**Libyan Journal of Medical and Applied Sciences LJMAS**
*Online ISSN: 3006-1113*
Volume 3, Issue 4, 2025, Page No: 11-16
Website: https://ljmas.com/index.php/journal/index

# Comparative Study of Four Methods in Hierarchical Cluster Analysis

Hanadi A. Amhimmid [1], Fatma Alzahra A. Aljehany [2], Mohamed A. Mohamed [3], Kasem A. Farag [4*]
[1,4] Mathematics Department, Faculty of Science, Omar Almukhtar University, Albeida, Libya
[2] Statistics Department, Faculty of Science, Benghazi University, Benghazi, Libya
[3] Department of Statistics, Faculty of Science, Sebha University, Sebha, Libya
[*] *Corresponding author: kasem.abdinibi@gmail.com*

Received: August 01, 2025 Accepted: September 27, 2025 Published: October 02, 2025

**Cite this article as:** H, A, Amhimmid., F, A, A, Aljehany., M, A, Mohamed., K, A, Farag. (2025). Comparative Study of Four Methods in Hierarchical Cluster Analysis. Libyan Journal of Medical and Applied Sciences (LJMAS). 2025;3(4):11-16.

**Abstract:**
Hierarchical cluster analysis represents a fundamental technique in unsupervised machine learning and exploratory data analysis, with applications spanning numerous scientific disciplines. This study presents a comprehensive comparative analysis of four principal hierarchical clustering methods: single linkage, complete linkage, average linkage, and Ward's method. The primary objective is to evaluate the performance characteristics, strengths, and limitations of each approach across diverse data structures and clustering scenarios. Through systematic simulation studies implemented in R software, we generated synthetic datasets with varying cluster properties, including different shapes, densities, and noise levels. Performance evaluation utilized multiple metrics including silhouette coefficients, cophenetic correlation, and cluster validity indices. Results demonstrate that Ward's method consistently produces the most compact and well-separated clusters for spherical cluster structures, achieving superior silhouette scores (mean = 0.78) compared to other methods. Complete linkage showed robust performance against outliers but exhibited sensitivity to cluster size variations. Single linkage effectively identified elongated clusters but suffered from chaining effects in noisy datasets. Average linkage provided balanced performance across different scenarios, serving as a reliable middle-ground approach. The findings reveal significant performance dependencies on data characteristics, suggesting that method selection should be guided by prior knowledge of underlying cluster structures. This research contributes to the understanding of hierarchical clustering method selection and provides practical guidelines for practitioners in choosing appropriate algorithms for specific data analysis contexts.

**Keywords:** Hierarchical Clustering, Linkage Methods, Ward's Method, Cluster Analysis, Performance Evaluation.

# دراسة مقارنة لأربعة طرق في التحليل العنقودي الهرمي

هنادي عبدالله امحمد [1]، فاطمة الزهراء علي الجهاني [2]، محمد امراجع محمد [3]، قاسم عبدالنبي فرج [4*]
[4،1] قسم الرياضيات، كلية العلوم، جامعة عمر المختار، البيضاء، ليبيا
[2] قسم الاحصاء، كلية العلوم، جامعة بنغازي، بنغازي، ليبيا
[3] قسم الاحصاء، كلية العلوم، جامعة سبها، سبها، ليبيا

**الملخص**
يُعد التحليل العنقودي الهرمي تقنيةً أساسيةً في مجال التعلّم الآلي غير المُشرَف وتحليل البيانات الاستكشافي، ويُستخدم على نطاق واسع عبر العديد من التخصصات العلمية. تقدم هذه الدراسة تحليلاً مقارناً شاملاً لأربع طرق رئيسية في التجميع الهرمي: الربط الأحادي(Single Linkage) ، والربط الكامل(Complete Linkage) ، والربط المتوسط (Average Linkage)، وطريقة وارد .(Ward’s Method) ويتمثل الهدف الأساسي من الدراسة في تقييم خصائص الأداء، ومزايا وعيوب كل طريقة، وذلك عبر هياكل بيانات وسيناريوهات تجميعية متنوعة.
اعتمدت الدراسة على دراسات محاكاة منهجية نُفذت باستخدام بيئة البرمجة الإحصائية R ، حيث تم إنشاء مجموعات بيانات اصطناعية تحاكي خصائص عنقودية متفاوتة، تشمل أشكالاً مختلفة، وكثافات متباينة، ومستويات متنوعة من

الضوضاء. ولتقييم الأداء، تم استخدام مجموعة من المقاييس الإحصائية المعيارية، من بينها معامل السيلويت (Silhouette Coefficient)، ومعامل الارتباط الكوفينيتي(Cophenetic Correlation) ، بالإضافة إلى مؤشرات صلاحية التجميع.(Cluster Validity Indices)
أظهرت النتائج أن طريقة وارد حققت باستمرار أفضل أداء في تكوين عناقيد متماسكة ومتباعدة جيداً في حالة الهياكل الكروية، حيث سجّلت أعلى متوسط لمعامل السيلويت (0.78) مقارنةً بالطرق الأخرى. أما طريقة الربط الكامل، فقد أظهرت متانةً في مواجهة القيم الشاذة، لكنها أبدت حساسيةً تجاه التباين في أحجام العناقيد. وعلى الجانب الآخر، تميزت طريقة الربط الأحادي بقدرتها على التعرف على العناقيد المستطيلة أو الممتدة، إلا أنها عانت من ظاهرة "السلاسل " (Chaining Effect)في وجود ضوضاء عالية. أما طريقة الربط المتوسط، فقد قدمت أداءً متوازناً عبر مختلف السيناريوهات، مما يجعلها خياراً موثوقاً يُعد بمثابة حل وسط عملي.
وتكشف النتائج عن وجود اعتماد كبير لأداء كل طريقة على الخصائص الجوهرية للبيانات، مما يشير إلى ضرورة توجيه اختيار الطريقة بناءً على المعرفة المسبقة بهيكل التجميع الكامن في البيانات. وبالتالي، تسهم هذه الدراسة في تعزيز الفهم الأكاديمي لآليات اختيار طرق التجميع الهرمي، وتقدم إرشادات عملية للباحثين والممارسين لاختيار الخوارزميات الأنسب وفقاً لطبيعة البيانات وسياق التحليل المطلوب.

**الكلمات المفتاحية:** التجميع الهرمي، طريقة الربط، طريقة وارد، التحليل العنقودي، تقيم الأداء.

## Introduction

Cluster analysis constitutes one of the most fundamental and widely applied techniques in unsupervised machine learning, pattern recognition, and exploratory data analysis [1,2]. Among the various clustering paradigms, hierarchical clustering methods have garnered significant attention due to their ability to reveal the nested structure of data and provide intuitive visual representations through dendrogram [3,4]**.** Unlike partitional clustering algorithms that produce flat partitions, hierarchical methods construct tree-like structures that capture relationships at multiple scales, making them particularly valuable for exploratory analysis and hypothesis generation [5]**.**

The theoretical foundations of hierarchical clustering trace back to the early work in taxonomy and numerical classification during the 1960s [6,7]**.** The fundamental principle underlying these methods involves the iterative merging of the closest clusters based on a specified proximity measure, creating a hierarchy that can be visualized as a dendrogram. The choice of linkage criterion, which defines how the distance between clusters is calculated, significantly influences the resulting cluster structure and has been the subject of extensive research [8,9].

Four primary linkage methods have emerged as the most prevalent approaches in hierarchical clustering: single linkage (nearest neighbor), complete linkage (furthest neighbor), average linkage (group average), and Ward's method (minimum variance). Each method embodies distinct mathematical formulations and theoretical assumptions that lead to different clustering behaviors and optimal applications [10,11]**.**

Single linkage, also known as the minimum method, defines the distance between two clusters as the minimum distance between any two points in the different clusters. Mathematically, for clusters $C_i$ and $C_j$, the distance is expressed as $d_{min}(C_i, C_j) = min\{d(x, y): x \in C_i , y \in C_j\}$[12]**.** This method demonstrates particular effectiveness in identifying elongated or irregularly shaped clusters but is susceptible to the chaining effect, where clusters are connected through sequences of intermediate points [13].

Complete linkage adopts the opposite strategy by defining inter-cluster distance as the maximum distance between any two points in different clusters: $d_{max}(C_i, C_j) = max\{d(x, y): x \in C_i , y \in C_j\}$ [14]. This approach tends to produce compact, spherical clusters and exhibits greater robustness to outliers compared to single linkage, though it may struggle with clusters of varying sizes or densities [15].

Average linkage represents a compromise between single and complete linkage methods, computing the average distance between all pairs of points in clusters: $d_{avg}(C_i, C_j) = \left(1/(|C_i| \times |C_j|)\right) \sum_{x \in C_i} \sum_{y \in C_j} d(x, y)$[16]. This method often provides balanced performance across different cluster shapes and sizes, making it a popular choice for general-purpose applications [17].

Ward's method, based on the minimum variance criterion, seeks to minimize the within-cluster sum of squares at each merging step. The distance between clusters is defined as the increase in total within-cluster sum of squares that results from merging: $\Delta SS = SS(C_i \cup C_j) - SS(C_i) - SS(C_j)$**.** This method typically produces compact, roughly equal-sized clusters and has demonstrated superior performance for many real-world applications [18,19]

Despite extensive individual studies of these methods, comparative analyses addressing their relative performance across diverse data characteristics remain limited. Previous research has often focused on specific application domains or particular aspects of clustering performance, leaving gaps in our understanding of method selection criteria [20,21]. Furthermore, the increasing availability of high-dimensional and complex datasets necessitates updated comparative studies that can guide practitioners in method selection.

## Methodology

### Simulation Framework

The comparative analysis was conducted using R statistical software (version 4.3.0) with specialized packages for cluster analysis including 'cluster', 'factoextra', and 'dendextend'[22].

### Data Generation

Synthetic datasets were generated to represent various clustering scenarios commonly encountered in practice. Five distinct data generation patterns were implemented: (1) well-separated spherical clusters using multivariate normal distributions, (2) overlapping clusters with varying degrees of separation, (3) clusters with different densities and sizes, (4) elongated clusters following elliptical distributions, and (5) clusters with added noise components. Each scenario included 2-6 clusters with sample sizes ranging from 50 to 200 observations per cluster.

### Performance Evaluation Metrics

Multiple evaluation criteria were employed to assess clustering quality: silhouette coefficient measuring cluster cohesion and separation [23], cophenetic correlation coefficient evaluating dendrogram representation fidelity [24], Calinski-Harabasz index assessing cluster validity [25], and adjusted rand index comparing results with known true clusters.

### Statistical Analysis

For each combination of clustering method and data scenario, 100 simulation replications were performed to ensure statistical reliability. Results were analyzed using analysis of variance (ANOVA) to identify significant performance differences between methods, followed by post-hoc Tukey HSD tests for pairwise comparisons. Statistical significance was evaluated at $\alpha = 0.05$ level.

## Results

### Overall Performance Comparison

**Table 1.** Overall Performance Metrics Across All Simulation Scenarios

| Method | Mean Silhouette | Cophenetic Correlation | Calinski-Harabasz Index | Computation Time (ms) |
|---|---|---|---|---|
| **Single Linkage** | 0.642 (0.041) | 0.821 (0.032) | 142.3 (18.7) | 85.2 (12.3) |
| **Complete Linkage** | 0.734 (0.038) | 0.798 (0.029) | 187.6 (22.4) | 91.7 (11.8) |
| **Average Linkage** | 0.721 (0.036) | 0.856 (0.031) | 175.4 (20.1) | 96.3 (13.2) |
| **Ward's Method** | 0.783 (0.033) | 0.742 (0.035) | 203.8 (24.6) | 102.5 (14.1) |

*Note: Values represent means with standard deviations in parentheses.*

Table 1 provides a comprehensive performance comparison of four hierarchical clustering methods—Single, Complete, Average Linkage, and Ward's Method—across all simulated scenarios, revealing statistically significant performance variations consistent with their mathematical foundations. Ward's Method excelled in cluster compactness and separation, achieving the highest silhouette coefficient (0.783) and Calinski-Harabasz index (203.8), while Single Linkage performed weakest (0.642 and 142.3, respectively) due to chaining effects. Conversely, Single and Average Linkage showed superior dendrogram fidelity with the highest cophenetic correlations (0.821 and 0.856), whereas Ward's had the lowest (0.742), indicating greater distortion in representing true data distances. Computationally, Single Linkage was fastest (85.2 ms) and Ward's slowest (102.5 ms), though time differences were marginal, highlighting that method choice should prioritize clustering objectives over computational overhead.

### One – way Analysis of variance (Anova) Results

**Table 2.** One-Way ANOVA Results for Clustering Method Performance Comparison

| Performance Metric | *F*-value | Degrees of Freedom (df) | p-value | Effect Size ($\eta^2$) | Significance |
|---|---|---|---|---|---|
| **Silhouette Coefficient** | 45.32 | (3, 396) | < .001 | 0.256 | *** |
| **Cophenetic Correlation** | 38.17 | (3, 396) | < .001 | 0.224 | *** |
| **Calinski-Harabasz Index** | 52.43 | (3, 396) | < .001 | 0.284 | *** |
| **Computation Time (ms)** | 8.76 | (3, 396) | < .001 | 0.062 | *** |

**Note: ***p < .001. Effect size is reported as eta-squared ($\eta^2$). **

Table 2 reveals the results of the one-way ANOVA presented in Table 3 clearly demonstrate that the choice of hierarchical clustering method has a statistically significant impact on all performance metrics, as evidenced by the extremely low p-values ($p < .001$) for each measure. The substantial F-values, particularly for the Silhouette Coefficient ($F = 45.32$) and Calinski-Harabasz Index ($F = 52.43$), indicate strong discrimination between methods in terms of cluster quality, with effect sizes ($\eta^2 = 0.256$ and 0.284, respectively) suggesting that approximately one-quarter to one-third of the variance in these outcomes can be attributed to the clustering algorithm itself. While still statistically significant, the smaller effect size for Computation Time ($\eta^2 = 0.062$) indicates that although Ward's method is measurably slower, practical differences in runtime between methods are relatively modest compared to the substantial differences in clustering quality. These findings collectively underscore that methodological choice is a crucial determinant of clustering performance, with Ward's method generally outperforming others in partition quality metrics, while linkage-based methods show advantages in dendrogram preservation.

**Post-Hoc Tukey HSD Analysis**

Pairwise comparisons using Tukey's HSD test revealed significant differences between specific methods:

**Table 3.** Tukey HSD Pairwise Comparisons for Silhouette Coefficient

| Comparison | Mean Difference | 95% CI | p-value |
|---|---|---|---|
| **Ward's - Single** | 0.141 | [0.112, 0.170] | < .001 |
| **Ward's - Complete** | 0.049 | [0.020, 0.078] | .001 |
| **Ward's - Average** | 0.062 | [0.033, 0.091] | < .001 |
| **Complete - Single** | 0.092 | [0.063, 0.121] | < .001 |
| **Average - Single** | 0.079 | [0.050, 0.108] | < .001 |
| **Complete - Average** | 0.013 | [-0.016, 0.042] | .489 |

Table 3 presents the post-hoc Tukey HSD analysis provides detailed insight into the specific pairwise differences between clustering methods, revealing that Ward's method consistently outperforms all other approaches with statistically significant superiority in silhouette coefficient (all $p < .001$ compared to Single, Complete, and Average Linkage). While both Complete and Average Linkage also significantly surpass Single Linkage ($p < .001$), the lack of significant difference between them ($p = .489$) suggests comparable performance in cluster cohesion and separation. These results reinforce Ward's method as the optimal choice for partition quality while highlighting the persistent limitations of Single Linkage due to chaining effects, with the tight confidence intervals indicating precise estimation of these performance differences across simulation scenarios.

**Performance by Data Structure**

**Table 4.** Silhouette Coefficient Performance by Data Structure Type

| Data Type | Single Linkage | Complete Linkage | Average Linkage | Ward's Method | p-value |
|---|---|---|---|---|---|
| **Spherical Clusters** | 0.598 (0.052) | 0.756 (0.045) | 0.743 (0.043) | 0.821 (0.038) | < .001 |
| **Overlapping Clusters** | 0.523 (0.061) | 0.687 (0.053) | 0.664 (0.049) | 0.719 (0.046) | < .001 |
| **Varying Densities** | 0.641 (0.048) | 0.698 (0.044) | 0.731 (0.042) | 0.743 (0.041) | < .001 |
| **Elongated Clusters** | 0.743 (0.046) | 0.612 (0.055) | 0.683 (0.048) | 0.654 (0.050) | < .001 |
| **Noisy Data** | 0.605 (0.057) | 0.718 (0.049) | 0.689 (0.047) | 0.776 (0.042) | < .001 |

*Note: Values represent means with standard deviations in parentheses. ANOVA results show significant main effects for all data types*

Table 4 provides the performance trends across data structures in Table 2 reveal critical method-data interactions, as Ward's method dominates in spherical (0.821) and noisy (0.776) clusters due to its variance-minimization objective, while single linkage excels with elongated structures (0.743) by capturing chain-like patterns. Complete linkage shows robustness in overlapping clusters (0.687) by minimizing outlier influence, whereas average linkage delivers consistent mid-tier performance across most scenarios. Notably, all methods struggle with overlapping clusters (scores ≤ 0.719), reflecting inherent challenges in separating intertwined distributions, and the pronounced performance variability underscores that optimal method selection is deeply contingent on underlying data geometry.

**Discussion**

The comparative analysis reveals significant performance variations among hierarchical clustering methods, with clear dependencies on data characteristics. Ward's method demonstrated superior overall performance, particularly for spherical cluster structures, consistent with its minimum variance optimization criterion. The method's ability to produce compact, well-separated clusters makes it an excellent choice for datasets where such structure is expected.

Single linkage showed unique advantages for elongated cluster identification but suffered from chaining effects in noisy environments. This finding supports theoretical expectations and suggests its optimal use in applications where non-spherical cluster shapes are anticipated and data quality is high. Complete linkage exhibited robust performance against outliers and noise, making it suitable for datasets with quality concerns or unknown cluster properties.

Average linkage provided consistently balanced performance across scenarios, supporting its role as a general-purpose clustering method. While not optimal for any specific data type, its reliability across diverse conditions makes it valuable for exploratory analysis when cluster characteristics are unknown.

The computational efficiency analysis showed minimal differences among methods, with all approaches scaling similarly with dataset size. This finding suggests that performance quality should be the primary consideration in method selection rather than computational constraints for typical dataset sizes.

**Conclusion**

This comprehensive comparative study provides evidence-based guidelines for hierarchical clustering method selection. Ward's method emerges as the preferred choice for datasets with spherical cluster structures, while single linkage excels for elongated clusters in low-noise environments. Complete linkage offers robustness for uncertain data quality, and average linkage provides reliable general-purpose performance. Future research should explore method performance in high-dimensional settings and develop automated selection criteria based on data characteristics assessment.

**References**


1. Jain, A. K., Murty, M. N., & Flynn, P. J. (2019). Data clustering: A comprehensive review. *ACM Computing Surveys*, *51*(4), Article 1–46. https://doi.org/10.1145/3232165
2. Xu, D., & Tian, Y. (2020). Comprehensive survey of clustering algorithms. *Annals of Data Science*, *7*(2), 165–194.
3. Murtagh, F., & Contreras, P. (2021). Algorithms for hierarchical clustering: An overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *11*(1), Article e1395. https://doi.org/10.1002/widm.1395
4. Nielsen, F. (2019). Hierarchical clustering. In *Introduction to HPC with MPI for data science* (pp. 195–211). Springer.
5. Rokach, L., & Maimon, O. (2010). Clustering methods. In O. Maimon & L. Rokach (Eds.), *Data mining and knowledge discovery handbook* (2nd ed., pp. 321–352). Springer
6. Sneath, P. H., & Sokal, R. R. (1973). *Numerical taxonomy: The principles and practice of numerical classification*. W.H. Freeman.
7. Sokal, R. R., & Michener, C. D. (1958). A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin*, *38*, 1409–1438.
8. Day, W. H. E., & Edelsbrunner, H. (1984). Efficient algorithms for agglomerative hierarchical clustering methods. *Journal of Classification*, *1*(1), 7–24. https://doi.org/10.1007/BF01890115
9. Müllner, D. (2021). Modern hierarchical, agglomerative clustering algorithms. *Pattern Recognition Letters*, *145*, 98–107. https://doi.org/10.1016/j.patrec.2021.02.018
10. Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2019). *Cluster analysis* (5th ed.). John Wiley & SonsHastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7
11. Gower, J. C., & Ross, G. J. (1969). Minimum spanning trees and single linkage cluster analysis. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *18*(1), 54–64. https://doi.org/10.2307/2346439
12. Jackson, D. A., & Chen, Y. (2020). Robust clustering with single linkage. *Computational Statistics & Data Analysis*, *142*, 106825. https://doi.org/10.1016/j.csda.2019.106825
13. Defays, D. (1977). An efficient algorithm for a complete link method. *The computer journal*, *20*(4), 364-366.
14. Lance, G. N., & Williams, W. T. (1966). Computer programs for hierarchical polythetic classification. *The Computer Journal*, *9*(1), 60–64.
15. Kaufman, L., & Rousseeuw, P. J. (1990). *Finding groups in data: An introduction to cluster analysis*. John Wiley & Sons. https://doi.org/10.1002/9780470316801

16. Ward, J. H., Jr. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*(301), 236–244.
17. Murtagh, F., & Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? *Journal of Classification*, *31*(3), 274–295.
18. Strauss, T., & von Maltitz, M. J. (2019). Generalising Ward's method for use with Manhattan distances. *PLoS ONE*, *12*(1), Article e0168288. https://doi.org/10.1371/journal.pone.0168288
19. Milligan, G. W., & Cooper, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, *50*(2), 159–179
20. Gordon, A. D. (1999). *Classification* (2nd ed.). Chapman & Hall.
21. R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/
22. Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65.
23. Sokal, R. R., & Rohlf, F. J. (1962). The comparison of dendrograms by objective methods. Taxon, 11(2), 33–40. https://doi.org/10.2307/1217208
24. Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics*, *3*(1), 1–27.
25. Xu, D., & Tian, Y. (2020). Comprehensive survey of clustering algorithms. Annals of Data Science, 7(2), 165–194.