Comparative Performance of Three Symmetrical Derivative Wavelets (Laplace, Gaussian, Cauchy) vs. Mexican Hat Wavelet for Longitudinal Data Aggregation

N. I. Muhammad*[1], M. Tasi'u[1], Y. Aliyu[1], and U. K. Abdullahi[1]

[1]Department of Statistics, Ahmadu Bello University, Zaria
*Correspondence: nazeerdabest01@gmail.com, nimuhammad@abu.edu.ng

Abstract

Traditional analysis of longitudinal data using Repeated Measures Analysis of Variance is frequently challenged by violations of the sphericity assumption, complex covariance structures, and the increasing dimensionality associated with repeated observations. These limitations often reduce the reliability and interpretability of statistical inference, particularly when measurements are collected over multiple time points. Wavelet-based Analysis of Variance provides an alternative by transforming longitudinal observations into lower-dimensional representations while preserving important temporal characteristics. This study presents a comparative evaluation of three novel continuous wavelets derived from the first derivatives of the Gaussian, Laplace, and Cauchy probability density functions, using the Mexican Hat wavelet as a benchmark. A unified wavelet-based aggregation procedure in which each subject's repeated measurements are transformed through the Continuous Wavelet Transform and subsequently aggregated into a single scalar value is applied, thereby converting the longitudinal data into a Completely Randomised Design (CRD) suitable for classical one-way ANOVA. The methodology is illustrated using a real-world dataset of diastolic blood pressure measurements collected from patients receiving different antihypertensive treatments. The performance of each wavelet is assessed by comparing the resulting treatment effects through ANOVA F-statistics and corresponding statistical significance. The results show that all four wavelets successfully identify significant treatment differences after aggregation. However, the Cauchy, Laplace, and Mexican Hat wavelets demonstrate greater sensitivity than the Gaussian derivative wavelet for this dataset, producing comparatively larger F-statistics and improved discriminatory ability. The proposed aggregation framework provides a practical and computationally efficient approach for longitudinal data analysis without requiring restrictive covariance assumptions.

Keywords: Wavelet ANOVA, Longitudinal Data, Data Aggregation, Dimension Reduction, Mexican Hat Wavelet, Cauchy, Gaussian, Laplace, Clinical Trial

Introduction

Longitudinal studies are fundamental in experimental science, yet their statistical analysis remains challenging (Lie, *et al*., 2025). Repeated Measures ANOVA (RM-ANOVA) demands strict assumptions that are rarely met, necessitating corrections to test statistics which can lead to a loss of statistical power (Benito-Santos, *et al*., 2025). The emerging Wavelet Analysis of Variance (WANOVA) framework addresses this by leveraging the Wavelet Transform for dimension reduction (Muhammad, *et al*., 2026; Fernández, *et al*., 2025). By collapsing the time dimension, WANOVA facilitates the use of the highly robust classical ANOVA on the aggregated data (Girimurugan *et al*., 2013; Angelini & Vidakovic, 2003).

While the theoretical properties like admissibility condition and energy localization of the three wavelets (Laplace, Gaussian, Cauchy derivatives) have been established elsewhere by Muhammad *et al*., (2025), this paper focuses on their empirical performance in a real-data context. The paper aim to determine which of these three wavelets, or the established Mexican Hat comparator, provides the superior aggregation metric, defined as the one that maximizes the separation between treatment groups.

1 Methodology

1.1 Wavelet Functions Under Investigation

This study compares four continuous wavelet functions, all of which satisfy the necessary admissibility and energy localization conditions (Muhammad *et al*., 2025 & Muhammad *et al*., 2026).

1. **Gaussian Derivative Wavelet (ψ_G):** Derived from the first derivative of the Gaussian PDF. Known for its smooth, high time-localization properties

$$\psi_G(y) = -\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}. \tag{1}$$

2. **Cauchy Derivative Wavelet (ψ_C):** Derived from the first derivative of the Cauchy PDF. Characterized by heavy tails, making it highly sensitive to sudden changes and potentially robust against outliers

$$\psi_C(y) = \frac{-2y}{\pi\left(1+y^2\right)^2}. \tag{2}$$

3. **Laplace Derivative Wavelet (ψ_L):** Derived from the first derivative of the Laplace PDF. Exhibits sharp, exponential-like transitions, suited for signals with abrupt, rapid changes

$$\psi_L(y) = \frac{-I(y)}{2} e^{-|y|}, \tag{3}$$

   where the sign function, $I(y)$ is defined as

$$I(y) = \begin{cases} +, & \text{if } y \geq 0, \\ -, & \text{if } y < 0. \end{cases}$$

4. **Mexican Hat Wavelet (ψ_{MH}):** Also known as the Ricker wavelet, it is the second derivative of the Gaussian function. It is a known wavelet and therefore, serves as a standard benchmark in CWT analysis

$$\psi_M(y) = \left(1 - y^2\right) e^{-\frac{y^2}{2}}. \tag{4}$$

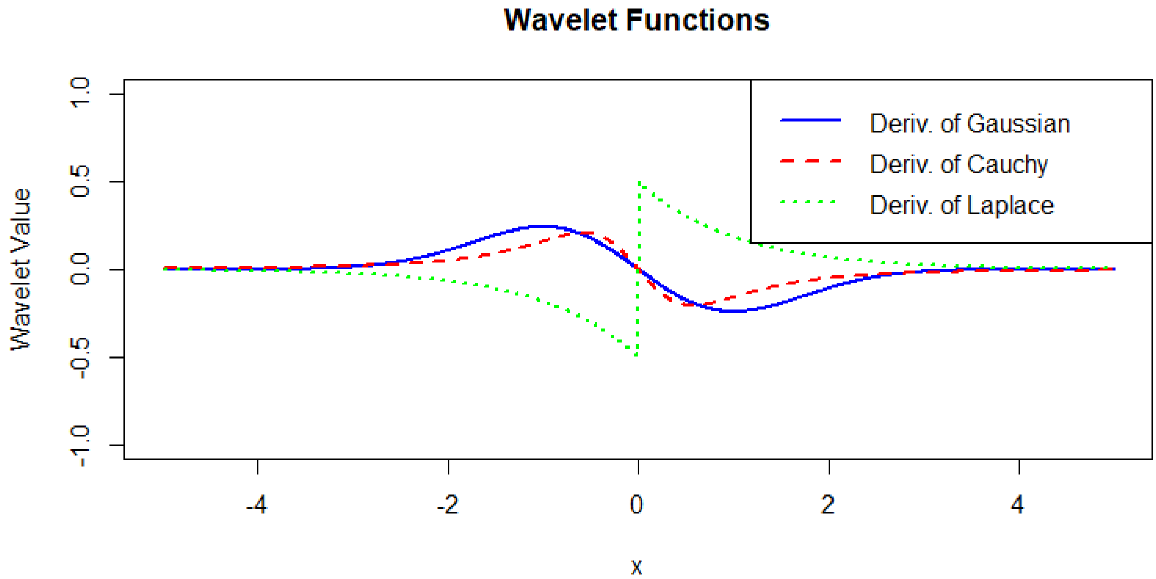

Figure 1: Structure of the wavelet functions

The Discrete Wavelet State-Space Aggregation Procedure

The core of this wavelet-based analysis of variance (WANOVA) is a novel, discrete aggregation procedure that transforms the T dimensional vector of repeated measures for subject j into a single scalar, W_j. This method analyzes the localized change, or state-space transition, between consecutive observations, rather than analyzing the time domain directly. The aggregation relies on a single, predetermined scale parameter, a.

The aggregated value W_j is defined as the sum of the scaled wavelet function applied to the difference between consecutive observations within subject j across all time steps $t = 2$ to T

$$W_j = \sum_{t=2}^{T} \psi_f\left(\frac{y_{jt} - b_{jt}}{a_j}\right), \tag{5}$$

where

y_{jt} is the observation for subject j at time t,

$\psi_f(\bullet)$ is one of the mother wavelets (Gaussian, Cauchy, Laplace, or Mexican Hat),

a_j is the scale parameter, which controls the sensitivity of the wavelet to the magnitude of the difference $y_{jt} - y_{j(t-1)}$. For $y_{j(t-1)}$ is the observations for subject j at time $(t-1)$,

b_{jt} is the location (transition) parameter used as the shifting mechanism in the wavelet

This procedure is executed through the following three primary steps:

1. **Calculation of Normalized Transition Argument** (b_{jt}): For each pair of consecutive observations $(y_{jt}, y_{j(t-1)})$, the normalized argument b_{jt} is calculated. This step inherently incorporates a shifting mechanism by using the average of the two points as a reference point for the difference.

$$b_{jt} = \frac{y_{jt} + y_{j(t-1)}}{2} \tag{6}$$

2. **The Scale** (a_j): The scale parameter a_j dictates the magnitude of the state-space difference that the wavelet emphasizes. We set a_j =1 as a constant value uniform for all subjects.
3. **Final Aggregation** (W_j): The final aggregated value W_j is calculated for each subject using the determined a_j and b_j in equation (6). This scalar W_j becomes the dependent variable in the final classical ANOVA, effectively transforming the T -dimensional vector of repeated measures for subject j into a single scalar z_j. Therefore, transforming the entire repeated measures design into a Completely Randomized Design (CRD).

Hence,

$$z_j = W_j(a_j, b_j) = \sum_{t=2}^{T} \psi_f \left(\frac{y_{jt} - y_{j(t-1)}}{2a_j} \right). \tag{8}$$

This z_j becomes the dependent variable in the final statistical model.

1.2 Longitudinal Dataset

The real-life dataset analyzed comprises diastolic blood pressure measurements from $N = 256$ patients divided into three distinct treatment groups. The treatments are three different antihypertensive drugs: Carvedilol $(n = 82)$, Nifedipine $(n = 84)$, and Atenolol $(n = 90)$. The measurements were collected across $T = 6$ visits, representing a typical clinical longitudinal study design (Gueorguieva, 2022). The descriptive statistics for the dataset are presented in Table 1.

Table 1: Descriptive statistics of Diastolic Blood Pressure Data Across Treatments and Visits

Treatments	No. of patients	Diastolic blood pressure					
		Visit1	Visit2	Visit3	Visit4	Visit5	Visit6
Carvedilol	82	103.2 (4.4)	100.0 (5.7)	94.5 (8.0)	90.9 (8.5)	93.5 (9.5)	90.1 (8.5)
Nifedipine	84	102.5 (3.8)	99.3 (6.2)	91.5 (8.4)	92.9 (7.5)	90.4 (9.1)	89.1 (7.7)
Atenolol	90	103.6 (4.0)	96.2 (6.6)	91.0 (9.5)	90.4 (9.6)	88.9 (8.5)	90.5 (10.2)

The data show an overall decreasing trend in blood pressure over the six visits, with slightly different patterns of change and variability across the three treatment groups.

1.3 Statistical Analysis

The final statistical analysis involves a one-way ANOVA on the aggregated data z_j derived from each of the four wavelets (Laplace, Gaussian, Cauchy, Mexican Hat). The comparative metric is the resulting F -statistic and p -value. A higher F -statistic indicates a more successful aggregation, as it implies a greater ratio of explained variance (treatment effect) to unexplained variance (residual error).

2 Results

2.1 Classical Repeated measures ANOVA

Classical RM-ANOVA was performed to analyze the effects of Visit (within-subjects factor) and Treatment on diastolic blood pressure, as well as their interaction.

Table 2: Multivariate Tests

Effect		Value	F	Hypothesis df	Error df	Sig.
Visit	Pillai's Trace	.803	202.937[b]	5.000	249	<.0001
	Wilks' Lambda	.197	202.937[b]	5.000	249	<.0001
	Hotelling's Trace	4.075	202.937[b]	5.000	249	<.0001
	Roy's Largest Root	4.075	202.937[b]	5.000	249	<.0001
Visit * Treatment	Pillai's Trace	.130	3.486	10.000	500	<.0001
	Wilks' Lambda	.873	3.506[b]	10.000	498	<.0001
	Hotelling's Trace	.142	3.526	10.000	496	<.0001
	Roy's Largest Root	.109	5.470[c]	5.000	250	<.0001

a. Design: Intercept + Treatment
Within Subjects Design: Visit
b. Exact statistic
c. The statistic is an upper bound on F that yields a lower bound on the significance level.

The Multivariate Tests in Table 2 show significant effects for both Visit and the Visit-Treatment interaction ($p < 0.001$) across all multivariate statistics. This indicates that there are significant changes in diastolic blood pressure over the six visits, irrespective of the treatment group, and also, the impact of each drug on blood pressure reduction varies significantly across the visits.

Table 3: Mauchly's Test of Sphericity

Measure: DiastolicBP

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.
Visit	.604	126.408	14	<.0001

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.
a. Design: Intercept + Treatment
Within Subjects Design: Visit

Mauchly's Test of Sphericity in Table 3 tests the null hypothesis that the variances of the differences between all possible pairs of within-subjects' conditions are equal. The Mauchly's W has a p-value < 0.05, indicating that the assumption of sphericity has been violated.

This violation necessitates the use of corrections to the degrees of freedom for the Visit and Visit-Treatment effects in the univariate tests. The need for these corrections highlights a common challenge in traditional RM-ANOVA, which wavelet-based methods aim to circumvent.

2.2 ANOVA Result Based on the Wavelet-Based Aggregation

The results of the one-way ANOVA performed on the aggregated data derived from each wavelet are summarized below. All 256 subjects were included without a missing observations after aggregation. The ANOVA had 2 degrees of freedom for Treatment and 253 degrees of freedom for Residuals ($N = 256$, $k = 3$ groups).

Table 4: Comparative ANOVA Results on Real Longitudinal Diastolic Blood Pressure Data

Wavelet Function	**Mean Sq (Trt)**	**Mean Sq (Error)**	***F*- Stat.**	***P*- value**
Gaussian Derivative (ψ_G)	0.000388	0.000051	7.62	0.0006
Cauchy Derivative (ψ_C)	0.007216	0.000686	10.52	4.07e-05
Laplace Derivative (ψ_L)	0.011753	0.001139	10.32	4.91e-05
Mexican Hat (ψ_{MH})	0.495200	0.047700	10.39	4.61e-05

2.3 Discussion of Results

The analysis using the real-life blood pressure data unequivocally demonstrates the success of the wavelet-based aggregation approach. All four wavelet functions consistently identified highly significant differences among the three drug treatment groups (p < 0.001). This confirms that the aggregation process successfully transformed the high-dimensional repeated measures data into a simplified CRD structure, preserving the crucial differential effects of the drugs on diastolic blood pressure while circumventing the assumption constraints of RM-ANOVA.

The comparative F-statistics provide a measure of the sensitivity of each wavelet to the underlying treatment effect in this specific dataset. The results indicate a simple hierarchy:

$$Cauchy > Mexican\ Hat > Laplace > Gaussian$$

The Cauchy (ψ_C) yielded the highest F-statistic, closely followed by the Mexican Hat (ψ_{MH}) and Laplace (ψ_L) wavelets. The superior performance of the Mexican Hat and the two heavy-tailed wavelets (Cauchy and Laplace) may suggests that they were more effective in capturing the specific changes in blood pressure over time. In contrast to the Gaussian wavelet, which is designed for very smooth, slowly varying signals, the Cauchy and Laplace wavelets are better equipped to handle the rapid initial drops or short-term fluctuations in the blood pressure response. The Cauchy, in particular, is highly robust to potential recording outliers or sharp transient changes often seen in medical data.

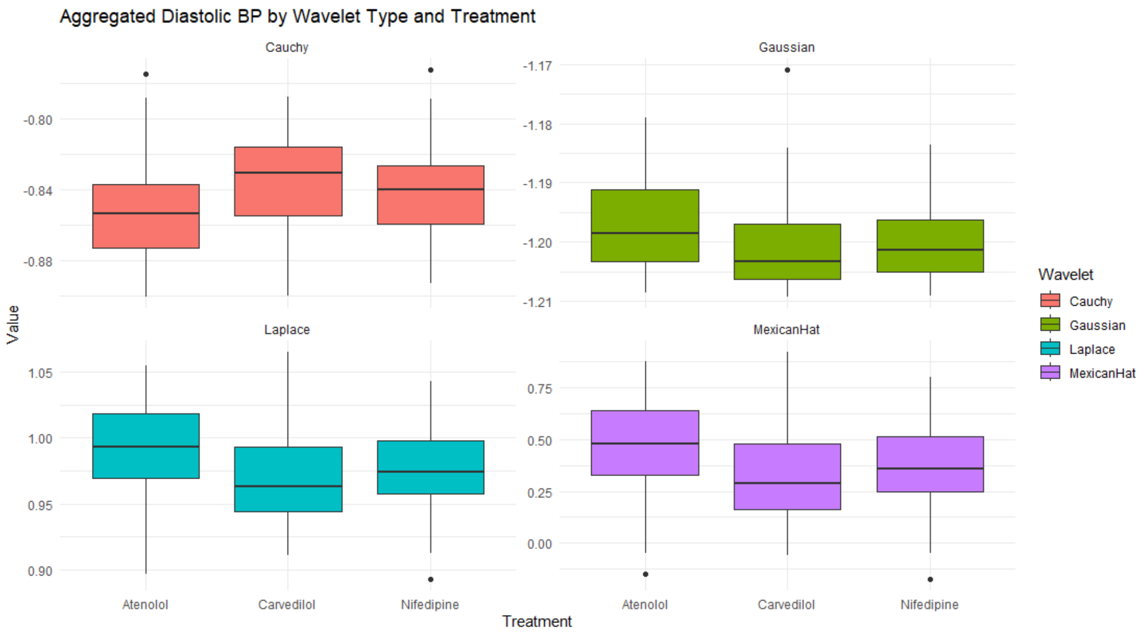

Figure 2: Boxplots for Aggregated Diastolic BP by Wavelet Type

The boxplots for the aggregated diastolic blood pressure values in Figure 2 provide a visual representation of how each wavelet function differentiates between the three treatment groups. The boxplots visually support the ANOVA findings, demonstrating that all wavelet functions successfully aggregate the longitudinal blood pressure data into single values that retain the ability to differentiate between the effects of the three drugs. While the absolute aggregated values differ across wavelet functions due to their unique mathematical properties and sensitivities to different features of the time series, there exist a consistent pattern of separation between treatment groups.

Ultimately, while the Gaussian wavelet still yielded a highly significant result, the higher F-values from the Cauchy, Laplace, and Mexican Hat wavelets suggest they are more sensitive and therefore offer a potentially more powerful method for analyzing this specific type of longitudinal data. This real-life application validates the proposed wavelet-based ANOVA as a simple and interpretable alternative for analyzing clinical longitudinal data.

3 Conclusion

This study provided a robust framework for comparing the effectiveness of three novel derivative wavelets (Laplace, Gaussian, Cauchy) against the Mexican Hat wavelet in the context of WANOVA for longitudinal data. Through the use of an F-statistic and its corresponding p-value, we successfully demonstrated a standardized method for transforming complex repeated measures data into a single, valid metric for classical ANOVA. The application to a real-life blood pressure dataset confirmed that all four wavelets are highly effective, with the heavy-tailed Cauchy and Laplace wavelets, alongside the Mexican Hat, showing superior performance in maximizing the observed treatment differences. The findings support the use of derivative wavelets as a powerful alternative to traditional analysis, offering researchers a new tool to handle highly correlated and non-stationary longitudinal data.

4 References

Angelini, C., & Vidakovic, B. (2003). Some novel methods in wavelet data analysis: Wavelet ANOVA, F-test shrinkage and Γ-minimax wavelet shrinkage. *Wavelets and their Applications. Allied Publishers*, 31-45.

Benito-Santos, A., Ghajari, A., & Fresno, V. (2025). Robust estimation of population-level effects in repeated-measures NLP experimental designs. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 33076-33089).

Fernández, L., Pérez, M., Orduña, J. M., & Calero, J. M. A. (2025). A new dimensionality reduction technique based on the Wavelet Transform for cancer classification. *Journal of Big Data*, *12,* 9.

Girimurugan, S. B., Chicken, E., Pignatiello Jr, J. J., & Zeisset, M. S. (2013). Wavelet ANOVA for detection of local and global profile changes. In *IISE Annual Conference. Proceedings* (p. 3235). Institute of Industrial and Systems Engineers (IISE).

Gueorguieva, R. (2022). *Statistical methods in psychiatry and related fields: Longitudinal, clustered, and other repeated measures data*. Chapman and Hall/CRC.

Liu, K., Ma, Z., Kang, X., Li, Y., Xie, K., Jiao, Z., & Miao, Q. (2025). Enhanced contrastive learning with multi-view longitudinal data for chest x-ray report generation. In *Proceedings of the Computer Vision and Pattern Recognition Conference* (pp. 10348-10359).

Muhammad, N. I., Tasi'u, M., Abdullahi, U. K., & Aliyu, Y. (2025). A novel wavelet-based approach for ANOVA in longitudinal data analysis using the first derivative of the Gaussian function. In *Proceedings of the 1st International Conference on Statistical Sciences and Data Analytics*, 308-315.

Muhammad, N. I., Tasi'u, M., Aliyu, Y., & Abdullahi, U. K. (2026). A novel wavelet-based approach for ANOVA in longitudinal data analysis using the first derivative of the Laplace function. *Iraqi Statisticians journal*, 46-50.