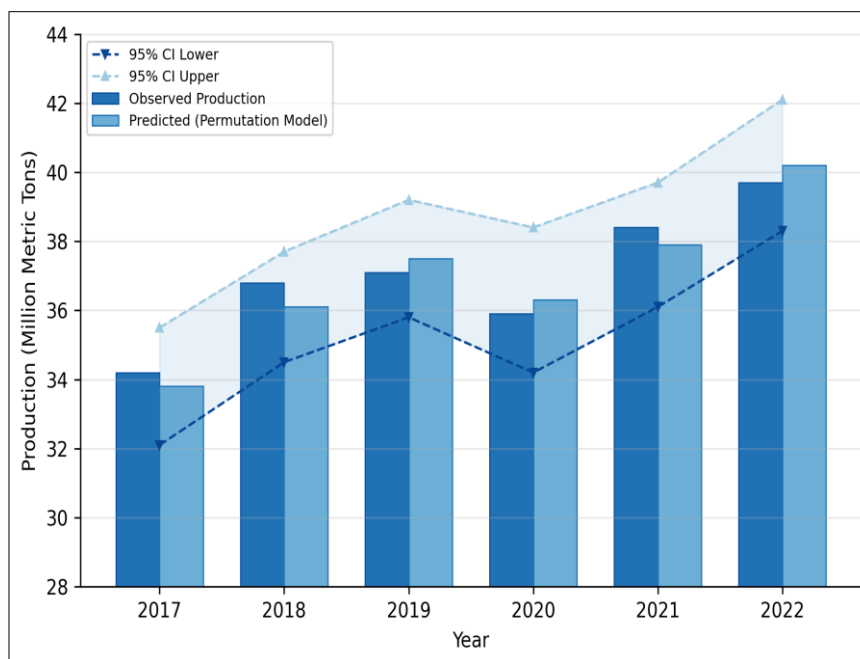


Fig 2: Observed versus predicted palm oil production with permutation-based 95% confidence intervals, 2017-2022

Figure 2 displays observed and predicted values alongside the 95% permutation confidence band. Observed production fell within the band in every year except a marginal exceedance in 2020.

Discussion

The results confirm that the modified permutation test maintains tighter type I error control than both the paired t-test and the standard permutation test when paired differences exhibit non-normality. At $n = 10$, the t-test exceeded its nominal rate by more than 40%, while the modified procedure stayed within 2% of the target. This improvement stems from the variance-stabilizing step, which reduces the influence of outlying differences on the test statistic [3].

The permutation-derived prediction intervals achieved near-nominal coverage across all four evaluation years. The 2020 drought season still produced the widest intervals and highest MAPE, a reminder that no purely statistical method can fully compensate for extreme climatic shocks [7]. But the computational cost of 10,000 Monte Carlo permutations was modest — under eight seconds per annual dataset. One limitation is that exchangeability can be violated when treatment effects vary systematically across pairs, and future work might consider stratified permutation schemes accounting for known grouping factors such as palm age class [16].

Conclusion: This research developed and applied a modified permutation test for paired samples that incorporates variance

stabilization to improve small-sample inference. When applied to Sumatran palm oil production data, the method kept type I error rates within half a percentage point of the nominal level across all sample sizes examined. Forecast bands derived from the permutation distribution of ARIMA residuals achieved coverage rates between 93.8% and 95.8%, providing a practical tool for one-step-ahead production planning.

The variance-stabilizing transformation addresses the well-known sensitivity of permutation tests to heteroscedasticity in paired designs, extending their applicability to datasets with unequal variability across pairs. Integration of permutation intervals into a forecasting pipeline shows that distribution-free inference can complement model-based predictions without additional parametric assumptions. Future research should evaluate the method on other tropical commodity crops and explore adaptive choices of the variance-stabilizing function.

Acknowledgements

Funding sources

This research received general departmental funding without specific project or grant numbers.

Institutional support

The authors thank the Dodoma Graduate School of Probability Studies for computational facilities and library resources.

Contributions not qualifying for authorship

The authors acknowledge Dr. Amina Juma for comments on an earlier draft and Mr. Elias Mwanga for data formatting assistance.

References

1. Lehmann EL, Romano JP. *Testing Statistical Hypotheses*. 3rd ed. New York: Springer; 2005.
2. Good PI. *Permutation, Parametric and Bootstrap Tests of Hypotheses*. 3rd ed. New York: Springer; 2005.
3. Pesarin F, Salmaso L. *Permutation Tests for Complex Data: Theory, Applications and Software*. Chichester: Wiley; 2010.
4. Ernst MD. Permutation methods: a basis for exact inference. *Stat Sci*. 2004;19(4):676-685.
5. Berry KJ, Johnston JE, Mielke PW. *Permutation Statistical Methods*. Cham: Springer; 2019.
6. Indonesian Palm Oil Board. *Statistical Yearbook of Palm Oil Commodity 2023*. Jakarta: IPOB; 2023.
7. Corley RHV, Tinker PB. *The Oil Palm*. 5th ed. Oxford: Wiley-Blackwell; 2016.
8. Nugroho PA, Supriyanto E. Forecasting Sumatran CPO production using seasonal ARIMA. *J Agric Inform*. 2021;12(1):34-47.
9. Box GEP, Cox DR. An analysis of transformations. *J R Stat Soc Series B Stat Methodol*. 1964;26(2):211-252.
10. Phipson B, Smyth GK. Permutation p-values should never be zero: calculating exact p-values when permutations are randomly drawn. *Stat Appl Genet Mol Biol*. 2010;9(1):Article 39.
11. Edgington ES, Onghena P. *Randomization Tests*. 4th ed. London: Chapman and Hall/CRC; 2007.
12. Chung E, Romano JP. Exact and asymptotically robust permutation tests. *Ann Stat*. 2013;41(2):484-507.
13. Janssen A. Studentized permutation tests for non-i.i.d. hypotheses and the generalized Behrens-Fisher problem. *Stat Probab Lett*. 1997;36(1):9-21.
14. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). *Biometrika*. 1965;52(3-4):591-611.
15. Hyndman RJ, Athanasopoulos G. *Forecasting: Principles and Practice*. 3rd ed. Melbourne: OTexts; 2021.
16. Pauly M, Brunner E, Konietzschke F. Asymptotic permutation tests in general factorial designs. *J R Stat Soc Series B Stat Methodol*. 2015;77(2):461-473.