Building of Informatics, Technology and Science (BITS)
Volume 7, No 4, March 2026 Page: 2205−2217
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i4.9054

Stacking-Correspondence Analysis for Fuzzy Data: Computational Framework for Analyzing Complex Qualitative Survey Data

Disa Rahma Kirana, Irlandia Ginanjar*, Bertho Tantular

Faculty of Mathematics and Natural Sciences, Department of Statistics, Universitas Padjadjaran, Bandung, Indonesia
Email: [1]dissa21002@mail.unpad.ac.id, [2,*]irlandia@unpad.ac.id, [3]bertho@unpad.ac.id
Correspondence Author Email: irlandia@unpad.ac.id
Submitted: **28/12/2025**; Accepted: **05/03/2026**; Published**: 05/03/2026**

Abstract−Bandung Regency faces a significant challenge in achieving Sustainable Development Goal (SDG) 12, marked by a critically low score of 14.53 out of 100. Uniform policies are often ineffective due to regional diversity and uncertainty in categorical survey data, which reflects real-world conditions. This study aims to identify sub-district characteristics based on consumption and production patterns to provide precise policy recommendations. The research utilizes data from the 2024 Supporting Area Survey (SWP), covering 280 villages across 31 sub-districts. A computational framework combining stacking techniques and Correspondence Analysis for Fuzzy Data (CAFD) is implemented to analyze four qualitative variables representing key SDG 12 indicators: household waste disposal sites, waste/sewage recycling activities, river sewage pollution status, and the combined consumption-production patterns. The stacking phase transforms the multi-way data structure into a two-way structure, while CAFD effectively handles qualitative uncertainty using membership degrees. Analysis results indicate that two principal dimensions capture 73.35% of the total information variance and successfully identify 17 sub-district clusters with similar problem profiles. The fuzzy approach unveils multi-characteristic profiles, identifying both dominant and secondary traits. This research contributes a two-dimensional perceptual map, enabling the government to transition from generic policies to tailored interventions for each sub-district. This solution represents a concrete step toward improving the SDG 12 achievement score through data-driven strategic planning.

Keywords: Stacking; Correspondence Analysis for Fuzzy Data; Consumption and Production Patterns; SDGs; Goal 12

1. INTRODUCTION

The achievement of the Sustainable Development Goals (SDGs) has become a crucial, including the Bandung Regency Government, in formulating sustainable development policies [1]. This commitment is realized through the integration of 44 priority SDG indicators into the Regional Medium-Term Development Plan 2021–2026 [2]. Nevertheless, the evaluation indicates that Bandung Regency is still far from its targets, recording a score of 46.62 [3]. The most significant challenge lies in the achievement of SDG 12, namely Responsible Consumption and Production, which registers a critical score of 14.53 [3]. This extremely low figure positions SDG 12 as one of the lowest scores in Bandung Regency, highlighting the existence of structural problems that require immediate attention [4]. Accordingly, this research expects to provide a computational solution to map sub-district profiles accurately despite data uncertainty. This uncertainty arises from the subjectivity of human perception in survey responses as well as the statistical noise generated when aggregating disparate village-level datasets into sub-district representations.

The low score for SDG 12 reflects the actual conditions on the ground, showing an extreme gap between targets and reality, particularly across three priority indicators: waste management, garbage disposal, and the provision of adequate facilities [2]. Bandung Regency generates approximately 1,282 tons of household waste daily [5]. However, the volume of waste managed and disposed of at the final disposal site is only around 200 tons per day [5]. This failure is compounded by the ineffectiveness of uniform government intervention programs, such as the TPS3R program, where only 70 out of 160 units constructed remain active [6]. This situation underscores the failure of a "*one-size-fits-all*" policy approach, as it fails to align with the specific characteristics and problems of each sub-district [7].

Given the failure of uniform programs, there is a clear need to identify unique, multi-characteristic profiles for each sub-district based on three priority SDG 12 indicators [8]. This identification must consider the interdependence between categories, rather than merely examining variables in isolation [9]. This approach necessitates the use of a multivariate statistical method specifically designed for qualitative data, namely Correspondence Analysis (CA), as the primary focus is analyzing inter-category dependency [10]. However, standard CA is challenged by the inherent uncertainty of aggregated survey data, which originally form a complex multi-way contingency table [11].

To overcome these limitations, this study proposes a framework called Stacking-Correspondence Analysis for Fuzzy Data (Stacking-CAFD). This framework combines two methodological pillars: the Stacking method, which transforms the multi-way contingency table into a two-way super-matrix, essential for enabling the analysis [12]; and the Correspondence Analysis for Fuzzy Data (CAFD), which is an extension of CA capable of analyzing contingency tables with fuzzy membership degrees [13]. By combining these, Stacking-CAFD offers significant novelty by allowing it to provide information on the multi-characteristic profile of each sub-district. This more nuanced profile will serve as the foundation for the Bandung Regency Government to transition from ineffective general policies to adaptive interventions specific to each region, aiming to effectively improve the SDG 12 achievement score [14].

This research presents substantial novelty. Although studies on consumption and production patterns in Indonesia have been conducted, such as by Maulana et al. on 2023 [15], which focused on community empowerment in Suntenjaya Village using descriptive quantitative methods, or studies by Pinem on 2023, which identified sub-district characteristics in Bandung Regency but focused on SDG Goal 9 about slum areas [16], none have specifically examined the multi-indicator interdependence of SDG 12 in Bandung Regency using a multivariate statistical approach. Overall, the previous studies failed to capture the simultaneous interaction between qualitative variables at the sub-district level and ignored the presence of uncertainty in the data. While recent research by Warolemba et al. (2023) has addressed data uncertainty in Indonesian welfare indicators using Fuzzy C-Means, their application is primarily focused on object clustering [17].

The novelty of this research lies in the Stacking-CAFD framework. Theodorou et al. (2012) applied CAFD for fuzzy membership but lacked stacking for multi-way structures [13]. Furthermore, while Hwang et al. (2010) focused on clustering through Fuzzy Cluster MCA [18], this research prioritizes inter-category dependencies without clustering. Similarly, Chevenet et al. (1994) utilized fuzzy coding for ecological data but omitted stacking for complex contingency tables [19]. Consequently, this study is the first to explicitly integrate stacking with CAFD in an SDG 12 context, providing a robust framework for precise spatial mapping [13]. This innovation identifies nuanced sub-district profiles, enabling the Bandung Regency Government to implement data-driven interventions to improve SDG 12 scores [20]. Therefore, the primary objective of this study is to identify the multi-indicator interdependence of SDG 12 across 31 sub-districts in Bandung Regency. The research aims to produce a high-precision perceptual map that effectively captures complex consumption and production patterns while accommodating data uncertainty.

2. RESEARCH METHODOLOGY

2.1 Research Stages

This research is designed to process complex multi-way qualitative data characterized by high uncertainty using the Stacking-Correspondence Analysis for Fuzzy Data (Stacking-CAFD) framework. The research workflow commences with the identification of critical issues surrounding the low Sustainable Development Goal (SDG) 12 achievement scores in Bandung Regency, followed by the collection of secondary data from the 2024 Supporting Area Survey.

a. Data Pre-processing (Stacking): This stage involves transforming the multi-way contingency table into a two-way super-matrix to satisfy the technical requirements of the correspondence algorithm.
b. Independence Testing**:** A verification of the statistical dependency between variables is conducted using the Pearson Chi-Square test to ensure that the data structure is suitable for further analysis.
c. CAFD Implementation: Fuzzy logic is integrated by constructing Triangular Fuzzy Numbers (TFN) to effectively accommodate the inherent ambiguity and uncertainty within the survey data.
d. Decomposition and Mapping: Geometric information, including inertia values and principal coordinates, is extracted via Eigen Value Decomposition (EVD) for visualization on a low-dimensional perceptual map.
e. Defuzzification and Interpretation: The fuzzy output is converted into crisp representative coordinates to facilitate nuanced policy decision-making and validate the research expectations.

2.2 Data Sources

The data utilized in this study are secondary data sourced from the 2024 Supporting Area Survey (SWP) of Bandung Regency, obtained from the Dinas Komunikasi dan Informasi (Diskominfo) of Bandung Regency. The units of observation for this research are 270 villages and 10 urban wards. These variables were strategically selected based on the priority indicators for SDG Goal 12, as formally outlined in the Bandung Regency Regional Medium-Term Development Plan (RPJMD) for 2021-2026 [2]. The study involves four characteristic variables and 31 sub-districts, all measured on a nominal scale. The specific variables used in this research are further detailed in Table 1.

Table 1. Qualitative Data

No.	Variable	Category
1.	Sub-district (K)	1. Sub-district Arjasari (K_1) 2. Sub-district Baleendah (K_2) ⋮ 31. Sub-district Soreang (K_{31})
2.	The Majority of Household Waste Disposal Sites (X_1)	1. The majority of household waste disposal sites are adequate ($X_{1,1}$) 2. The majority of household waste disposal sites are inadequate ($X_{1,2}$)
3.	Existence of Waste/Sewage Processing/Recycling Activities (X_2)	1. There are waste/sewage recycling activities ($X_{2,1}$) 2. There are no waste/sewage recycling activities ($X_{2,3}$)
4.	River Sewage Pollution Status (X_3)	1. There is no river sewage pollution ($X_{3,1}$) 2. There is river sewage pollution ($X_{3,2}$)

2.3 Stacking

Stacking is a crucial data pre-processing stage before applying Correspondence Analysis for Fuzzy Data (CAFD). This method is selected based on two main justifications: Firstly, technical justification, CAFD requires the input data to be in the form of a two-way contingency table [13]. However, the original data consists of four qualitative variables, which naturally form a multi-way table. Stacking addresses this technical requirement. Secondly, conceptual justification, stacking transforms the perspective of the analysis to highlight the simultaneous interaction pattern among the three SDG 12 characteristic variables, making the analysis of inter-dependency more holistic [21]. The Stacking algorithm transforms the original multi-way contingency table into a two-way super-matrix $\left(\mathbf{N}_{q_1 \times q_{\tilde{X}}}\right)$ [12].

In this process, the object variable Sub-district (K) is maintained as the row variable ($q_1 = 31$). A new column variable, the Combined Characteristic ($\tilde{X}$), is formed by aggregating all $2 \times 2 \times 2$ combinations of the categories from the three characteristic variables: X_1, X_2, and X_3 (River Pollution Status). The total number of new column categories is $q_{\tilde{X}} = 8$ categories ($j = 1, 2, \dots, 8$). The resulting output matrix, $\mathbf{N}$, is a 31×8 contingency table where the element n_{ij} represents the frequency of villages/urban wards in Sub-district i that possess the characteristic combination j. The matrix $\mathbf{N}$, which serves as the input for the Chi-Square Test and CAFD, is represented in Table 2.

Table 1. Four-Way Contingency Table Before Stacking

Sub-district	$(X_{1,1})$				$(X_{1,2})$			
	$(X_{2,1})$		$(X_{2,2})$		$(X_{2,1})$		$(X_{2,2})$	
	$(X_{3,1})$	$(X_{3,2})$	$(X_{3,1})$	$(X_{3,2})$	$(X_{3,1})$	$(X_{3,2})$	$(X_{3,1})$	$(X_{3,2})$
1	n_{11}	n_{12}	n_{13}	n_{14}	n_{15}	n_{16}	n_{17}	n_{18}
2	n_{21}	n_{22}	n_{23}	n_{24}	n_{25}	n_{26}	n_{27}	n_{28}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
q_1	$n_{q_1 1}$	$n_{q_1 2}$	$n_{q_1 3}$	$n_{q_1 4}$	$n_{q_1 5}$	$n_{q_1 6}$	$n_{q_1 7}$	$n_{q_1 8}$

The two-way contingency table after the Stacking process, which is the super-matrix $\mathbf{N}$, is shown in Table 3.

Table 2. Two-Way Contingency Table After Stacking

Sub-district	$\tilde{X}_1$	$\tilde{X}_2$	$\tilde{X}_3$	$\tilde{X}_4$	$\tilde{X}_5$	$\tilde{X}_6$	$\tilde{X}_7$	$\tilde{X}_8$
1	n_{11}	n_{12}	n_{13}	n_{14}	n_{15}	n_{16}	n_{17}	n_{18}
2	n_{21}	n_{22}	n_{23}	n_{24}	n_{25}	n_{26}	n_{27}	n_{28}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
q_1	$n_{q_1 1}$	$n_{q_1 2}$	$n_{q_1 3}$	$n_{q_1 4}$	$n_{q_1 5}$	$n_{q_1 6}$	$n_{q_1 7}$	$n_{q_1 8}$

Notation:
$\tilde{X}_1 = \{X_{1,1},\ X_{2,1},\ X_{3,1}\}$ $\quad$ $\tilde{X}_5 = \{X_{1,2},\ X_{2,1},\ X_{3,1}\}$
$\tilde{X}_2 = \{X_{1,1},\ X_{2,1},\ X_{3,2}\}$ $\quad$ $\tilde{X}_6 = \{X_{1,2},\ X_{2,1},\ X_{3,2}\}$
$\tilde{X}_3 = \{X_{1,1},\ X_{2,2},\ X_{3,1}\}$ $\quad$ $\tilde{X}_7 = \{X_{1,2},\ X_{2,2},\ X_{3,1}\}$
$\tilde{X}_4 = \{X_{1,1},\ X_{2,2},\ X_{3,2}\}$ $\quad$ $\tilde{X}_8 = \{X_{1,2},\ X_{2,2},\ X_{3,2}\}$

The specific definitions and descriptions of the eight new categories for the Combined Characteristic ($\tilde{X}$), which resulted from the aggregation of SDG 12 indicators during the stacking process, are detailed in Table 4.

Table 3. New Categories for the Characteristic Variable ($\tilde{X}$) After Stacking

Category	Description
$\tilde{X}_1$	All indicators adequate/safe
$\tilde{X}_2$	There is river sewage pollution
$\tilde{X}_3$	There are no waste/sewage recycling activities
$\tilde{X}_4$	There is river sewage pollution and there are no waste/sewage recycling activities
$\tilde{X}_5$	The majority of household waste disposal sites are inadequate
$\tilde{X}_6$	The majority of household waste disposal sites are inadequate and there is river sewage pollution
$\tilde{X}_7$	The majority of household waste disposal sites are inadequate and there are no waste/sewage recycling activities
$\tilde{X}_8$	All indicators inadequate/unsafe

2.4 Contingency Table

The characteristic variables are transformed into a two-way contingency table where the object variable is used for the rows and the characteristic variables for the columns. The resulting contingency table is as presented in Table 5.

Table 4. Contingency Table

Sub-district	Characteristic Variable ($\tilde{X}$)						Total
	1	2	...	j	...	$q_{\tilde{X}}$	
1	n_{11}	n_{12}	$\cdots$	n_{1j}	$\cdots$	$n_{1q_{\tilde{X}}}$	$n_{1\bullet}$
2	n_{21}	n_{22}	$\cdots$	n_{2j}	$\cdots$	$n_{2q_{\tilde{X}}}$	$n_{2\bullet}$

Sub-district	Characteristic Variable ($\tilde{X}$)						
	1	2	...	j	...	$q_{\tilde{X}}$	Total
⋮	⋮	⋮	⋱	⋮	⋱	⋮	⋮
i	n_{i1}	n_{i2}	⋯	n_{ij}	⋯	$n_{iq_{\tilde{X}}}$	$n_{i\bullet}$
⋮	⋮	⋮	⋱	⋮	⋱	⋮	⋮
q_1	$n_{q_1 1}$	$n_{q_1 2}$	⋯	$n_{q_1 j}$	⋯	$n_{q_1 q_{\tilde{X}}}$	$n_{q_1 \bullet}$
Total	$n_{\bullet 1}$	$n_{\bullet 2}$	⋯	$n_{\bullet j}$	⋯	$n_{\bullet q_{\tilde{X}}}$	n

Here, q_1 is the number of categories for the row variable (sub-district), and $q_{\tilde{X}}$ is the number of categories for the column variable (combined consumption and production pattern characteristics), where ($\tilde{X} = 1, 2, \dots, 8$), while n represents the total number of observations in the data.

2.5 The Chi-Square Independence Test

The Chi-Square Independence Test is a crucial verification step to determine whether there is a significant dependence between the row variable, Sub-district (*K*), and the column variable, the Combined Characteristic ($\tilde{X}$) [22]. If the variables are found to be statistically independent (H_0 is accepted), the CAFD cannot be logically pursued. The hypotheses used in the independence test based on the Pearson Chi-Square statistic are as follows [23]:

$H_0 : \pi_{ij} = \pi_{i\bullet} \bullet \pi_{\bullet j}$ (The row and column categorical variables are mutually independent)
$H_1 : \pi_{ij} \neq \pi_{i\bullet} \bullet \pi_{\bullet j}$ (The row and column categorical variables are mutually dependent)

In this study, a significance level of $\alpha = 5\% = 0.05$ is adopted. A significance level of 0.05 is used considering that the case being analyzed is intended for government policy formulation, thus a confidence level of 95% is deemed appropriate [24]. The independence test is performed on the Contingency Matrix **N** resulting from the Stacking process. Based on the structure in Table 5, the cross-tabulation matrix **N** is defined by the following equation.

$$\mathbf{N}_{(q_1 \times q_{\tilde{X}})} = \left(n_{ij}\right);\ n_{ij} \geq 0,\ \ i = 1, 2, \dots, q_1 \text{ and } j = 1, 2, \dots, q_{\tilde{X}} \tag{1}$$

Based on Equation (1), the matrix of joint probabilities $\widetilde{\mathbf{P}}$ can be obtained. This matrix satisfies the property that the sum of the elements in each row is 1, defined as follows:

$$\widetilde{\mathbf{P}}_{q_1 \times q_{\tilde{X}}} = \left(\frac{1}{n}\right)\mathbf{N} = \left(\frac{n_{ij}}{n}\right) = \left(\tilde{p}_{ij}\right) \tag{2}$$

The total marginal values for the rows and columns of matrix $\widetilde{\mathbf{P}}$ are vectors $\tilde{\mathbf{p}}_{i\bullet}$ and $\tilde{\mathbf{p}}_{\bullet j}$. Let the vector of the row marginal distribution (row mass) be $\tilde{\mathbf{p}}_{i\bullet}$, and the vector of the column marginal distribution (column mass) be $\tilde{\mathbf{p}}_{\bullet j}$, marginal values for the rows and columns are is defined by the following equation [25].

$$\tilde{\mathbf{p}}_{i\bullet(i\times 1)} = \left(\tilde{p}_{i\bullet}\right) \text{ where } \tilde{p}_{i\bullet} = \textstyle\sum_{j=1}^{q_{\tilde{X}}} \tilde{p}_{ij} \tag{3}$$

$$\tilde{\mathbf{p}}_{\bullet j_{(j\times 1)}} = \left(\tilde{p}_{\bullet j}\right) \text{ where } \tilde{p}_{\bullet j} = \textstyle\sum_{i=1}^{q_1} \tilde{p}_{ij} \tag{4}$$

Thus, the assumption of independence can be expressed by the following equation:

$$\tilde{p}_{ij} = \tilde{p}_{i\bullet}\, \tilde{p}_{\bullet j}\,;\ \ i = 1, 2, \dots, q_1\,;\ \ j = 1, 2, \dots, q_{\tilde{X}}$$

The joint probability π_{ij} is estimated by $\hat{\pi}_{ij} = \tilde{p}_{ij}$, where $\tilde{p}_{ij}$ is the estimate of the probability (proportion) as shown in Equation (2) and the marginal proportions $\tilde{p}_{i\bullet}$ and $\tilde{p}_{\bullet j}$ are given by Equations (3) and (4). The test statistic for the chi-square test is therefore obtained as follows [25]:

$$\chi^2 = n \textstyle\sum_{i=1}^{q_1} \sum_{j=1}^{q_{\tilde{k}}} \frac{\left(\tilde{p}_{ij} - \tilde{p}_{i\bullet}\, \tilde{p}_{\bullet j}\right)^2}{\tilde{p}_{i\bullet}\, \tilde{p}_{\bullet j}} \tag{5}$$

where $\pi_{i\bullet}$ is the marginal probability of the object variable (row), $\pi_{\bullet j}$ is the marginal probability of the characteristic variable (column), π_{ij} is the joint probability of the object and characteristic variables, $\tilde{p}_{ij}$ is the joint proportion of the i-th row and j-th column, $\tilde{p}_{i\bullet}$ is the marginal proportion of the i-th row, and $\tilde{p}_{\bullet j}$ is the marginal proportion of the j-th column. The Pearson Chi-Square test criterion with $\chi^2 \geq \chi^2_{\alpha(q_1-1)(q_{\tilde{X}}-1)}$ is to reject H_0 if the p-value < α, accept otherwise. The rejection of H_0 is identified based on the p-value using the following equation.

$$p - value = P\left\{\chi^2_{(q_1-1)(q_{\tilde{X}}-1)} \geq \chi^2\right\} \tag{6}$$

Based on Equation (6), if the conclusion indicates that there is a dependency between the two qualitative variables, the Correspondence Analysis for Fuzzy Data can proceed, utilizing the Stacking results ($\tilde{X}$) to obtain the principal coordinates for the rows and columns.

2.6 Correspondence Analysis for Fuzzy Data (CAFD)

Correspondence Analysis for Fuzzy Data (CAFD) is an extension of Correspondence Analysis (CA) specifically chosen to analyze the dependency among categories of qualitative variables when the data contain uncertainty [13]. The CAFD method allows each observational unit to have a degree of membership in several categories, rather than having absolute membership in only one category [19]. The primary input for CAFD is the Fuzzy Contingency Table, where each frequency element $(\tilde{n}_{ij})$ is represented by a Triangular Fuzzy Number (TFN), denoted by $\tilde{n}_{ij} = \left(\tilde{n}_{(1)ij}, \tilde{n}_{(2)ij}, \tilde{n}_{(3)ij}\right)$. The central value $\tilde{n}_{(2)ij}$ is the observed frequency, while the lower bound $\tilde{n}_{(1)ij}$ and upper bound $\tilde{n}_{(3)ij}$ are obtained through the Binomial Confidence Interval (Agresti-Coull) using Equation (7) to get the TFNs:

$$\tilde{n}_{(1)ij} = n_{i\bullet} \times p_{lower} \text{ and } \tilde{n}_{(3)ij} = n_{i\bullet} \times p_{upper}. \tag{7}$$

Next, the fuzzy contingency table is normalized into a fuzzy probability matrix (**P**) with elements $p_{ij} = \left(p_{(1)ij}, p_{(2)ij}, p_{(3)ij}\right)$. The elements of the **P** matrix represent the fuzzy proportion relative to the row total $(n_{i\bullet})$ [13]. The **P** matrix has a size of $q_1 \times q_{\tilde{X}}$ calculated according to the following element-wise equation, ensuring that the range of uncertainty is fully represented in the results [13]:

$$p_{(1)ij} = \frac{n_{(1)ij}}{n_{(3)i\bullet}} = \frac{n_{(1)ij}}{\sum_j n_{(3)ij}}; \quad p_{(2)ij} = \frac{n_{(2)ij}}{n_{(2)i\bullet}} = \frac{n_{(2)ij}}{\sum_j n_{(2)ij}}; \quad p_{(3)ij} = \frac{n_{(3)ij}}{n_{(1)i\bullet}} = \frac{n_{(3)ij}}{\sum_j n_{(1)ij}} \tag{8}$$

The subsequent step is to calculate the diagonal matrices of the fuzzy marginal distributions, namely the row mass $(\mathbf{D}_r)$ and the column metric $(\mathbf{D}_c)$ respectively, according to the following equations [13]:

$$\mathbf{D}_{r(q_1 \times q_1)} = diag(\mathbf{r}_{i\bullet}) \tag{9}$$

$$\mathbf{D}_{c(q_{\tilde{k}} \times q_{\tilde{k}})} = diag\left(\boldsymbol{c}_{\bullet j}\right) = diag\left(\frac{1}{p_{\bullet j}}\right) \tag{10}$$

The elements of the column metric $\boldsymbol{c}_{\bullet j}$ are the inverse of the column mass $p_{\bullet j}$. The row mass $\mathbf{r}_{i\bullet}$ and the column metric $\boldsymbol{c}_{\bullet j}$ are computed element-wise for the TFNs. From these computed components, the fuzzy matrix (**S**) is constructed. The fuzzy matrix. The fuzzy matrix (**S**) describes the association structure between categories in a fuzzy form and serves to extract geometric information in the form of fuzzy eigenvalues and fuzzy eigenvectors. The calculation of the fuzzy matrix **S** can be written as follows in Equation (11) [13]:

$$\mathbf{S}_{(q_{\tilde{k}} \times q_{\tilde{k}})} = \left(\mathbf{P}^T_{(q_{\tilde{k}} \times q_1)} \mathbf{D}_{r(q_1 \times q_1)} \mathbf{P}_{(q_1 \times q_{\tilde{k}})}\right) \mathbf{D}_{c(q_{\tilde{k}} \times q_{\tilde{k}})} \tag{11}$$

The matrix **S** is generally non-symmetric, making it complex to decompose directly. However, the problem of decomposing matrix **S** can be simplified by transforming it into an equivalent and symmetric matrix **A.** Since the symmetric matrix **A** is easier to analyze and guarantees stable solutions, the Eigen Value Decomposition (EVD) is applied to this matrix. The matrix **S** is equivalent to the corresponding matrix **A** as shown in Equation (12) [13]:

$$\mathbf{A}_{(q_{\tilde{k}} \times q_{\tilde{k}})} = \mathbf{D}^{\frac{1}{2}}_{c(q_{\tilde{k}} \times q_{\tilde{k}})} \left(\mathbf{P}^T_{(q_{\tilde{k}} \times q_1)} \mathbf{D}_{r(q_1 \times q_1)} \mathbf{P}_{(q_1 \times q_{\tilde{k}})}\right) \mathbf{D}^{\frac{1}{2}}_{c(q_{\tilde{k}} \times q_{\tilde{k}})} \tag{12}$$

In this analysis, the formed symmetric fuzzy matrix **A** must be transformed into a single crisp matrix representation to be processed in the Eigen Value Decomposition (EVD) stage. This transformation process is carried out through a systematic two-step mechanism. The first step is the η-cut, which functions as a mechanism to manage the level of uncertainty. By setting an η value between 0 and 1, every fuzzy element in matrix **A** is converted into a crisp interval. The application of the η-cut on matrix **A** yields an interval matrix, denoted as $\mathbf{A}_\eta$, with left and right bounds according to Equation (13) [13]:

$$\mathbf{A}_{(q_{\tilde{k}} \times q_{\tilde{k}})\eta} = [\mathbf{A}_{\eta(p_{(1)ij})}, \mathbf{A}_{\eta(p_{(3)ij})}] = \left[a_{(1)ij} + \eta\left(a_{(2)ij} - a_{(1)ij}\right),\ a_{(3)ij} - \eta\left(a_{(3)ij} - a_{(2)ij}\right)\right] \tag{13}$$

The second step, θ-interpolation, serves to select a single representative point from within the interval established by the η-cut. The parameter θ, also valued between 0 and 1, acts as a weighting factor determining the position of the final point within that interval. A choice of $\theta = 0$ will select the left bound of the interval, while $\theta = 1$ elects the right bound. These two steps allow the conversion of the fuzzy matrix **A** into a crisp matrix $\mathbf{A}(\eta, \theta)$ ready for analysis in the EVD stage. The θ-interpolation for the two interval bounds is presented below [13]:

$$\mathbf{A}_{(q_{\tilde{k}} \times q_{\tilde{k}})\eta} = \mathbf{A}(\eta, \theta) = \left\{(1 - \theta)\mathbf{A}_{\eta(p_{(1)ij})} + \theta \mathbf{A}_{\eta(p_{(3)ij})}\right\} \tag{14}$$

This approach is highly effective because the eigenvalues generated from matrix **A** are identical to the eigenvalues of the original matrix **S** where **S** and **A** have the same eigenvalues (λ), and the eigenvectors u from **S** and w from **A** have the following relationship: $u = \mathbf{D}^{\frac{1}{2}}_{c(q_{\tilde{k}} \times q_{\tilde{k}})} w$ [13]. Thus, it is often simpler to diagonalize **A** and then compute u from the relationship above.

The eigenvalues (λ_ℓ) extracted from matrix **S** are called inertia values, representing the magnitude of the data variation (information) carried by each dimension. The process of decomposition is carried out to obtain these eigenvalues. The decomposition method used in this research is the Eigen Value Decomposition (EVD), which aims to determine the eigenvalues matrix ($\mathbf{\Lambda}$) and eigenvectors ($\mathbf{u}$) of the fuzzy association matrix **S**. Therefore, the eigen values decomposition is presented below [13]:

$$\mathbf{S}u = \mathbf{\Lambda}u \tag{15}$$

Where $u^T = \left(u_1, u_2, \dots, u_{q_{\tilde{X}}}\right)$ is the fuzzy eigenvector, and $\mathbf{\Lambda}$ is the diagonal matrix of the fuzzy eigenvalues. After the eigenvalues and eigenvectors are obtained from the decomposition process, the principal coordinates for the object categories can be obtained using Equation (16) as follows [13]:

$$\mathbf{F} = \mathbf{P}\mathbf{D}_{\mathbf{c}}^{\frac{1}{2}} w_\ell = (\boldsymbol{f}_1, \boldsymbol{f}_2 \dots, \boldsymbol{f}_\ell) = (f_{i\ell}) \tag{16}$$

Where **P** is the fuzzy correspondence matrix, $\mathbf{D}_c$ is the diagonal row mass matrix as in Equation (10), and w_ℓ a is the eigenvector of **A** with $\ell = 1, \dots, L$. Each row of the matrix **F** reflects a category, while each column represents the coordinate for each dimension. $\boldsymbol{f}_\ell$ is the principal coordinate vector of the ℓ-th dimension, and $f_{i\ell}$ are the elements of the principal coordinates.

Inertia represents the percentage of total variation and is used as a measure of the quality of the mapping [26]. An inertia value indicating sufficiently good mapping quality is generally set at a minimum of 70% [27]. The inertia for D dimension where $D \leq L$ or the cumulative percentage of variance (diversity), can be expressed by Equation (17) [28]:

$$\tau_D = \left(\frac{\lambda_1 + \lambda_2 + \dots + \lambda_d}{\sum_{\ell=1}^{L} \lambda_\ell}\right) = \left(\frac{\sum_{d=1}^{D} \lambda_d}{\sum_{\ell=1}^{L} \lambda_\ell}\right) \tag{17}$$

The inertia values indicate the contribution of the k-th row to the total inertia. Once the inertia values are obtained, the variance explained by each dimension can be determined. The variance explained by the D-th dimension depends on the percentage contribution ϕ_d of the respective eigenvalue. The variance coverage for each dimension can be denoted as in Equation (18) [28]:

$$\phi_d = \left(\frac{\lambda_d}{\sum_{\ell=1}^{L} \lambda_\ell}\right) \tag{18}$$

Where ϕ_d the variance coverage of the d-th dimension, λ_d is the d-th eigenvalue obtained from EVD, and the denominator is the total inertia. If the cumulative inertia percentage in two or three dimensions reaches 70%, the resulting two-or three-dimensional map is considered to represent the data adequately [29]. The representation of dependency obtained from the principal coordinates is based on the distances between categories [30]. To ensure that the represented dependency is objective, information is extracted based on distance metrics.

To allow the CAFD results to be visualized in a low-dimensional map, a final step, defuzzification, is required. A robust fuzzy analysis method must be capable of transforming uncertain outputs into a concrete decision [31]. Therefore, this stage aims to convert the fuzzy coordinates into a representative set of crisp coordinates for each category, enabling the creation of an analyzable correspondence map.

a. *Method of Centroid*

In this study, the Centroid method will be used for defuzzification, utilizing the values $\eta = \frac{1}{4}, \frac{1}{2},$ and $\frac{3}{4}$ [13].

$$\mathbf{F}_{Centroid\ a} = \mathbf{P}_{\left(\eta=\eta, \theta=\frac{1}{2}\right)} \left(\mathbf{D}_{\boldsymbol{c}}^{\frac{1}{2}} w_{\boldsymbol{\ell}}\right)_{(\eta=1, \theta=0)} \tag{19}$$

$$\mathbf{F}_{Centroid\ b} = \left(\mathbf{P}\mathbf{D}_{\boldsymbol{c}}^{\frac{1}{2}} w_{\boldsymbol{\ell}}\right)_{\left(\eta=\eta, \theta=\frac{1}{2}\right)} \tag{20}$$

b. *Method of Arithmetic Mean*

$$\mathbf{F}_{Arithmetic\ Mean} = \frac{1}{(\nu+1)(\mu+1)} \sum\nolimits_{\eta,\theta \in [0,1]} \left\{\mathbf{P}_{(\eta,\theta)} \mathbf{D}_{\boldsymbol{c}(\eta,\theta)}^{\frac{1}{2}} w_{\boldsymbol{\ell}(\eta,\theta)}\right\} \tag{21}$$

Where, $\eta = \frac{0}{\nu}, \frac{1}{\nu}, \dots, \frac{\nu}{\nu}$, $\theta = \frac{0}{\mu}, \frac{1}{\mu}, \dots, \frac{\mu}{\mu}$, and $\nu, \mu \in \mathbb{N}^*$. In this research, for the Arithmetic Mean method, the final crisp coordinate is calculated using the following parameters:$(\eta, \theta) = (0,0), (0,1), (1,0), (1,1)$. This defuzzification process facilitates easier interpretation of the relationships between categories by presenting the results in a factor map similar to standard CA.

To ensure comprehensive interpretability, this study implements eight defuzzification scenarios as a stringent form of sensitivity analysis. This approach evaluates the stability of the perception map by establishing the SCA results as the critical baseline that ignores uncertainty ($\eta = 1$), The Arithmetic Mean method is then used to obtain a holistic view by averaging the sharpest boundaries to consider the entire possibility space. Finally, the Centroid method is applied using $\theta = \frac{1}{2}$ and varying η at three different confidence levels: high $\left(\eta = \frac{3}{4}\right)$, moderate $\left(\eta = \frac{1}{2}\right)$, and low $\left(\eta = \frac{1}{4}\right)$. By comparing the consistency of the association patterns across these eight scenarios, the research can convincingly validate the findings, demonstrating that the drawn conclusions remain robust under various levels of data uncertainty.

2.7 Euclidean Distance

If the two dimensional map generated from the CAFD analysis does not achieve 70% cumulative variance, the Euclidean distance matrix can be used to identify the characteristics of the object variables with the column categories. Euclidean distance is used to measure the degree of data similarity [34]. The smaller the distance, the more similar the characteristics are, and vice versa. The Euclidean distance between the principal coordinate vector $\mathbf{f}_{i\ell} = (f_{i1}, f_{i2}, \dots, f_{iL})$ and $\mathbf{f}_{j\ell} = \left(f_{j1}, f_{j2}, \dots, f_{jL}\right)$ can be calculated using the formula presented in Equation (22) [32].

$$d(\mathbf{f}_{i\ell}, \mathbf{f}_{j\ell}) = \sqrt{\left(\mathbf{f}_{i\ell} - \mathbf{f}_{j\ell}\right)^T \left(\mathbf{f}_{i\ell} - \mathbf{f}_{j\ell}\right)} \tag{22}$$

Where $\mathbf{f}_{i\ell}$ is the principal coordinate vector of the i-th sub-district in the ℓ-th dimension, and $\mathbf{f}_{j\ell}$ is the principal coordinate vector of the *j*-th qualitative category in the ℓ-th dimension.

3. RESULT AND DISCUSSION

This section presents the comprehensive results of the Stacking-CAFD computational framework applied to SDG 12 indicators in Bandung Regency. The analysis begins with statistical validation and proceeds through various fuzzy defuzzification scenarios to generate a robust spatial perception map.

3.1 Prerequisite Statistical Validation

The initial step is to verify the relationship between the sub-district (K) and the combined characteristics resulting from the stacking process ($\tilde{X}$) using the Pearson Chi-Square test, as performed according to Equations (5) and (6). The results of the Pearson Chi-Square test, utilizing R software and its packages, are presented in Table 6.

Table 5. Pearson Chi-Square Test Results

Qualitative Variable	χ^2	df	*p-value*	Decision
Sub-district vs X_1	94,34	30	$1{,}42 \times 10^{-8}$	H₀ rejected
Sub-district vs X_2	48,25	30	0,01873	H₀ rejected
Sub-district vs X_3	55,61	30	0,00303	H₀ rejected
Sub-district vs $\tilde{X}$	339,84	210	$3{,}46 \times 10^{-8}$	H₀ rejected

Based on the Pearson Chi-Square test at a 5% significance level, the results indicate that all variables lead to the decision to reject $\boldsymbol{H_0}$. This implies that all characteristic variables exhibit dependency on the Sub-district variable ($\boldsymbol{K}$). The original variables that show dependency on the Sub-district variable include: The Majority of Household Waste Disposal Sites ($\boldsymbol{X_1}$),the Existence of Waste/Sewage Processing/Recycling Activities ($\boldsymbol{X_2}$), and the River Sewage Pollution Status ($\boldsymbol{X_3}$). Consequently, all original variables can be included in the subsequent analysis. Similarly, the stacked variable ($\tilde{\boldsymbol{X}}$) is also found to be dependent on the Sub-district variable.

3.2 Application Implementation

Following the validation, the framework proceeds to implement the CAFD algorithm. This implementation captures the multi-dimensional interactions by calculating principal coordinates across eight different defuzzification scenarios.

3.2.1 Inertia and Dimensionality Analysis

To effectively visualize the categorical interactions among the characteristic variables, the Correspondence Analysis for Fuzzy Data (CAFD) is continued using the stacked variable ($\tilde{X}$). Once the prerequisite test is satisfied, the analysis proceeds to the CAFD stage, starting with the calculation of Equations (7). The CAFD yields the principal coordinates, calculated using various defuzzification scenarios according to Equations (16), (19), (20), and (21).

$$\mathbf{F}_{\mathrm{Kec}\,(31\times2)} = \begin{bmatrix} 0{,}2205 & 0{,}5319 \\ -0{,}4497 & 0{,}3415 \\ -0{,}0015 & 0{,}3486 \\ \vdots & \vdots \\ 0{,}4214 & 1{,}1998 \end{bmatrix}$$

$$\mathbf{F}_{\text{Char\_SCA}(8\times2)}=\begin{bmatrix}0{,}2156 & 0{,}1941\\ 0{,}9491 & 0{,}0806\\ -0{,}2267 & -1{,}655\\ \vdots & \vdots\\ -0{,}4106 & 0{,}1841\end{bmatrix}\quad \mathbf{F}_{\text{Char\_AM}\,(8\times2)}=\begin{bmatrix}-0{,}0670 & -0{,}4321\\ -0{,}3853 & -0{,}0135\\ -0{,}2338 & 1{,}3989\\ \vdots & \vdots\\ 3{,}4649 & -2{,}443\end{bmatrix}$$

$$\mathbf{F}_{\text{Char\_Ca\_}\eta=\frac{1}{4}(8\times2)}=\begin{bmatrix}0{,}3421 & -0{,}2116\\ 0{,}9370 & -0{,}0638\\ 0{,}1403 & 1{,}9189\\ \vdots & \vdots\\ -0{,}0159 & -0{,}5715\end{bmatrix}\quad \mathbf{F}_{\text{Char\_Cb\_}\eta=\frac{1}{4}(8\times2)}=\begin{bmatrix}-0{,}0923 & -0{,}4931\\ -0{,}5776 & -0{,}0276\\ -0{,}2644 & 0{,}6334\\ \vdots & \vdots\\ 3{,}7092 & -2{,}5774\end{bmatrix}$$

$$\mathbf{F}_{\text{Char\_Ca\_}\eta=\frac{1}{2}(8\times2)}=\begin{bmatrix}0{,}2999 & -0{,}2058\\ 0{,}9410 & -0{,}0694\\ 0{,}0180 & 1{,}8310\\ \vdots & \vdots\\ -0{,}1475 & -0{,}4424\end{bmatrix}\quad \mathbf{F}_{\text{Char\_Cb\_}\eta=\frac{1}{2}(8\times2)}=\begin{bmatrix}-0{,}0923 & -0{,}4297\\ -0{,}4988 & -0{,}0533\\ -0{,}1956 & 0{,}4615\\ \vdots & \vdots\\ 2{,}5823 & -1{,}8661\end{bmatrix}$$

$$\mathbf{F}_{\text{Char\_Ca\_}\eta=\frac{3}{4}(8\times2)}=\begin{bmatrix}0{,}2578 & -0{,}1999\\ 0{,}9451 & -0{,}0750\\ -0{,}1043 & 1{,}7423\\ \vdots & \vdots\\ -0{,}2791 & -0{,}3133\end{bmatrix}\quad \mathbf{F}_{\text{Char\_Cb\_}\eta=\frac{3}{4}(8\times2)}=\begin{bmatrix}-0{,}0883 & -0{,}3589\\ -0{,}4137 & -0{,}0722\\ -0{,}1240 & 0{,}3064\\ \vdots & \vdots\\ 1{,}6053 & -1{,}2307\end{bmatrix}$$

To determine the reliability of the correspondence map in representing the original data, an inertia analysis must be performed. Inertia analysis is calculated using the eigenvalues (λ_ℓ) for each dimension obtained from Equation (15). The percentage of variance coverage(ϕ_d) is calculated using Equation (18) and accumulated into the cumulative variance (τ_D) using Equation (17) to validate the appropriateness of the number of dimensions used, and the results summarized in Table 7.

Table 6. Eigenvalues, Variance Explained, and Cumulative Diversity

ℓ	λ_ℓ	ϕ_d	τ_D
1	118,08	48,94	48,94
2	58,87	24,40	73,35
3	36,73	15,22	88,57
$\vdots$	$\vdots$	$\vdots$	$\vdots$
30	$-1{,}56 \times 10^{-13}$	$-6{,}48 \times 10^{-14}$	100

Based on Table 7, the optimal number of dimensions chosen is $D = 2$. This decision is based on the established criterion that the cumulative variance must be ≥ 70%. The first two dimensions successfully explain 73.35% of the total inertia, making them an adequate representation for visualization and interpretation.

3.2.2 Multi-Perspective Defuzzification

To identify the specific degree of dependency between the Sub-district (K) and the combined characteristics resulting from stacking ($\tilde{X}$) the Euclidean distance is calculated between each sub-district and the principal coordinates derived from eight different defuzzification methods. This yields eight sets of distance measures. Three out of the eight sets of distance measures obtained according to Equation (22) are displayed in Table 8 (SCA), Table 9 (Arithmetic Mean), Table 10 (Centroid a).

Table 7. Euclidean Distance Results from the SCA Method

Sub-district	$\tilde{X}_1$	$\tilde{X}_2$	$\tilde{X}_3$	$\tilde{X}_4$	$\tilde{X}_5$	$\tilde{X}_6$	$\tilde{X}_7$	$\tilde{X}_8$
K_1	0,36	0,86	2,23	0,77	1,01	1,01	0,99	0,72
K_2	0,68	1,42	2,01	1,08	0,46	0,34	0,33	0,16
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
K_{31}	1,03	1,24	2,93	1,39	1,63	1,56	1,41	1,31

Table 8. Euclidean Distance Results from the Arithmetic Mean Method

Sub-district	$\tilde{X}_1$	$\tilde{X}_2$	$\tilde{X}_3$	$\tilde{X}_4$	$\tilde{X}_5$	$\tilde{X}_6$	$\tilde{X}_7$	$\tilde{X}_8$
K_1	1,01	0,82	0,98	4,95	0,15	2,78	0,84	4,39
K_2	0,86	0,36	1,08	4,26	0,60	3,24	1,52	4,79
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
K_{31}	1,70	1,46	0,68	5,34	0,84	2,09	0,96	4,73

Table 9. Euclidean Distance Results from the Centroid a Method with $\eta = \frac{1}{2}$

Sub-district	$\tilde{X}_1$	$\tilde{X}_2$	$\tilde{X}_3$	$\tilde{X}_4$	$\tilde{X}_5$	$\tilde{X}_6$	$\tilde{X}_7$	$\tilde{X}_8$
K_1	0,74	0,94	1,31	1,34	0,83	1,24	1,43	1,04
K_2	0,93	1,45	1,56	1,98	0,26	0,69	0,92	0,84

Sub-district	$\tilde{X}_1$	$\tilde{X}_2$	$\tilde{X}_3$	$\tilde{X}_4$	$\tilde{X}_5$	$\tilde{X}_6$	$\tilde{X}_7$	$\tilde{X}_8$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
K_{31}	1,41	1,37	0,75	1,46	1,44	1,87	2,08	1,74

Based on Figures 1, 2, and 3 (the correspondence maps) and the Euclidean distance results, the primary characteristics for each sub-district are derived across all eight defuzzification methods, as summarized in Table 11.

Table 11. Sub-district Characteristics Based on Eight Deffuzification Method

No	K	AM	SCA	$\text{Ca}_{\eta=\frac{1}{4}}$	$\text{Ca}_{\eta=\frac{1}{2}}$	$\text{Ca}_{\eta=\frac{3}{4}}$	$\text{Cb}_{\eta=\frac{1}{4}}$	$\text{Cb}_{\eta=\frac{1}{2}}$	$\text{Cb}_{\eta=\frac{3}{4}}$
1	K_1	$\tilde{X}_5$	$\tilde{X}_1$	$\tilde{X}_1$	$\tilde{X}_1$	$\tilde{X}_1$	$\tilde{X}_5$	$\tilde{X}_5$	$\tilde{X}_5$
2	K_2	$\tilde{X}_2$	$\tilde{X}_8$	$\tilde{X}_5$	$\tilde{X}_5$	$\tilde{X}_5$	$\tilde{X}_5$	$\tilde{X}_3$	$\tilde{X}_3$
3	K_3	$\tilde{X}_5$	$\tilde{X}_1$	$\tilde{X}_5$	$\tilde{X}_5$	$\tilde{X}_5$	$\tilde{X}_5$	$\tilde{X}_5$	$\tilde{X}_3$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
31	K_{31}	$\tilde{X}_3$	$\tilde{X}_1$	$\tilde{X}_3$	$\tilde{X}_3$	$\tilde{X}_3$	$\tilde{X}_5$	$\tilde{X}_5$	$\tilde{X}_6$

The CAFD approach CAFD enables multi-perspective uncertainty representation through defuzzification, generating realistic regional profiles via Triangular Fuzzy Numbers (TFN). Table 12 summarizes these potential characteristics for each sub-district, with the most frequent (dominant) traits marked in bold for easier interpretation.

Table 12. Multi-Characteristic Profile of Sub-districts Based on Consumption and Production Patterns

Groups	Sub-district	Consumption and Production Pattern Characteristics
1	Arjasari (K_1) Pacet (K_{23}) Pangalengan (K_{25})	All indicators adequate/safe ($\tilde{X}_1$) The majority of household waste disposal sites are inadequate ($\tilde{X}_5$)
	Kertasari (K_{17})	All indicators adequate/safe ($\tilde{X}_1$) The majority of household waste disposal sites are inadequate ($\tilde{X}_5$)
2	Rancaekek (K_{29}) Solokanjeruk (K_{30})	All indicators adequate/safe ($\tilde{X}_1$) There are no waste/sewage recycling activities ($\tilde{X}_3$) All indicators inadequate/unsafe ($\tilde{X}_8$)
3	Ciwidey (K_{13})	All indicators adequate/safe ($\tilde{X}_1$) There is river sewage pollution ($\tilde{X}_2$) There are no waste/sewage recycling activities ($\tilde{X}_3$) All indicators inadequate/unsafe ($\tilde{X}_8$)
4	Cileunyi (K_9) Dayeuhkolot (K_{14})	All indicators adequate/safe ($\tilde{X}_1$) There is river sewage pollution there are no waste/sewage recycling activities ($\tilde{X}_4$)
5	Ibun (K_{15})	All indicators adequate/safe ($\tilde{X}_1$) There is river sewage pollution ($\tilde{X}_2$) There is river sewage pollution there are no waste/sewage recycling activities ($\tilde{X}_4$)
6	Paseh (K_{26})	All indicators adequate/safe ($\tilde{X}_1$) There is river sewage pollution ($\tilde{X}_2$) There is river sewage pollution there are no waste/sewage recycling activities ($\tilde{X}_4$) All indicators inadequate/unsafe ($\tilde{X}_8$)
7	Katapang (K_{16}) Kutawaringin (K_{18}) Margahayu (K_{21}) Rancabali (K_{28})	All indicators adequate/safe ($\tilde{X}_1$) There is river sewage pollution there are no waste/sewage recycling activities ($\tilde{X}_4$) All indicators inadequate/unsafe ($\tilde{X}_8$)
8	Cimenyan (K_{11})	All indicators adequate/safe ($\tilde{X}_1$) The majority of household waste disposal sites are inadequate ($\tilde{X}_5$) The majority of household waste disposal sites are inadequate and There is river sewage pollution ($\tilde{X}_6$) All indicators inadequate/unsafe ($\tilde{X}_8$)
9	Cikancung (K_7)	All indicators adequate/safe ($\tilde{X}_1$) There is river sewage pollution ($\tilde{X}_2$) The majority of household waste disposal sites are inadequate ($\tilde{X}_5$)

Groups	Sub-district	Consumption and Production Pattern Characteristics
		The majority of household waste disposal sites are inadequate and there are no waste/sewage recycling activities $(\tilde{X}_7)$
10	Pameungpeuk (K_{24})	All indicators adequate/safe $(\tilde{X}_1)$ The majority of household waste disposal sites are inadequate $(\tilde{X}_5)$ All indicators inadequate/unsafe $(\tilde{X}_8)$
11	Majalaya (K_{19})	There is river sewage pollution $(\tilde{X}_2)$ The majority of household waste disposal sites are inadequate $(\tilde{X}_5)$ The majority of household waste disposal sites are inadequate and There is river sewage pollution $(\tilde{X}_6)$ The majority of household waste disposal sites are inadequate and there are no waste/sewage recycling activities $(\tilde{X}_7)$
12	Bojongsoang (K_4)	There is river sewage pollution $(\tilde{X}_2)$ There is river sewage pollution there are no waste/sewage recycling activities $(\tilde{X}_4)$ The majority of household waste disposal sites are inadequate and There is river sewage pollution $(\tilde{X}_6)$ The majority of household waste disposal sites are inadequate and there are no waste/sewage recycling activities $(\tilde{X}_7)$
	Cangkuang (K_5)	There is river sewage pollution $(\tilde{X}_2)$ There is river sewage pollution there are no waste/sewage recycling activities $(\tilde{X}_4)$ The majority of household waste disposal sites are inadequate and There is river sewage pollution $(\tilde{X}_6)$ The majority of household waste disposal sites are inadequate and there are no waste/sewage recycling activities $(\tilde{X}_7)$
13	Ciparay (K_{12})	There is river sewage pollution $(\tilde{X}_2)$ The majority of household waste disposal sites are inadequate $(\tilde{X}_5)$ The majority of household waste disposal sites are inadequate and there are no waste/sewage recycling activities $(\tilde{X}_7)$
14	Cicalengka (K_6)	There is river sewage pollution $(\tilde{X}_2)$ There are no waste/sewage recycling activities $(\tilde{X}_3)$ The majority of household waste disposal sites are inadequate $(\tilde{X}_5)$ The majority of household waste disposal sites are inadequate and there are no waste/sewage recycling activities $(\tilde{X}_7)$
15	Baleendah (K_2)	There is river sewage pollution $(\tilde{X}_2)$ There are no waste/sewage recycling activities $(\tilde{X}_3)$ The majority of household waste disposal sites are inadequate $(\tilde{X}_5)$ All indicators inadequate/unsafe $(\tilde{X}_8)$
16	Cimaung (K_{10}) Pasirjambu (K_{27}) Soreang (K_{31})	All indicators adequate/safe $(\tilde{X}_1)$ There are no waste/sewage recycling activities $(\tilde{X}_3)$ The majority of household waste disposal sites are inadequate $(\tilde{X}_5)$ The majority of household waste disposal sites are inadequate and There is river sewage pollution $(\tilde{X}_6)$
17	Margaasih (K_{20})	There are no waste/sewage recycling activities $(\tilde{X}_3)$ The majority of household waste disposal sites are inadequate and There is river sewage pollution $(\tilde{X}_6)$ The majority of household waste disposal sites are inadequate and there are no waste/sewage recycling activities $(\tilde{X}_7)$
18	Banjaran (K_3)	All indicators adequate/safe $(\tilde{X}_1)$ There are no waste/sewage recycling activities $(\tilde{X}_3)$ The majority of household waste disposal sites are inadequate $(\tilde{X}_5)$
19	Cilengkrang (K_8) Nagreg (K_{22})	There is river sewage pollution $(\tilde{X}_2)$ There is river sewage pollution there are no waste/sewage recycling activities $(\tilde{X}_4)$ The majority of household waste disposal sites are inadequate and there are no waste/sewage recycling activities $(\tilde{X}_7)$ All indicators inadequate/unsafe $(\tilde{X}_8)$

The principal coordinates from all defuzzification scenarios form the basis of the 2D perceptual map, visualizing the association structure between categories. Figures 1, 2, and 3 display the combined results, where coordinate points are color-coded by method: Black (SCA), Purple (Centroid 'a'), Blue (Centroid 'b'), and Red (Arithmetic Mean).

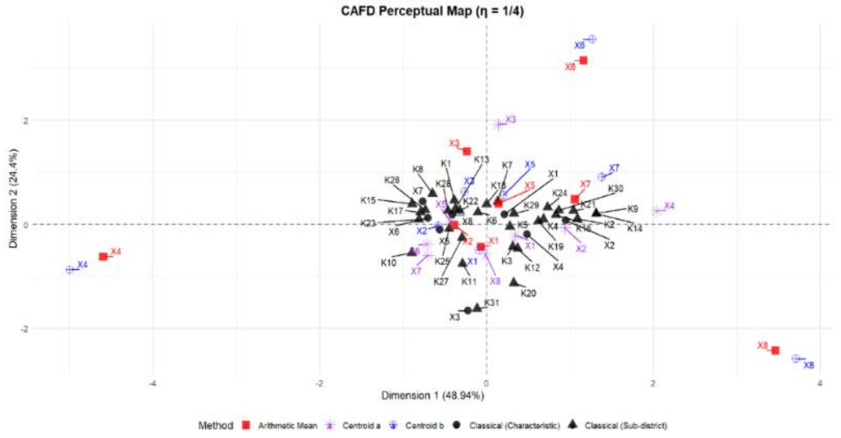

Figure 1. CAFD Perceptual Map with Pessimistic Defuzzification (η = 0.25).

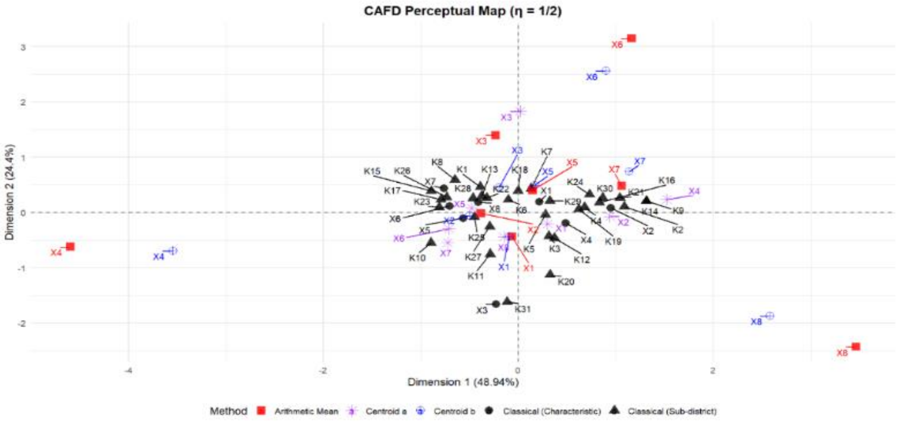

Figure 2. CAFD Perceptual Map with Balanced Defuzzification (η= 0.25).

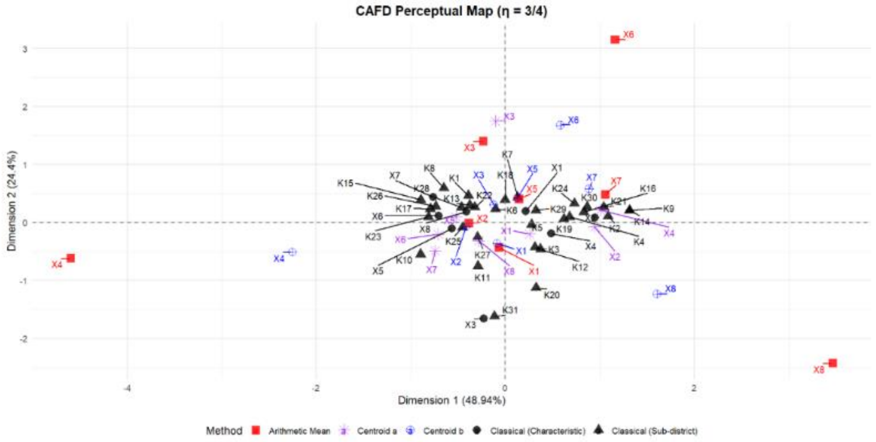

Figure 3. CAFD Perceptual Map with Optimistic Defuzzification (η= 0.75).

3.3 Discussion

The discussion focuses on the interpretation of the 19 sub-district groups and the stability of the framework.

3.3.1 Interpretation of Sub-district Clusters

The fuzzy membership approach unveiled multi-characteristic profiles that standard analysis would overlook. For instance, Group 1 (Arjasari, Pacet, Pangalengan) shows a dominant characteristic of $\tilde{X}_5$ but also has potential in $\tilde{X}_1$. This nuanced view allows the government to apply secondary interventions before a sub-district's status worsens.

3.3.2 Computational Robustness and Policy Implications

The variation of the η parameter in Figures 1, 2, and 3 demonstrates the framework's flexibility. In the pessimistic scenario ($\eta = 0{,}25$), coordinate points appear widely dispersed, making it ideal for identifying the most vulnerable sub-districts, whereas the balanced scenario ($\eta = 0{,}50$) provides general planning projections, and the optimistic scenario ($\eta = 0{,}75$) shifts points toward higher values, reflecting potential regional strengths. By addressing the "uncertainty gap," Stacking-CAFD recognizes varied functional degrees of waste management rather than absolute regional traits. This enables the Bandung Regency Government to transition from generic "one-size-fits-all" policies to localized, cluster-based interventions tailored to the specific profiles identified in Table 12.

3.3.3 Interpretative Framework and Policy Transition

To facilitate the transition from complex fuzzy mathematical computations to practical policy grouping, this framework utilizes three essential interpretative dimensions 4. First, the 2D Perceptual Map (Biplot) Logic visualizes the simultaneous plotting of 31 sub-districts and 8 combined characteristics within the same geometric space, where the proximity or distance between points represents their statistical dependency. Second, the Multi-Characteristic Profile analysis, as evidenced by the fuzzy membership degrees. Third, the Policy Transition Model illustrates the shift from ineffective 'one-size-fits-all' generic policies toward cluster-based interventions.

4. CONCLUSION

This research successfully implemented the Stacking-Correspondence Analysis for Fuzzy Data (Stacking-CAFD) framework as a robust computational approach for analyzing multi-way qualitative survey data, effectively achieving the objective of identifying specific multi-characteristic profiles of 31 sub-districts in Bandung Regency regarding SDG 12. The integration of stacking techniques with fuzzy uncertainty modeling explained a significant 73.35% of the total information variance, leading to the identification of 19 distinct sub-district clusters. These findings provide a critical, data-driven foundation for the Bandung Regency Government to address the low SDG 12 score of 14.53 by transitioning from ineffective 'one-size-fits-all' policies to adaptive, cluster-based interventions, such as urgent infrastructure upgrades for areas with inadequate indicators ($\tilde{X}_8$) and social programs for those lacking recycling initiatives ($\tilde{X}_3$). While the framework is robust, its limitations include a reliance on the Arithmetic Mean defuzzification method and moderate uncertainty parameters, suggesting that future research should explore alternative defuzzification scenarios to further validate the stability of these policy-oriented spatial clusters.

REFERENCES

[1] M. N. Afandi, E. T. Anomsari, and A. Novira, "Sustainable Development Goals (SDGs) Perspective in Regional Development Planning and Implementation," in Proceedings of the 2nd International Conference on Administration Science 2020 (ICAS 2020), vol. 558, 2021, pp. 43–47. doi: 10.2991/assehr.k.210629.009.

[2] Government of Bandung Regency, *Regional Medium-Term Development Plan (RPJMD) of Bandung Regency 2021–2026*. Government of Bandung Regency. (2021). Regional Medium-Term Development Plan (RPJMD) of Bandung Regency 2021–2026. Bandung Regency Government., 2021.

[3] Kemendesa PDTT, "Village Information System (SID): Achievement Scores of Sustainable Development Goals in Bandung Regency," Sustainable Development Goals Desa Database, 2023. [Online]. Available: https://sid.kemendesa.go.id. [Accessed: Jul. 8, 2025].

[4] A. Shostya and N. Kakar, "Toward Sustainable Consumption and Production," in *Fulfilling the Sustainable Development Goals*, London: Routledge, 2021, pp. 274–293. doi: 10.4324/9781003144274-23.

[5] Diskominfo Bandung Regency, "The Regent of Bandung is Continuously Innovating in Waste Management Within the Bandung Regency," Official News Portal of Bandung Regency, 2023. [Online]. Available: https://diskominfo.bandungkab.go.id/#. [Accessed: Jul. 1, 2025].

[6] Diskominfo Bandung Regency, "Emphasizing Collaboration, the Regent Is Optimistic That Issues Ranging From Stunting to Waste Can Be Overcome," Official News Portal of Bandung Regency, 2024. [Online]. Available: https://diskominfo.bandungkab.go.id/#. [Accessed: Jul. 1, 2025].

[7] M. Salomaa, D. Charles, and G. Bosworth, "Universities and innovation strategies in rural (UK)," *Industry and Higher Education*, vol. 37, no. 1, pp. 67–79, Feb. 2023, doi: 10.1177/09504222221096279.

[8] L. Salazar-Llano, M. Rosas-Casals, and M. I. Ortego, "An Exploratory Multivariate Statistical Analysis to Assess Urban Diversity," *Sustainability*, vol. 11, no. 14, p. 3812, Jul. 2019, doi: 10.3390/su11143812.

[9] D. A. Pratiwi and I. Ginanjar, "Characterization method for more than three dimensions based on dependence and correlation using hybrid multiple correspondence analysis," *Communications in Mathematical Biology and Neuroscience*, 2025, doi: 10.28919/cmbn/9034.

[10] N. Sourial *et al.*, "Correspondence analysis is a useful tool to uncover the relationships among categorical variables," *J Clin Epidemiol*, vol. 63, no. 6, pp. 638–646, Jun. 2010, doi: 10.1016/j.jclinepi.2009.08.008.

[11] B. Jefmański, E. Roszkowska, and M. Kusterka-Jefmańska, "Intuitionistic Fuzzy Synthetic Measure on the Basis of Survey Responses and Aggregated Ordinal Data," *Entropy*, vol. 23, no. 12, p. 1636, Dec. 2021, doi:10.3390/e23121636.

[12] E. J. Beh and R. Lombardo, "Multiple and multiway correspondence analysis," *WIREs Computational Statistics*, vol. 11, no. 5, Sep. 2019, doi: 10.1002/wics.1464.

[13] Y. Theodorou, C. Drossos, and P. Alevizos, “Correspondence analysis with fuzzy data: The fuzzy eigenvalue problem,” *Fuzzy Sets Syst*, vol. 158, no. 7, pp. 704–721, Apr. 2007, doi: 10.1016/j.fss.2006.11.011.

[14] L. L. Orr, R. B. Olsen, S. H. Bell, I. Schmid, A. Shivji, and E. A. Stuart, “Using the Results from Rigorous Multisite Evaluations to Inform Local Policy Decisions,” *Journal of Policy Analysis and Management*, vol. 38, no. 4, pp. 978–1003, Sep. 2019, doi: 10.1002/pam.22154.

[15] I. , A. M. A. , & A.-H. R. M. Maulana, “Implementation of sustainable development goals (SDGs) no. 12: Responsible production and consumption by optimizing lemon commodities and community empowerment to reduce household waste,” *ASEAN Journal of Community Service and Education*, vol. 2, no. 2, pp. 141–146, 2023.

[16] F. F. Pinem, “Identifying Characteristics of Slum Settlements in Bandung Regency Sub-districts via Multiple Correspondence Analysis,” *BIAStatistics Journal of Statistics Theory and Aplication*, 2023.

[17] F. Chevenet, S. Doléadec, and D. Chessel, “A fuzzy coding approach for the analysis of long-term ecological data,” Freshw. Biol., vol. 31, no. 3, pp. 295–309, Jun. 1994. doi: 10.1111/j.1365-2427.1994.tb01742.x.

[18] H. Hwang, W. R. Dillon, and Y. Takane, “Fuzzy Cluster Multiple Correspondence Analysis,” Behaviormetrika, vol. 37, no. 2, pp. 111–133, Jul. 2010. doi: 10.2333/bhmk.37.111.

[19] Y. Theodorou, C. Drossos, and P. Alevizos, “Correspondence analysis with fuzzy data: The fuzzy eigenvalue problem,” Fuzzy Sets Syst., vol. 158, no. 7, pp. 704–721, Apr. 2007. doi: 10.1016/j.fss.2006.11.011.

[20] A. Sopelana, S. Urra-Uriarte, I. Oregi, and A. Nacu, "Prioritising Critical Factors for Local Economic Development in Urban Regeneration Strategies," Urban Sci., vol. 9, no. 9, Art. no. 342, 2025. doi: 10.3390/urbansci9090342.

[21] C. J. Merz, “Using Correspondence Analysis to Combine Classifiers,” Mach. Learn., vol. 36, no. 1–2, pp. 33–58, Jul. 1999. doi: 10.1023/A:1007559205422.

[22] A. Agresti, *An Introduction to Categorical Data Analysis Second Edition*, Second. John Wiley & Sons, 2007.

[23] E. J. Beh and R. Lombardo, *Correspondence Analysis Theory, Practice and New Strategies*. John Wiley & Sons, 2014.

[24] A. Neugut and T. Fojo, "The statistical significance revolution," JNCI Cancer Spectr., vol. 8, no. 4, Art. no. pkae035, Aug. 2024. doi: 10.1093/jncics/pkae035.

[25] I. Ginanjar, N. Sunengsih, and Sudartianto, “Contribution of a categorical statistical test in examining dependencies among qualitative variables,” *J Phys Conf Ser*, vol. 1265, Jul. 2019, doi: 10.1088/1742-6596/1265/1/012022.

[26] V. Choulakian, "Quantification of intrinsic quality of a principal dimension in correspondence analysis and taxicab correspondence analysis," Commun. Stat. Simul. Comput., vol. 50, no. 12, pp. 4111–4122, 2021. doi: 10.1080/03610918.2019.1639735.

[27] G. Li, S. Lu, H. Zhang, and S. Lo, “Correspondence Analysis on Exploring the Association between Fire Causes and Influence Factors,” Procedia Eng., vol. 62, pp. 581–591, 2013. doi: 10.1016/j.proeng.2013.08.103.

[28] T. S. Nauli, I. Ginanjar, and D. Y. Faidah, “Dependency-based clustering using correspondence and Ward’s hierarchical analysis: a case study on clean water and sanitation indicators in West Java’s cities and regencies,” Commun. Math. Biol. Neurosci., vol. 2025, Art. no. 9235, 2025. doi: 10.28919/cmbn/9235.

[29] E. J. Beh and R. Lombardo, "Multiple and multiway correspondence analysis," WIREs Comput. Stat., vol. 11, no. 5, Art. no. e1464, Sep./Oct. 2019. doi: 10.1002/wics.1464.

[30] K. E. Lestari, M. R. Utami, and M. R. Yudhanegara, “Simple Algorithm to Construct Circular Confidence Regions in Correspondence Analysis Using R,” *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, vol. 16, no. 1, pp. 065–074, Mar. 2022, doi: 10.30598/barekengvol16iss1pp065-074.

[31] J. Verstraete, W. Radziszewska, K. Kaczmarek-Majer, and S. Bykuć, “Combined Defuzzification Under Shared Constraint,” *IEEE Transactions on Fuzzy Systems*, vol. 32, no. 5, pp. 3049–3058, May 2024, doi: 10.1109/TFUZZ.2024.3367008.

[32] A.C. Rencher, *Method of Multivariate Analysis*. Wiley, 2002.

[33] M. W. Warolemba, Resmawan, and D. R. Isa, " Cluster Fuzzy C-Means Analysis and Discrimination for Collection of People's Welfare Data," Jurnal Sainsmat, vol. XII, no. 2, pp. 141-152, Sep. 2023, URL: http://ojs.unm.ac.id/index.php/sainsmat

[34] A. Li, C. Fan, F. Xiao, and Z. Chen, "Distance measures in building informatics: An in-depth assessment through typical tasks in building energy management," Energy Build., vol. 256, Art. no. 111817, Feb. 2022. doi: 10.1016/j.enbuild.2021.111817.